



我国城市公共服务提供绩效评价

——面向 L 市的探索性研究

尚虎平 陈星宇

摘要: 城市公共服务绩效从内容结构上来说包括人民满意与物质提高两个方面。针对这两个方面,利用单轮德尔菲法、隶属度计算、变差系数、相关分析在 L 市进行实证筛选,产生了 56 个绩效评估指标,进而采用该指标体系对 L 市进行了实际评估。结果显示,整体绩效有所提升,但与标杆城市相比,仅达到其值的 13.122%,且与民生关系密切的公共服务绩效比历史平均数据还差。据此推断,我国其它城市也可能存在着民生服务项目绩效恶化、公共服务在“捏软柿子”、在为讨好上级而“服务”的问题。因此建议民生类公共服务可以采取特别项目、绩效预算拨款、公开招标制度,甚至可以将连续两年绩效不佳的民生项目对应的税收由地方税划归中央税。

关键词: 城市公共服务; 人民满意; 物质提高; 绩效评价

从 2002 年党的十六大第一次将提供“公共服务”明确界定为我国政府的四大职能开始,全面加强公共服务职能,追求公共服务的高效率和低成本,已经成为全国各界的共识,对公共服务提供进行常规性的绩效评估也逐渐被人们所接受。尽管公共服务绩效评价成为了热点,但目前还缺乏科学、合用的指标体系,缺乏基于科学指标体系展开、立足于主客观数据进行的实证研究。本文正是基于这种认识,力图以实证研究的方式来弥补此种缺憾。

一、公共服务绩效评价的逻辑

正如 Boyne(2003)、Ancarani(2009)所发现,传统的公共服务绩效评估都过分强调了产出的数量、质量和资金使用价值,却忽视了顾客的满意等“软”问题^{①②}。针对这种不足,丁元竹提出基本公共服务是政府的基本责任,客观的基本公共服务绩效和主观的基本公共服务绩效共同构成科学意义上的公共服务绩效,两者缺一不可^③。厦门大学陈振明课题组在推进国家自然科学基金重点项目“公共服务提供机制与方式研究”的过程中提出了公共服务绩效的“二元模型”,该模型与丁元竹所提的基础公共服务的内涵基本一致,也强调了公共服务既具有人民主观满意的属性,也具有客观物质性数据提高的属性。总体而言,人民满意主要是指人民对公共服务提供的数量、方式、质量以及提供者工作态度、作风的满意;物质性数据提高主要指社会保障、医疗卫生、公共教育、公共安全、公共信息等领域承载了公共服务功能的客观物质数量的提升。笔者认为这些研究已经抓住了公共服务

① G. A. Boyne. "Sources of Public Service Improvement: A Critical Review and Research Agenda", *Journal of Public Administration Research and Theory* 2003, 13(3), p. 367~394.

② A. Ancarani. "Supplier Evaluation in Local Public Services: Application of A Model of Value for Customer", *Journal of Purchasing and Supply Management* 2009, 15(1), pp. 33~42

③ 丁元竹,《界定基本公共服务及其绩效》,载《国家行政学院学报》2009 年第 2 期,第 18~20 页。

的“牛鼻子”,对公共服务绩效评价必须遵循既要把握主观满意的绩效、也要测量客观数据的提高的内在逻辑,这是一种完整的评估,它体现了刚柔并济的原则,既保证了对物的重视,也保证了对人的尊重,避免了传统评价中见物不见人的弊端。具体而言,基本逻辑如图1所示。

图1展示了公共服务绩效评估的基本逻辑实际上是实现公共服务的两个构成要素(人民满意、物质提高)的评判过程,它涉及如何将这两类要素以可操作化、可获得证据数据的方式进行评估。在分别评估了人民满意和物质提高的基础上,就可以将这二者按照特定的比例(权重)合成为统一的公共服务绩效结果,共同构成了公共服务绩效的“阴”、“阳”两极^①。具体而言,在评估人民满意过程中,由于“满意”、“不满意”、“比较满意”等均属于主观判断,需要设计出合适的评估方法才能挖掘到较为接近真实情况的评价结果,这就是图1的“评估操作化”过程。目前来看,人民满意度的可操作性评价一般都采取评判公民满意度的方式,这方面的技术已经比较成熟,主要通过采用已有的或者编制新的公民满意度问卷进行大样本调研,并以大样本数据来推断全集总体的(人民)满意情况。由于大样本调研存在着“从众行为”,且由于问卷的结构化限制,常常不能了解个体的深层次想法,为了解决此问题,问卷调研一般都要以深度访谈作为辅助手段,两种工具同时派上用场。评估公共服务绩效的“物质提高”方面则容易得多,因为所有的“物质提高”都具有有案可稽的性质,我们可以在经济统计年鉴、社会统计年鉴、公用事业统计年鉴、政府公文、政府档案等物质性载体中查找到它的蛛丝马迹。

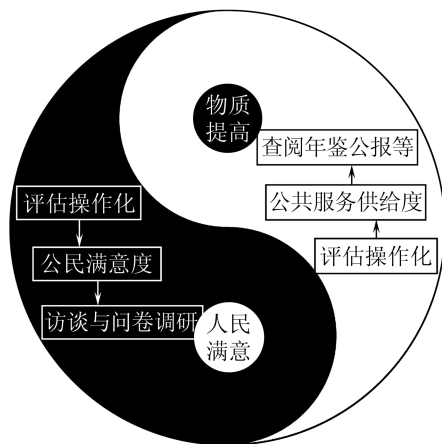


图1 公共服务绩效评估逻辑

等物质性载体中查找到它的蛛丝马迹。

二、我国城市公共服务绩效评估指标构建

城市公共服务绩效是一个度量城市公共服务提供的目标性词汇,从内容结构上包括人民满意与物质提高两个方面。要评价这两方面内容,就需要科学、合理地反映这两方面现实的指标体系,它实际上是个跨学科测量问题,既涉及经济学、管理学所强调的投入产出和效率问题,也涉及心理学、社会学所要求的人民满意、心理满足问题,还涉及政治学、法学、公共行政学所要求的权力合法运行、人民权利正确保障问题,对于这些问题的正确评估需要从这些不同的学科视角提供不同的指标池。

(一)确定指标池(Indicator pool)

按照城市公共服务绩效的内容结构与跨学科特点,本研究在“人民满意”与“物质提高”两个领域分别设置了“印象满意”、“政府效能”和“企业经营环境”、“基本公共服务”4个维度,并在这些维度中基于已有的各类公共服务绩效指标体系、公共效能评价指标体系、政府绩效评估指标体系遴选了99个横跨经济学、管理学、社会学、心理学、政治学、公共行政学、法学等学科的具体指标作为指标筛选的“指标池”^②。

(二)指标池筛选

本文在筛选中利用单轮德尔菲法结合隶属度计算的方式来消除噪音。在单轮德尔菲的专家选择方面,我们采取了向全国在公共服务研究领域具有较深造诣专家发送Email的方式发送指标池,请他们对自己认为属于该评估维度的指标进行圈选,在Email发出后以电话提醒帮忙填答。本研究按照理论专

^①按照太极原理,阴与阳的重要性是相当的,其比例必然为1:1(权重为50%;50%),但公共服务绩效毕竟不是纯理论意义上的太极,它必须反映特定时期公共服务发展的侧重点。如果正处于物资短缺时期,则可能“物质提高”部分的比例要大一些,而如果处于物质丰富人民幸福感薄弱阶段,则可能“人民满意”的权重要高一些。

^②由于指标池是一张非常长的表格,此处为了行文简洁,略去此表,在后文我们会呈现筛选确定的指标表,如果同行需要查询指标池,可与作者本人联系(zhouyafu8210@sina.com)。

家、实践专家、其他类专家共 3 个口径各选择了 30 人(如表 1)。“其他类专家”是指“理论专家”和“实践专家”之外,对于公共服务领域颇为了解或者颇有研究的群体,比如专门关注公共服务的记者、第三部门志愿者、学生等。

表 1 单轮德尔菲专家组成

来源地区 \ 分类	理论专家		实践专家		其他类专家		合计	
	发送	回收	发送	回收	发送	回收	发送	回收
东北	2	2	3	3	3	3	8	8
华北	8	7	6	6	6	5	20	18
中南	3	3	3	3	3	3	9	9
华东	8	6	6	6	6	4	20	16
华南	4	4	6	6	6	6	16	16
西北	3	3	3	3	3	3	9	9
西南	2	3	3	3	3	2	8	7
合计	30	27	30	30	30	26	90	82

第一步,专家临界值确定。专家临界值下线的计算公式为 $M = \mu + \frac{S}{\sqrt{\sum_{x=1}^n N}} t_a (x \in R^+)$, 其中, μ 表

示欲选择专家的期望数量; S 表示专家对备选问题选择数量的标准差; N 代表备选问题被专家选中的频数; t_a 是置信度 α 取特定值时的 t 检验值, 它可以通过查阅 t 检验表获得。最后计算出临界专家数量为

$$M = 30 + \frac{19.77}{\sqrt{4118}} \times 2.36 \approx 31.$$

第二步,隶属度确定。依照模糊数学的观点,城市公共服务绩效评估指标体系(I)本身就是一个模糊集合,每个指标就是集合中的一个元素,对每个指标进行隶属度分析即可确定出哪些指标在合格线之上。隶属度可以表示为 $R_n = \frac{M_n}{T} (n = 1, 2, \dots, J)$, 其中 R_n 为隶属度, J 为指标的数量, T 为反馈消息的专家总数。通过对反馈回来的 82 份有效问卷的统计分析,分别得到了 99 个指标池中指标的隶属度,并将其与临界值进行了对比。由于临界专家数量为 31, 临界值便为 $R = \frac{31}{82} = 0.378$, 这样,“事业单位满意度”、“市民感知的政府服务态度”等隶属度小于 0.378 的 23 项指标被剔除, 剩下了 76 个指标。

第三步,指标的鉴别力判断。指标的鉴别力涉及两方面的问题:其一是指标之间是否相关性过强,以至于一个或者多个指标实际上都指向同一个问题;其二是指标是否正确表达了所要评估的标的物。对于第一个问题,主要是通过通过对各个评价指标间的相关分析,删除一些相关系数比较大的评估指标。而对第二个问题,比较常用的方式是借助变差系数来消除这个问题。

按照统计学原理,相关系数一般可分为 3 级: $|R| < 0.4$ 为低度线性相关, $0.4 \leq |R| < 0.7$ 为显著性相关, $0.7 \leq |R| < 1$ 为高度线性相关^①。本研究按照这种分类,将所有两两相关中 $|R| \geq 0.4$ 的一项指标都进行了删除。结果,“交通警察处理小案速度”、“公民参与听证的频率”、“固定、移动与网络通讯的平均价格”、“年底平均房价”等 14 个不合格指标被删除,如表 2。

表 2 “印象满意”维度下指标的相关性

指标(标志)	市民满意度 i_1	第三部门满意度 i_2	企业主满意度 i_4	隶属度
市民满意度 i_1	1.00			0.817
第三部门满意度 i_2	0.334***	1.00		0.537
企业主满意度 i_4	0.213	0.199	1.00	0.671

说明:①***表示显著性水平为 0.01,其它均为 0.000;②标志之所以为 i_1, i_2, i_4 , 是因为在隶属度筛选中已经删除了 i_3 。

①《相关系数(Correlation coefficient)》,载 <http://wiki.mbalib.com/wiki/%E7%9B%B8%E5%85%B3%E7%B3%BB%E6%95%B0>。

变差系数计算一般式为 $V_x = \frac{S_x}{\bar{I}_x} \times 100\%$, 其中, $\bar{I}_x = \frac{1}{n} \sum_{x=1}^n I_x$ 是指标的均值; $S_x = \sqrt{\frac{1}{n-1} (I_x - \bar{I}_x)^2}$

是指标的标准差。一般而言, 一个指标的变差系数越大, 则该指标的鉴别力越强, 反之亦然。按照 Foster、Greer、Thorbecke(1984)^①、章炼(1992)^②等的研究, 如果测量指标都为正值, 只要不同的指标总体在 $\bar{I}_x \pm 3S_x$ 的范围内, 就可以形成正态分布, 其变异系数也达到了最大。从本研究的数据来看, $\bar{I}_x \pm 3S_x \leq 0.590$ 是变差系数的临界下线, 也就是本研究删除指标的依据。经过 SPSS18.0 软件分析, 我们删掉了“城市地税办税厅税务登记速度(0.52)”、“企业家对政府官员的信任度(0.44)”、“在岗职工社会平均工资(0.36)”等 6 个变差系数小于 0.590 的指标, 最终形成了我们的指标体系(表 3)。

三、L 市公共服务绩效实证评估

在获得指标体系之后, 本研究采取了一种不同于传统绩效评估价的方式来具体实施评估。首先按照“人民满意”与“物质提高”的评价域将需要评价的城市绩效分类展开; 其次, 将数据分为软数据与硬数据, 软数据对应“人民满意”, 硬数据对应“物质提高”; 第三, 将传统上单独进行的指标赋权过程与问卷调查、硬数据获取过程有机结合起来; 第四, 结合测量平差理论的已有成果, 利用标准差倒数来确定权重并在此基础上开发出绩效评估的一般模型, 进而评估出 L 市公共服务的整体绩效^③; 最后, 为了判断该城市公共服务绩效的现状, 我们还采取了与特定标杆值逐项、总体比较的方式。

(一) 数据的来源与搜集

评估指标数据采集是城市公共服务绩效评价实证阶段最为关键的环节之一, 本研究的数据采集包括软数据和硬数据的采集两部分。对于难以通过查阅文献记录获得的软数据, 需要通过问卷调查、座谈会等实证调查方法采集。我们编制了面向普通市民、企业、第三部门的问卷^④, 在小范围试调研的基础上完善了该问卷, 然后在 L 市发放了 1000 份针对市民的问卷、200 份针对企业主的问卷和 50 份针对第三部门的问卷。硬数据的搜集主要通过查询政府公布的权威资料得来。本研究主要通过查阅《2009 年中国统计年鉴》、《G 省年鉴 2008/2009》、《L 市政府工作报告 2008/2009》以及 G 省、L 市政府门户网站的数据获得。

(二) 评价方法和评估一般式确定

笔者采用测量平差理论中利用标准差来客观赋权的方式(喻国荣, 2007^⑤; 王中宇等, 2008^⑥; 王穗辉, 2010^⑦), 利用问卷获得的实际数据、各类统计年鉴、行政记录等实在性数据对方差进行分析, 然后按照测评理论中方差越小、其权越大的性质将赋权与评价结果结合进行。由于公共服务绩效评估涉及到的对象广泛, 本文采用标准化方式实行无量纲化。其过程如下:

设 \bar{I} 为一列单位不一的指标实在值数据 I 的均值, σ 为其标准差, 那么对这列指标实在值 I 做无量纲化处理(标准化处理), 求其无量纲值 I' 的公式便可以记为式(1):

$$I' = \frac{I - \bar{I}}{\sigma} \quad (1)$$

在解决了无量纲问题之后, 就可以确定出最终的绩效评估一般式(2)了。本研究中最终的绩效评估结果都采用式(2)作为绩效评估的一般式。

① J. Foster, J. Greer, E. Thorbecke. "A Class of Decomposable Poverty Measures", *Econometrica* 1984, 52(3), pp. 761~766.

② 章 炼:《“变差系数”可比的前提条件及公式修正》,载《上海教育科研》1992 年第 2 期,第 21、63 页。

③ L 市是 G 省省会, 政治、经济、文化中心, 在西北地区处于“座中四联”位置, 是黄河上游经济区重要的经济中心和西陇海兰新经济带重要的支撑点和辐射源, 是我国东中部地区联系西部地区的桥梁和纽带, 是大西北的交通通信枢纽。2010 年, L 市全年完成生产总值 1100.39 亿元, 同比增长 13.7%, 人均 GDP 达 5162 美元, 全年完成社会消费品零售总额 545.11 亿元, 同比增长 18.61%。

④ 如要查阅问卷, 请与作者本人联系。

⑤ 喻国荣:《论测量平差中的权和权阵》,载《测绘通报》2007 年第 7 期,第 39~42 页。

⑥ 王中宇、刘智敏、夏新涛、祝连庆:《测量误差与不确定度评定》,科学出版社 2008 年,第 66~77 页。

⑦ 王穗辉:《误差理论与测量平差》,同济大学出版社 2010 年,第 23~27 页。

表3 经过相关度分析删除后的指标体系

域	维度	指标	标志	单位	变差系数
人民满意	印象满意	市民满意度	i_1	%	0.922
		第三部门满意度	i_2	%	0.727
		企业主满意度	i_4	%	0.692
	政府效能	一站式服务中心办结速度	i_6	分钟	1.903
		企业感知的政府服务态度	i_{10}	/	0.777
		市民感知的政府廉洁程度	i_{12}	/	0.991
		企业感知的政府廉洁程度	i_{13}	/	0.826
		政策信息依法公开程度	i_{16}	%	0.625
企业经营环境	企业家信心指数	i_{17}	点	0.638	
	工业用地平均价格	i_{21}	万元/亩	1.211	
	工商业用电平均价格	i_{22}	元/千瓦时	1.092	
	厂房租金增长率	i_{26}	%	0.598	
	企业成立手续办结平均实际时间	i_{29}	天	1.464	
	政府网上服务项目中可网上办理的项目比例	i_{30}	%	1.377	
	政府对企业的回应率	i_{34}	%	0.923	
	工商提醒告知服务制度化情况	i_{35}	/	0.699	
	税务提醒告知服务制度化情况	i_{36}	/	0.686	
	工作适龄人群非意愿性失业率	i_{37}	%	2.00	
	失业人员经培训后重新就业率	i_{39}	%	1.343	
	登记失业人员再就业率	i_{40}	%	1.269	
	住房服务	城镇居民人均居住面积	i_{43}	平方米	1.400
		城镇居民无房产比例	i_{44}	%	2.017
		城镇居民3代人居住一套住房的比例	i_{45}	%	1.939
		居民月平均工资性收入与平均房价比	i_{47}	%	1.303
	公共安全	年度刑事案件递增率	i_{50}	%	0.884
		年度治安案件递增率	i_{51}	%	0.792
		刑事案件破案率	i_{52}	%	1.001
治安案件查处率		i_{53}	%	0.887	
公共教育	教育经费占地方财政支出比重	i_{54}	%	1.004	
	教育经费足额拨款的比率	i_{55}	%	1.255	
	偏远区、县重点中学占全市重点中学的比率	i_{57}	%	0.878	
	适龄儿童入学(园)率	i_{58}	%	1.232	
医疗卫生	每万人专业技术人员数	i_{60}	人	0.761	
	每万人医疗机构床位数	i_{62}	张	0.692	
	患传染病者未治愈率	i_{65}	%	0.593	
环境保护	绿化覆盖率	i_{66}	%	0.664	
	每日噪音平均数	i_{67}	分贝	0.605	
	一年中空气污染指数为良好(或二级)的天数	i_{68}	天	0.740	
	垃圾无害化处理率	i_{70}	%	0.786	
	污水处理率	i_{72}	%	0.808	
社会保障	抚恤和社会福利救济费占财政支出比重	i_{73}	%	0.775	
	失去就业能力者未被救助的比例	i_{74}	%	0.932	
	社会保险参保人次与常住人口比例	i_{75}	%	0.817	
基础设施	水质综合达标率	i_{78}	%	0.621	
	燃气普及率	i_{79}	%	0.596	
	每平方公里公厕数量	i_{84}	个	0.645	
	人均城市道路面积	i_{85}	万平方米	0.688	
	道路维护、修建经费增长速度	i_{86}	%	0.873	
公共交通	公交补贴占财政支出比重	i_{89}	%	0.700	
	公交经费年增长速度	i_{90}	%	1.022	
	公交线路的平均等候时间	i_{91}	分钟	1.676	
	新型公共交通工具占所有工具的运营量比	i_{92}	%	0.619	
文体休闲	万人图书馆图书藏量	i_{94}	册	0.601	
	人均休闲广场的占有量	i_{95}	平方米	0.674	
	万人公园景点面积	i_{97}	公顷/万人	0.873	
	每百个社区拥有的休闲广场数量	i_{99}	个	0.920	

$$\begin{aligned}
 P_j &= \frac{I'_1}{\sigma_1^2} + \frac{I'_2}{\sigma_2^2} + \dots + \frac{I'_{n-1}}{\sigma_{n-1}^2} + \frac{I'_n}{\sigma_n^2} \\
 &= \frac{I_1 - \bar{I}_1}{\sigma_1^3} + \frac{I_2 - \bar{I}_2}{\sigma_2^3} + \dots + \frac{I_{n-1} - \bar{I}_{n-1} - \bar{I}_{n-1}}{\sigma_{n-1}^3} + \frac{I_n - \bar{I}_n}{\sigma_n^3} \\
 &= \frac{1}{\sigma_n^3} \sum_{j=1}^n I_j - I' \quad (j=1, 2, \dots, n)
 \end{aligned} \tag{2}$$

(三) 评估实施

在计算总体绩效的过程中,我们主要以(1)、(2)式为依据,这两个公式的展开都依赖于数据真值、数据均值、无量纲值,笔者利用 SPSS18.0 将这些值分项统计,其结果如表 4。在获得了历史均值、无量纲值之后,就可以利用(2)式来计算每个指标的绩效值和总体绩效值了,这些结果也都整合在表 4 中。

然而,仅仅获得 L 市的绩效值我们还很难弄清其政策含义,具体来说就是还很难依据它来提出改进 L 市公共服务绩效的措施,因为还欠缺可资比较的标杆。为了解决“有比较才有鉴别”的问题,本研究还引入了一个“理想值”变量作为标尺来判断 L 市公共服务绩效的好坏。理想值的产生基于 4 种逻辑:其一是根据理想化状态,比如人民对企业的满意问卷的最高值为 6,我们便将该指标的理想值设置为 6;其二是关于硬数据中的比率,如“燃气普及率”,我们按照理想状态将其设置为 100%;其三是硬数据中的比较性理想值,这类指标既没有理想化值,也不可能具有比率的最大值 100%,本研究选取了与 L 市同属于一个行政区,但发展势头更好、各方面情况均比 L 市更为理想的 X 市的值作为参照的“理想值”;其四是针对国家出台了明确标准的硬数据标准值如“每平方公里公共厕所数量为 3 个”^①,这种本身属于国家层面确定的理想值,我们直接将其引用过来。在理想值的基础上,我们将之进行了无量纲化,产生了无量纲值,同时也计算了理想值的均值,最终也计算出了总体理想绩效值,处于表 4 的右半列。表 4 的“标志”是我们对筛选出的 56 个指标重新编号后的排序。

表 4 L 市公共服务提供总体绩效评价结果

标志	数据真值	数据均值	无量纲值	总体绩效	标杆值	无量纲标杆值	总体标杆绩效值
i1	2.77	2.72	0.136612	0.373257	6	8.961748634	24.48565201
i2	2.69	2.66	0.081967	0.223954	6	9.12568306	24.93356027
i3	2.91	2.93	-0.05464	-0.1493	6	8.387978142	22.91797307
i4	2.3	2.2	0.273224	0.746514	2.2	0	0
i5	2.1	2.22	-0.32787	-0.89582	6	10.32786885	28.21822091
i6	2.3	2.26	0.10929	0.298606	6	10.21857923	27.9196154
i7	2.2	2.17	0.081967	0.223954	6	10.46448087	28.5914778
i8	2.5	2.16	0.846995	2.314193	6	10.49180328	28.66612918
i9	126.61	122.74	10.57377	28.89008	147.14	66.66666667	182.1493625
i10	15	13	5.464481	14.93028	10	-8.19672131	-22.39541342
i11	0.863	0.663	0.546448	1.493028	0.85	0.510928962	1.39598077
i12	6.00	5.00	2.73224	7.465138	5.24	0.655737705	1.791633074
i13	16	10	16.39344	44.79083	5	-13.6612022	-37.32568903
i14	16.00	14.558	3.939891	10.76473	70	151.4808743	413.8821703
i15	2.98	2.78	0.546448	1.493028	6	8.797814208	24.03774374
i16	53.00	51.00	5.464481	14.93028	100	133.8797814	365.7917525
i17	40	44	-10.929	-29.8606	100	153.0054645	418.0477172
i18	6	2.8	8.743169	23.88844	5	6.010928962	16.42330317
i19	76.00	73.80	6.010929	16.4233	100	71.58469945	195.5866105
i20	53.00	56.00	-8.19672	-22.3954	100	120.2185792	328.4660635
i21	15	16.7	-4.64481	-12.6907	26.4	26.50273224	72.41183672
i22	41.00	52.89	-32.4863	-88.7605	0	-144.508197	-394.8311386
i23	29.00	27.00	5.464481	14.93028	0	-73.7704918	-201.5587208
i24	25.5	42.49666667	-46.439	-126.882	40.82	-4.58105647	-12.51654772
i25	15.20	14.80	1.092896	2.986055	0	-40.4371585	-110.4840395

①《中国十二五规划 GDP 年均增长 7%, 登记失业率 5%》, 载 <http://finance.qq.com/a/20110307/003642.htm>。

标志	数据真值	数据均值	无量纲值	总体绩效	标杆值	无量纲标杆值	总体标杆绩效值
i26	16.30	15.90	1.092896	2.986055	0	-43.442623	-118.6956911
i27	44.20	41.90	6.284153	17.16982	45	8.469945355	23.1419272
i28	82.10	99.70	-48.0874	-131.386	92	-21.0382514	-57.48156111
i29	21.00	20.00	2.73224	7.465138	4	-43.715847	-119.4422049
i30	99.00	98.00	2.73224	7.465138	100	5.464480874	14.93027561
i31	16.00	14.50	4.098361	11.19771	19.05	12.43169399	33.96637702
i32	98.00	99.80	-4.91803	-13.4372	100	0.546448087	1.493027561
i33	79	80.6	-4.37158	-11.9442	77.37	-8.82513661	-24.11239511
i34	54	51.9	5.737705	15.67679	41.3	-28.9617486	-79.13046075
i35	38	37.1	2.459016	6.718624	0	-101.36612	-276.9566126
i36	23.16	21.17	5.437158	14.85562	44.99	65.08196721	177.8195825
i37	59	58	2.73224	7.465138	55.2	-7.65027322	-20.90238586
i38	250	268	-49.1803	-134.372	301	90.16393443	246.3495476
i39	70.00	75.00	-13.6612	-37.3257	100	68.30601093	186.6284452
i40	62	60.96	2.84153	7.763743	100	106.6666667	291.43898
i41	3	2.66	0.928962	2.538147	3.5	-24.4535519	-66.81298337
i42	16	13.7	6.284153	17.16982	0	-37.431694	-102.2723879
i43	24.27	20.73	9.672131	26.42659	100	216.5846995	591.7614739
i44	100	100	0	0	100	0	0
i45	72.00	0.68	194.8634	532.4136	100	271.3661202	741.4374869
i46	0.62	0.52	0.273224	0.746514	3	6.775956284	18.51354176
i47	0.000865	0.000765	0.000273	0.000747	0.00148	0.001953552	0.005337574
i48	5.00	4.00	2.73224	7.465138	10	16.39344262	44.79082684
i49	1.21	1.17	0.10929	0.298606	2.76	-254.754098	-696.0494491
i50	6.66	6.36	0.819672	2.239541	10	9.945355191	27.17310162
i51	20	16	10.92896	29.86055	10	-16.3934426	-44.79082684
i52	2.30	1.50	2.185792	5.97211	20	50.54644809	138.1050494
i53	1360	1359	2.73224	7.465138	135.9	0	0
i54	0.0004	0.0003	0.000273	0.000747	0.001	0.001912568	0.005225596
i55	0.6	0.4	0.546448	1.493028	1	1.639344262	4.479082684
i56	0.013	0.01	0.008197	0.022395	0.01	0	0
总体	—	—	—	309.9416	—	—	2361.997584

在总体绩效评估结果中之所以出现了 0 值与负值,是因为我们假定,不管哪个城市都应该努力改善人民满意和物质提高的状态,实现“市富民乐”的施政追求,如果该市的绩效不能达到历史数据的均值状态,我们就认为它们在当年是缺乏绩效的,甚至是负绩效的。需要指出的是,这种逻辑是“第三只看政府”的第三方评价属性,其主要目的在于为政府诊断问题,引起它们的警醒,并非要全盘否定它们的工作,在这里负值只具备咨询价值和参考价值。尽管这不太符合常理,因为政府忙碌一年不可能一事无成,但我们认为以数量测评的直观结果来看,应当就是这种反映方式。

(四) L 市公共服务提供绩效概况

总结表 4 来说,L 市的公共服务绩效整体状况可以概括为 4 个方面:

1. 市民满意度、第三部门满意度、一站式服务中心办结速度等 43 个指标的绩效现状相对历史平均数据有了不同程度的提升,绩效整体状况呈现上升势头。从评价结果来看,市民满意(0.373257)、第三部门满意度(0.223954)、一站式服务中心办结速度(0.746514)等 43 个指标的绩效值均大于 0,这说明在 56 个指标中,绝大多数指标(76.8%)所代表的公共服务绩效在提升,这也可以看作是我国政府持续改革的一个可喜成果。

2. 与标杆城市、理想值相比,L 市的总体绩效仍然相对落后,改进余地尚大。尽管有 43 个指标对应的公共服务绩效在持续提升之中,但从与标杆值的比较结果来看,几乎每一个指标的绩效值都与理想值有差距,甚至较大的差距,这使得 L 市最终的整体绩效值(309.9416)仅仅为理想绩效总值(2361.997584)的 13.122%,这充分显示了 L 市的单项绩效、总体绩效还有巨大的改进余地。

3. 企业主满意度、企业感知的政府服务态度、税务提醒告知服务制度化等 12 项与民生关系密切的

指标的绩效现状甚至比历史平均数据还糟糕。从评价结果可以看出,企业主满意度(-0.1493)、企业感知的政府服务态度(-0.89582)、税务提醒告知服务制度化情况(-29.8606)、登记失业人员再就业率(-22.3954)、城镇居民人均居住面积(-12.6907)、城镇居民无房户比例(-88.7605)、居民月平均工资性收入与平均房价比(-126.882)、治安案件查处率(-131.386)、适龄儿童入学(园)率(-13.4372)、每万人专业卫生技术人员数(-11.9442)、一年中空气污染指数为良好(或二级)的天数(-134.372)、垃圾无害化处理率(-37.3257)共12项指标的绩效值甚至出现了负值,其意涵是相对于历史平均数据,该年L市这些方面不仅仅没有出现明显改善,还有恶化的势头。进一步细究这些指标,我们还会发现,这些指标中很多都属于与民生关系密切、甚至人民反响很大的社会热点问题,比如住房、就业、治安,说明L市与其它城市一样,在这些方面的公共服务还很不到位,甚至出现了下降的势头。

4. 水质综合达标率相对历史平均数据没有出现绩效改进。该项指标的绩效分值为0,代表该项指标相对于历史数据并未出现提升,说明该项公共服务并未取得应得的绩效。

四、推论与建议

表4的评价结果实际上还蕴含了一个延伸性的结论,它对于评价、改善我国城市公共服务可能更具有普遍意义,我们认为这是本研究的一种推论,它来自于我们对理想标杆绩效值的分析。在标杆值中,竟然出现了0值、负值,说明我们选择的本行政区的“领头羊”城市与其历史数据相比较,公共服务绩效也存在着这样或那样的不足,如果推而广之,其他行政区的一般城市、“领头羊”城市可能也存在类似问题。

推论1:与民生关系密切的评价指标绩效倾向于产生低绩效,甚至0绩效、负绩效。从统计结果来看,标杆值的工业用地平均价格(-22.39541342)、企业成立手续办结平均实际时间(-37.32568903)、城镇居民无房户比例(-394.8311386)、城镇居民3代人居住一套住房的比例(-201.5587208)、居民月平均工资性收入与平均房价比(-12.51654772)、年度刑事案件递增率(-110.4840395)、年度治安案件递增率(-118.6956911)、治安案件查处率(-57.4815611)、教育经费占地方财政支出比重(-119.4422049)、每万人专业卫生技术人员数(-24.11239511)、每万人医疗机构床位数(-79.13046075)、患传染病者未治愈率(-276.9566126)、每日噪音平均数(-20.90238586)、抚恤和社会福利救济费占财政支出比重(-66.81298337)、社会保险参保人次与常住人口比例(-102.2723879)共15个指标的绩效值出现了负值,这说明即使作为该地区领头羊的X市在这些指标上取得的绩效结果也较差,甚至出现了低于历史平均数据的情况。与L市一样,X市的这些负绩效指标也几乎都是与民生关系密切的公共服务事项,这无疑透露了一个具有普遍性意义的现实:在目前公共服务提供中,城市中普遍存在着忽视民生指标的问题,这大抵也是各个城市因为住房、拆迁、治安等问题频频引起自焚、上访等热点事件的原因。

推论2:公共服务总体绩效提升更多采取了指标博弈的策略,即“捏软柿子”,“做更容易做的事情”,“不做对的,只做有分数的”。无论从L市还是标杆城市X来说,令人欣慰的是每个城市的公共服务总体绩效都在逐渐提升,这也符合我国各地的整体发展势头。然而我们应该看到,每个城市公共服务绩效得分更多来自于那些容易完成的指标,而对于像“城镇居民无房户比例”、“城镇居民3代人居住一套住房的比例”、“年度治安案件递增率”这种人民反映强烈、希冀有较大突破、需要逐渐降低数值以取得高绩效的指标,我国城市政府显然缺乏有效的办法,甚至其绩效还不如历史平均值;而对于像“居民月平均工资性收入与平均房价比”、“教育经费占地方财政支出比重”、“社会保险参保人次与常住人口比例”等人民热望改善、需要提高数值以取得高绩效的指标,我国城市政府也表现得一筹莫展,也有低于历史平均数据的现象。这些情况显示了我国城市政府在公共服务提供中采取了一种指标博弈的策略,也就是去做更容易完成的事情,“捏软柿子”,“不做人民热望的,只做能获得绩效分数的”。但毕竟提供公共服务不是参加托福考试,对于自己最没把握的听力可以放任自流,在公共服务提供中,如果人民反映强烈的公共服务持续提供不足,会引起政府的合法性危机,引起人民对政府的敌意。

推论3:获得的并非“公共”服务绩效,而是“领导”、“上级”眼里的服务绩效。这是与推论2一体两面的问题。从L、X市来看,人民迫切需要提供的公共服务提供很不理想,而其它的公共服务相对来说

具有较好的绩效，这其中就有很多属于上级政府、上级领导“严令”完成的公共服务，这是“压力性体制”在公共服务提供中的具体表现，在这种情况下，与其说提供“公共服务”，不如说提供了“上级”和“领导”眼里的服务。

推论4：进步慢了就是落后。从L市来看，大多数公共服务绩效指标的绩效值尽管为正，但也仅仅是比0值稍微大一些而已，比如“市民感知的政府廉洁程度”、“企业感知的政府廉洁程度”、“政策信息依法公开程度”、“人均城市道路面积”分别为0.298606、0.223954、2.314193、0.000747，而作为“榜样”的标杆城市X，其绩效值的增长幅度也不够大。在一个追赶型、赶超型的国家，尤其是在赶超型的地区，进步慢了就是落后，我国城市公共服务提供还需要更高的效率，更高的绩效。

如何破解结论与推论中的难题，提升L市乃至X市，甚至我国所有城市的公共服务提供绩效，笔者认为可以采取如下措施：

措施1：普及型提升措施。由于存在着“进步慢了就是落后”的问题，这就需要我国城市在提供公共服务的过程中，不能“点到为止”，不能蜻蜓点水，无论是“市民感知的政府廉洁程度”、“企业感知的政府廉洁程度”还是“人均城市道路面积”，都需要有大幅度的改善，而不是采取磨洋工的方式，只求有所进步而不求两步并作一步走，需要采取跨越式发展来提供公共服务，绩效才可能有显著的进展。

措施2：民生类公共服务提供需要采取特别项目型绩效评价制度。从L市、X市的绩效结果来看，与民生关系极端密切的住房、社会保障、治安保障类公共服务提供不仅改进较小，甚至还有衰退的现象。究其原因，与这些指标被其它容易完成、容易被捏的“软柿子”淹没不无关系。因为按照总体评价结果来看，城市公共服务还是在提升的，但这却掩盖了人民呼声最高的几种公共服务提供的恶化情况，这也是我国城市对房价怨言四起的原因。我们建议对于这些人民极端关注的公共服务，实行“特事特办”式的项目管理、项目评价，每一项都单独列入政府的决策、计划范围，甚至可以实行一票否决式评价，如果这些项目年度绩效评估结果不理想，就动用“一票否决”权，否决掉政府的公共服务提供工作。

措施3：民生类指标采取绩效预算拨款方式，甚至采取公开招标方式投入预算，解决预算软约束问题。这是与措施2相互配套的措施，即对于那些有意无意被忽视的民生类评价指标，在采取项目制绩效评价的基础上，绩效评估结果要在来年的预算拨款中反映出来。凡是绩效评价结果好的指标、好的民生项目，就在原有基数上增加拨款；而评估为负绩效、0绩效的，则取消拨款，采取西方国家的“日落法”的做法，将这些公共预算款项面向社会实行招标，寻求更高效的组织来承担这些民生项目，为民谋利。

措施4：将连续两年绩效为0、为负值的民生项目所对应的税收由地方税划归中央税，解决地方政府牟利而不服务的痼疾。从我们的评价来看，“城镇居民无房户比例”、“城镇居民3代人居住一套住房的比例”、“居民月平均工资性收入与平均房价比”等与住房有关的民生项目绩效结果竟然为负值，这无疑透露了地方政府在利用“土地财政”赚钱的过程中却放弃了应有的住房、土地的公共服务职能，我们设想如果连续两年该类绩效为负、0值时，就将未来3—5年内该市的该类型地方税种（房产税、土地税等）划归中央，将地方税种改为中央税，这样可以源头上打击“当官不为民服务”的劣政，使得地方政府不敢轻视民生项目。这样，我国政府的人民满意度一定会得到提高。

■作者简介：尚虎平，兰州大学管理学院副教授，管理学博士；甘肃 兰州 730000。

陈星宇，重庆市沙坪坝区天星桥街道干部。

■基金项目：教育部人文社会科学研究青年项目(09YJC630126)；国家社会科学基金重大招标项目(11&·ZD057)；国家社会科学基金重大招标项目(11&·ZD070)

■责任编辑：叶娟丽



