



LNU

辽宁大学经济学部  
Faculty of Economics, Liaoning University

讨论稿系列  
Working Paper Series

No. C2022008

2022-12-15

# 公共责任视角下政府购买公共服务绩效评估指标体系构建研究\*

董 杨

**摘要:** 如何有效地构建绩效评估指标体系,以帮助管理者做出正确决策和改善绩效管理问题,这是一个富有挑战性的难题。本文基于公共责任独特视角,系统构建政府购买公共服务绩效评估指标体系,并建立一个综合的分析框架,试图为公共责任下政府购买公共服务绩效评估的指标提供依据。本文最终确定行政责任、法律责任、专业责任、政治责任、道德责任五个评估维度,采用德尔菲法对绩效指标进行两轮筛选,并获得45个正式绩效指标,其信度和效度水平较高,且通过一致性的检验与分析,有效回应了政府购买公共服务绩效评估在多大程度上增强或降低了公共责任。研究结果进一步表明,健全法律规制、规范行政职能、回应社会需求在政府购买公共服务绩效评估工作中的重要程度。本研究最终的落脚点是所构建的指标体系将在实践应用中体现其实际效果和价值,因此从研究结果来看,强调制度建设的重要性将是未来研究的重点。

**关键词:** 公共责任 政府购买公共服务 评估维度 绩效指标

## 一、问题的提出

加快推进责任政府建设是实现国家治理体系与治理能力现代化的重要议题。自党的十八届三中全会提出《关于全面深化改革若干重大问题的决定》以来,中央政府提出一系列与建设责任政府相关的创新举措。政府购买公共服务绩效评估则是落实公共责任的重要方式,是实现责任政府的路径选择。绩效评估致力于提供关于项目和组织绩效的各种客观的相关信息,这些信息可以用来强化管理和为决策提供依据,达成工作目标和改进整体绩效,以及增加责任感。可见,绩效评估实质上是问责的一种工具,确保公共责任的履行(Behn, 2001;<sup>1</sup>Arie

\*国家社科项目:“基于公共责任视角的政府购买公共服务绩效评估研究”(项目编号:18CZZ033);辽宁省经济社会发展研究项目:“政策反馈视角下辽宁政府购买服务促进社会组织发展研究”(项目编号:20231s1qnkt-039);  
作者:董杨,辽宁大学公共管理学院副教授,沈阳110036;北京大学政府绩效评估中心副研究员,北京100871;  
<sup>1</sup> Behn, R. D. Rethinking Democratic Accountability[M], The Brookings Institution Press, Washington, 2001.

Halachmi, 2002;<sup>1</sup>Bovens, 2008;<sup>2</sup> Anna A. Amirkhanyan, 2011<sup>3</sup>)。因此,深入探究政府购买公共服务绩效评估研究,基于公共责任这一视角显得尤为重要。然而,现有研究中公共责任并没有得到足够重视,具体表现主要集中在以下三个方面:(1)少数成果以公共责任视角仅用于政府购买公共服务活动的考察(李洪佳, 2015;<sup>4</sup>叶托, 薛琬焯, 2018;<sup>5</sup>徐国冲, 赵晓雯, 2020<sup>6</sup>);(2)目前大多研究并没有把公共责任和政府购买公共服务绩效评估相结合,具体到评估维度和绩效指标的设计和实施上均没有体现(叶托, 胡税根, 2015;<sup>7</sup>王春婷等, 2013;<sup>8</sup>范炜烽, 2020<sup>9</sup>)。(3)虽然也有部分学者从公共责任独立视角对政府购买公共服务绩效评估进行考察,但仅仅局限于理论层面的探讨(徐家良, 2015;<sup>10</sup>姜爱华, 杨琼, 2020<sup>11</sup>),在系统性和科学性方面仍有改进余地。由此可见,多数研究把公共责任作为绩效评估体系一个部分,缺乏基于公共责任视角的独立考察与系统设计。同时,上述现有研究为公共责任视角的政府购买公共服务绩效评估探索提供了事实基础。

因此,本文基于公共责任这一视角,试图探讨政府购买公共服务绩效内涵与评估方法,并在此基础上,采用德尔菲法和层次分析法等工具构建基于公共责任的绩效评估体系,以期强化实践者的责任意识,落实公共责任,提高服务绩效。

## 二、公共责任视角下政府购买公共服务绩效评估的分析框架

在《预算绩效评价共性指标体系框架》指导下,从公共责任视角出发,结合政府购买公共服务这一特殊情境,本文只为建立一个综合的绩效评估分析框架。可以在此框架下展开政府购买公共服务项目的绩效评估,并为准确设定公共责任视角下政府购买公共服务的绩效指标提供依据。试图建立一个有用的概念化分析框架不仅与政府购买服务的理论知识相连接,同时为政府购买服务项目的具体目标、潜在的原则以及回应公共责任独特视角和公共部门的价值观搭建桥梁。首先,一系列的绩效评估维度来源于政府购买公共服务活动所涉及到的目标,并对可能的绩效维度、一些潜在的绩效指标进行说明。其次,还将讨论涉及到这些指标更为新颖的绩效信息,关于公共责任及其关系情境下的绩效领域同样值得关注,它们代表了通过公共部门项目寻求服务供给价值方面的绩效,例如,这种价值包括公平、平等、责任和回应性。

### (一) 基于政府购买公共服务目标的绩效范畴

从政府购买公共服务目标情况来看,似乎存在大量可能的目标。因为,政府购买公共服务的目标都与政府的政治、行政、社会和经济等职能履行密切相关。这些目标与我们对政府职能的理解,以及对政府其他更为具体目标的理解同样紧密。换一种角度的话,如果将政府购买公共服务与单纯的经济问题联系起来,就会严重低估其本质上意识形态的广度。尽管现有记录下明确的目标要比单纯的经济目标广泛得多,但已经可以确定政府购买公共服务确实

<sup>1</sup> Arie Halachmi. Performance Measurement, Accountability, and Improved Performance[J]. Public Performance & Management Review, 2002, 25(4).

<sup>2</sup> Mark Bovens, Thomas Schillemans, Paul' T Hart. Does Public Accountability Work? An Assessment Tool[J]. Public Administration, 2008, 86(1).

<sup>3</sup> Anna A. Amirkhanyan. What is the Effect of Performance Measurement on Perceived Accountability Effectiveness in State and Local Government Contracts?[J]. Public Performance & Management Review, 2011, 35(2).

<sup>4</sup> 李洪佳. 政府购买公共服务的公共责任流失及其治理[J]. 理论导刊, 2015(12): 87-91.

<sup>5</sup> 叶托, 薛琬焯. 政府购买公共服务的责任风险与问责模式[J]. 地方财政研究, 2018(4): 17.

<sup>6</sup> 徐国冲, 赵晓雯. 政府购买公共服务的“公共性拆解”风险及其规制[J]. 天津社会科学, 2020(03): 83-88.

<sup>7</sup> 叶托, 胡税根. 政府购买社会服务的绩效评估指标体系研究——基于德尔菲法和层次分析法的应用[J]. 广东行政学院学报, 2015, 27(02): 5-13+45.

<sup>8</sup> 王春婷, 李帆, 林志刚. 政府购买公共服务绩效结构模型建构与实证检测——基于深圳市与南京市的问卷调查与分析[J]. 江苏师范大学学报(哲学社会科学版), 2013, 39(1): 109-115.

<sup>9</sup> 范炜烽, 许燕. 政府向社会力量购买公共服务评估指标体系构建研究[J]. 科学决策, 2020(05): 1-24.

<sup>10</sup> 徐家良, 许源. 合法性理论下政府购买社会组织服务的绩效评估研究[J]. 经济社会体制比较, 2015(06): 187-195.

<sup>11</sup> 姜爱华, 杨琼. 政府购买公共服务“全过程”绩效评价探究[J]. 中央财经大学学报, 2020(03): 3-9+43.

有一些共同的目标，如政治目标提供更好的服务、鼓励参与、创造就业等；经济目标为私营部门或市场发展创造条件、增加竞争等；财政管理目标为减税创造空间、减少公共部门债务；社会目标涉及培育社会力量、增加选择机会等，以及视具体情况而定的其他目标。

### （二）基于公共责任分类及其责任关系的绩效范畴

公共责任分类及责任关系是影响公共责任绩效范畴的主要因素。西方学者将其主要划分为两大类：传统简单的公共责任模型与更加复杂的公共责任模型。其中更加复杂的模型较具代表性的有 Romzek 和 Dubnic (1987)<sup>1</sup>，所确定的等级责任、法律责任、专业责任和政治责任四种公共责任类型；Hodge 和 Coghill (1999, 2007)<sup>2</sup>较为突出的代表性观点认为，民营化状态下公共责任具有较低的政治责任，但具有较高的管理和市场责任，因此有必要强化公共责任道德和公共责任意识。可见，更加复杂的公共责任模型中蕴含着不同的责任关系。从分析中可以得出，公共责任分类及责任关系的绩效至少需从行政、法律、政治、专业和道德五个不同维度下考虑相关指标进行评估，如遵守组织指令、法律规定、公民偏好、尊重专业知识、个人道德等。随着政府购买公共服务进程的推进，所带来的根本变化是要求政府承担更多而不是更少的公共责任，这一点值得我们认真思考。<sup>3</sup>根据国际经验观察，社会公众对公共服务提供者的期望，包括应提供服务达到的经济、效率、质量、社会效益以及公平、回应性都在不断提高，而与公共服务由谁供给无关，这些越来越高的期望使得对政府购买公共服务与公共责任之间的有效性评估也变得更加复杂。

### （三）综合的绩效评估分析框架

理想情况下，前面提到的所有绩效维度都应该是对政府购买服务活动进行绩效评估的一部分。然而，这样的框架涉及到的内容是十分广泛的。本研究需要聚焦公共责任这一视角，在梳理以上多维度多层次含义的基础上进行归纳提炼，同时辅以对公共责任及其责任关系可能的绩效维度的理解，总结出一个综合的基于公共责任视角的分析框架，进一步说明公共责任视角下政府购买公共服务项目所需的评估维度，本文所构建最终的综合分析框架如图 1 所示。此处的绩效类别已尽可能归类为相似的领域，在这个最终框架图中，显示了公共责任视角的五个汇总维度以及每个维度的一些总括性解释性绩效指标。

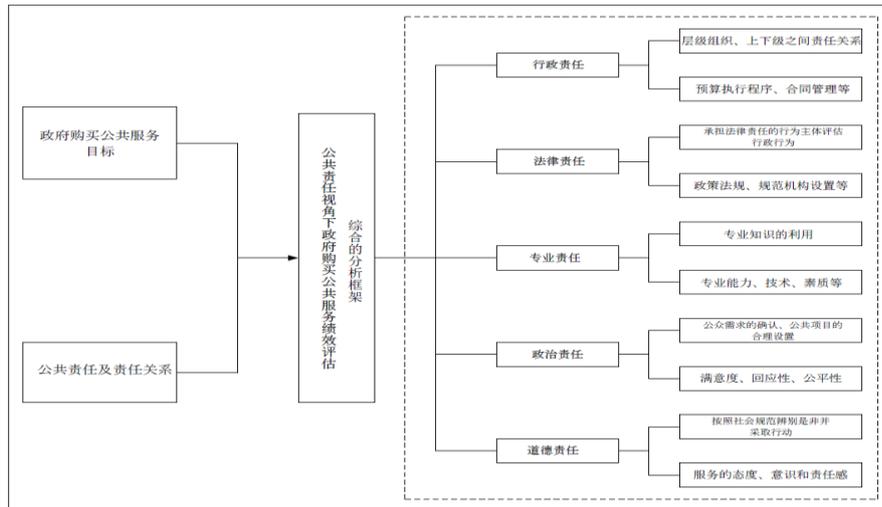


图 1. 综合的公共责任视角下政府购买公共服务绩效评估的分析框架

<sup>1</sup> Barbara S. Romzek. Dynamics of Public Sector Accountability in an Era of Reform. International Review of Administration Sciences, Vol. 66 2000, 21-44.

<sup>2</sup> Graeme A. Hodge, Ken Coghill. Accountability in the Privatized State. Governance: An International Journal of Policy, Administration, and Institutions, Vol. 20, No. 4, October 2007, 695-696.

<sup>3</sup> Graeme A. Hodge, Ken Coghill. Accountability in the Privatized State. Governance: An International Journal of Policy, Administration, and Institutions, Vol. 20, No. 4, October 2007, 685.

概括而言，公共责任视角下研究政府购买公共服务的绩效评估包含两个层面：内部与外部、正式与非正式。一是对购买公共服务的外部责任（行为）进行评估，也即确保责任实现的正式途径包括法律、行政规章所明确的责任归属机制，如行政命令、法院的判决等；二是对购买过程中的内部责任（意识）进行评估，也即实现责任的非正式途径包含责任要求的起源之处，如个人伦理道德、公民偏好、政治过程、专业规范等。可以理解为，公共责任视角下责任行为和责任意识、正式与非正式责任构成一个完整的框架。换言之，法律责任、行政责任、政治责任、专业责任与道德责任，在政府购买公共服务绩效评估活动中的作用均不容忽视，基于此，从而得到了一个理论上完整的公共责任视角下政府购买公共服务绩效评估的分析框架。

### 三、公共责任视角下政府购买公共服务绩效评估指标的拟定与筛选

#### （一）绩效评估的指标初步拟定

基于上述指标依据和公共责任视角下综合的分析框架，遵循政府购买公共服务绩效评估指标体系的建设原则，通过梳理相关文献，并结合有关部门与机构的调查研究，本文初步拟定公共责任视角下政府购买公共服务绩效评估指标，共计 5 个评估维度和 65 个评估指标（见表 2）。

针对指标体系设计的评估维度需要重点说明如下问题：第一，本研究确立的五个评估维度的内涵：法律责任，承担法律责任的行为主体是政府的行政机构和行政官员，以评估活动中行政行为是否符合法律；政治责任，其内容包括公众需求的确认，政策优先次序的科学排列，公共项目的合理设置，公共资源的有效配置，活动结果最大限度满足公众需求等，以评估活动涉及到的上述政治责任内容；行政责任，主要是层级组织内部上下级之间的责任关系，下级服从上级，向上级负责并接受上级的监督，以评估实现活动而投入有形与无形的资源；专业责任，要求行政部门中的专家以公益为首要目标，利用自己的专业知识最大限度地维护和促进公共利益，以评估能力、程序及目标达成度；道德责任，即“按照社会规范辨别是非并采取行动。责任的道德维度超越了法律要求，反映了公共服务的总体文化”。政府及政府官员不能只是按照法规做事，不违背法规并不是其责任的界限，他们负有广泛的道德责任，如公正、诚实、正直、积极主动等，以评估整个活动中服务的态度、意识和责任感。第二，指标所采用的测量方式重视量化方法的运用，通过数据分析方法能够快速展现客观的结果，例如涉及相关的比例、数量、有效性等信息；但同时也会关注一些与政治、社会、文化等因素相关的因果关系，例如满意度、回应性、服务态度与责任意识等。

#### （二）绩效评估的指标首轮筛选

根据初步拟定的指标建立指标体系后，通过发放《基于公共责任视角的政府购买公共服务绩效评估研究专家咨询问卷》进行绩效指标的首轮筛选。为了保证指标筛选的科学性，考虑到本项目的研究视角与研究内容，共向政府部门负责购买公共服务相关业务的工作人员、能够承接政府购买公共服务的社会组织工作人员与从事政府购买公共服务相关研究的专家学者等三类群体发放问卷 79 份。政府工作人员具有改进购买公共服务绩效意愿，并致力于健全购买机制建设，以及丰富的实践经验；社会组织工作人员作为服务承接方，具有供给服务的专业知识与技能，而专家学者则能在理论上予以指导，进而保证问卷设计的有效性和合理性。由于不同问卷填写者自身的专业背景与工作内容不同，在问卷发放时尽可能保证三类群体的填写比例相当。最终回收有效问卷 79 份，其中政府工作人员、社会组织工作人员与专家学者填写问卷数量分别为：39 份、20 份、20 份。为了在权重上匹配，本研究通过电脑生成随机数的方式删去政府工作人员问卷 19 份，最终结果如下（见表 1）：

表 1 德尔菲法专家咨询基本情况

| 调查对象     | 人数 | 职称   | 学历                               |
|----------|----|--|----------------------------------|
| 政府工作人员   | 20 | 处级及以上 3 人<br>副处级 5 人<br>科级 5 人<br>副科级及以下 7 人 | 博士研究生 2 人<br>硕士研究生 9 人<br>本科 9 人 |
| 专家学者     | 20 | 教授 4 人<br>副教授 7 人<br>讲师及以下 9 人               | 博士研究生 20 人                       |
| 社会组织工作人员 | 20 | 处级及以上 1 人<br>科级 4 人<br>副科级及以下 15 人           | 硕士研究生 5 人<br>本科 13 人<br>专科 2 人   |

问卷的指标设计采用五级量表的方式对指标的重要性进行评价,分别为“很不重要”、“不重要”、“一般”、“重要”、“很重要”五个等级,并计 1-5 分。在指标体系的初次筛选中,本研究采用平均数与标准差的计算方式,针对指标重要性的专家意见一致性和重要程度进行考察。在本研究中,受到层次分析法(AHP)指标体裁的限制,将指标筛选的标准界定为平均数 4.3 以下(重要和很重要之间)与标准差 1 以下,在首次筛选中,符合任意一个标准界定即删除指标。最终问卷结果如下(见表 2):

表 2 问卷结果的平均数与标准差统计

| 指标<br>维度 | 指标名称            | 平均数   | 标准差   | 指标<br>维度 | 指标名称      | 平均数   | 标准差   |
|----------|-----------------|-------|-------|----------|-----------|-------|-------|
| 行政<br>责任 | 购买方工作人员的数量      | 3.583 | 0.962 | 专业<br>责任 | 服务流程的透明度  | 4.417 | 0.72  |
|          | 购买方管理人员与行政人员的比例 | 3.383 | 0.922 |          | 服务合同的明确度  | 4.417 | 0.743 |
|          | 购买方工作人员的专业素质    | 4.45  | 0.91  |          | 服务组织的成长情况 | 4.183 | 0.701 |
|          | 购买服务预算编制的合理性    | 4.4   | 0.924 |          | 政府规模的缩小情况 | 3.717 | 0.94  |
|          | 购买服务方案设计的完整性    | 4.383 | 0.865 |          | 服务目标的明确性  | 4.233 | 0.81  |
|          | 购买服务过程的规范性      | 4.467 | 0.812 |          | 服务目标的达成度  | 4.35  | 0.709 |
|          | 购买服务预算执行进展程度    | 4.217 | 0.885 | 政治<br>责任 | 公众参与程度    | 4.067 | 0.821 |
|          | 购买服务计划目标的完成度    | 4.45  | 0.746 |          | 公众需求满足程度  | 4.317 | 0.651 |
|          | 服务提供成本的节省程度     | 4.1   | 0.796 |          | 政府信息公开程度  | 4.3   | 0.696 |

|            |              |       |          |              |              |           |       |       |
|------------|--------------|-------|----------|--------------|--------------|-----------|-------|-------|
|            | 合同管理成本的合理性   | 4.15  | 0.84     |              | 购买流程的公平性     | 4.467     | 0.724 |       |
|            | 购买服务中竞争的有效程度 | 4.283 | 0.976    |              | 服务提供的及时性     | 4.35      | 0.685 |       |
|            | 服务效率的改善水平    | 4.333 | 0.896    |              | 服务提供的回应性     | 4.35      | 0.709 |       |
|            | 服务设施的完善度     | 4.133 | 0.892    |              | 服务提供的精准性     | 4.483     | 0.596 |       |
|            | 购买服务信息资源的有效性 | 4.283 | 0.885    |              | 服务对象的满意度     | 4.567     | 0.533 |       |
| 法律<br>责任   | 组织机构设置的规范化   | 4.35  | 0.799    |              | 服务质量的改善水平    | 4.383     | 0.613 |       |
|            | 行政体制的健全性     | 4.3   | 0.83     |              | 政策目标的达成度     | 4.4       | 0.588 |       |
|            | 政策法规的完善度     | 4.467 | 0.833    |              | 社会组织的增长率     | 4.033     | 0.863 |       |
|            | 监督机制的有效性     | 4.583 | 0.766    |              | 承接方的成长情况     | 4.167     | 0.74  |       |
|            | 问责机制的清晰度     | 4.517 | 0.725    |              | 权力寻租状况       | 4.3       | 0.766 |       |
|            | 需求表达机制的准确性   | 4.433 | 0.745    |              | 服务完成情况及后续可行性 | 4.267     | 0.733 |       |
|            | 市场机制的成熟度     | 4.2   | 0.819    |              | 道德<br>责任     | 服务组织的责任文化 | 4.183 | 0.676 |
|            | 沟通协调机制的完备性   | 4.3   | 0.83     |              |              | 工作人员的服务态度 | 4.367 | 0.61  |
|            | 风险防控机制的健全性   | 4.35  | 0.84     |              |              | 工作人员的敬业精神 | 4.4   | 0.643 |
| 信息收集机制的完整性 | 4.35         | 0.88  | 服务提供的人性化 | 4.317        |              | 0.651     |       |       |
| 专业<br>责任   | 承接机构的数量      | 3.8   | 0.84     | 购买方管理人员的责任感  |              | 4.433     | 0.533 |       |
|            | 承接方工作人员胜任力水平 | 4.433 | 0.673    | 购买方行政人员的积极性  | 4.367        | 0.581     |       |       |
|            | 承接方工作人员的专业能力 | 4.533 | 0.623    | 承接方管理人员的责任感  | 4.45         | 0.565     |       |       |
|            | 承接方管理人员的组织能力 | 4.383 | 0.783    | 承接方行政人员的积极性  | 4.317        | 0.596     |       |       |
|            | 承接方资金使用的规范性  | 4.45  | 0.811    | 购买方工作人员的服务意识 | 4.333        | 0.655     |       |       |
|            | 承接方财务制度的健全程度 | 4.383 | 0.761    | 承接方工作人员的服务意识 | 4.333        | 0.655     |       |       |

|  |          |       |       |  |              |       |       |
|--|----------|-------|-------|--|--------------|-------|-------|
|  | 服务方案的完整性 | 4.383 | 0.715 |  | 购买方掌握公众需求的程度 | 4.333 | 0.681 |
|  | 服务提供的专业性 | 4.533 | 0.65  |  | 承接方承诺完成的及时性  | 4.317 | 0.676 |
|  | 服务流程的规范化 | 4.3   | 0.72  |  |              |       |       |

与此同时，为了保证问卷结果的信度与效度，在效度方面通过向专家学者和相关工作人员进行专业咨询与讨论，可以保证内容效度，而在信度方面，本研究通过 SPSS 软件进行问卷的信度检验，采用克隆巴赫系数与折半信度系数进行分析。经过分析，问卷的整体克隆巴赫系数值为 0.972，大于学界公认的 0.9，而问卷中五个评估维度的克隆巴赫系数均大于 0.9，分别为：

表 3 五个指标维度的克隆巴赫系数

|        | 行政责任  | 法律责任  | 专业责任  | 政治责任  | 道德责任  |
|--------|-------|-------|-------|-------|-------|
| 克隆巴赫系数 | 0.930 | 0.949 | 0.947 | 0.942 | 0.948 |

而从问卷的整体折半系数来看，由于该问卷有 65 个指标，折半分成两部分时，两部分分别的分析项数量并不相等，因而应该使用不等长折半系数(Spearman-Brown 系数)进行信度质量判断。经过计算，折半信度系数值为 0.911，大于 0.9，指标的信度质量具有相当的保证。

通过对问卷结果的描述性统计分析，在初始的 65 个指标中，一共有 19 个指标满足平均数小于 4.3 的标准，而没有指标的标准差大于 1。这 19 个满足平均数标准的指标分别为：“购买方工作人员的数量(3.58)”、“购买方管理人员与行政人员的比例(3.383)”、“购买服务信息资源的有效性(4.3)”、“服务流程的规范化(4.3)”、“服务完成情况及后续可行性(4.3)”、“服务提供的人性化(4.31)”、“市场机制的成熟度(4.26)”、“服务目标的明确性(4.26)”、“购买服务预算执行的进展程度(4.22)”、“合同管理成本的合理性(4.22)”、“服务设施的完善度(4.22)”、“承接方的成长情况(4.22)”、“服务提供成本的节省程度(4.16)”、“服务组织的成长情况(4.16)”、“服务组织的责任文化(4.14)”、“公众参与程度(4.1)”、“社会组织的增长率(4.1)”、“承接机构的数量(3.9)”以及“政府规模的缩小情况(3.717)”，这 19 个指标的平均数均小于 4.3。由此可见，在专家咨询中上述指标的相对重要性较低。

因此，经过指标体系的首轮筛选，本研究暂定将“购买服务信息资源的有效性、服务流程的规范化、服务完成情况及后续可行性、服务提供的人性化、市场机制的成熟度、服务目标的明确性、购买服务预算执行的进展程度、合同管理成本的合理性、服务设施的完善度、承接方的成长情况、服务提供成本的节省程度、服务组织的成长情况、服务组织的责任文化、公众参与程度、社会组织的增长率、承接机构的数量、政府规模的缩小情况、购买方工作人员的数量、购买方管理人员与行政人员的比例”19 个指标删去，保留 49 个指标。但考虑到本研究基于“公共责任”视角的特征，需要保证指标体系体现出“公共责任”的研究倾向，因此需要进一步对指标的删减和合并进行讨论。

### (三) 绩效评估指标的最终确定

在经过指标体系的首轮筛选之后，考虑到一些指标存在内容重复、相关度过高及重要指标被遗漏等情况，本研究经过多次组织专家讨论与咨询，对指标进行再次筛选和整合。在指标体系的维度层面，基于公共责任的研究视角，本研究并未从政府购买公共服务的生命周

期展开维度建设，而是以前文所建立综合的分析框架为依据，建立起“行政责任”、“法律责任”、“专业责任”、“政治责任”与“道德责任”五个指标维度。

而在具体指标的整合层面，考虑到前期的首轮筛选的 19 个指标中，“合同管理成本的合理性”、“购买服务中竞争的有效程度”以及“服务组织的责任文化”三个指标仍然具有一定的理论意义。具体表现为“合同管理成本的合理性”主要关注政府部门在购买公共服务过程中的财务状况，体现出对纳税人的负责意识。“购买服务中竞争的有效程度”则体现出购买服务的过程中是否从公正、客观的维度设立准入门槛。“服务组织的责任文化”则是直接体现服务组织是否存在道德责任底线的重要衡量指标。因此本研究考虑将这三个指标予以保留。此外，本研究将其他相关指标进行整合，把“购买方管理人员的责任感”与“购买方行政人员的积极性”整合为“购买方工作人员的责任意识”，把“承接方管理人员的责任感”与“承接方行政人员的积极性”整合为“承接方工作人员的责任意识”，把“购买服务预算编制的合理性”整合到“购买服务方案设计的完整性”中。同时考虑到“服务流程的规范化”指标实际上体现在“购买流程的公平性、服务提供的及时性、服务提供的回应性”等指标中，故予以删去。经过多次的修改和完善，本研究最终确定的指标体系包含 45 个指标，具体情况如表 4：

表 4 最终指标体系

| 编号 | 评估维度 | 编号  | 评估指标         | 编号 | 评估维度 | 编号  | 评估指标          |
|----|------|-----|--------------|----|------|-----|---------------|
| A1 | 行政责任 | B1  | 购买方工作人员的专业素质 | A3 | 专业责任 | B17 | 承接方工作人员的胜任力水平 |
|    |      | B2  | 购买服务方案设计的完整性 |    |      | B18 | 承接方工作人员的专业能力  |
|    |      | B3  | 购买服务过程的规范性   |    |      | B19 | 承接方管理人员的组织能力  |
|    |      | B4  | 购买服务计划目标的完成度 |    |      | B20 | 承接方资金使用的规范性   |
|    |      | B5  | 合同管理成本的合理性   |    |      | B21 | 承接方财务制度的健全程度  |
|    |      | B6  | 购买服务中竞争的有效程度 |    |      | B22 | 服务方案的完整度      |
|    |      | B7  | 服务效率的改善水平    |    |      | B23 | 服务提供的专业性      |
| A2 | 法律责任 | B8  | 组织机构的设置规范化   | A4 | 政治责任 | B24 | 服务流程的透明度      |
|    |      | B9  | 行政体制的健全性     |    |      | B25 | 服务合同的明确度      |
|    |      | B10 | 政策法规的完善度     |    |      | B26 | 服务目标的达成度      |
|    |      | B11 | 监督机制的有效性     |    |      | B27 | 公众需求满足程度      |
|    |      | B12 | 问责机制的清晰度     |    |      | B28 | 政府信息公开程度      |
|    |      | B13 | 需求表达机制的准确性   |    |      | B29 | 购买流程的公平性      |
|    |      | B14 | 沟通协调机制的完备性   |    |      | B30 | 服务提供的及时性      |
|    |      | B15 | 风险防控机制的全性    |    |      | B31 | 服务提供的回应性      |

|    |      |     |              |  |  |     |           |
|----|------|-----|--------------|--|--|-----|-----------|
|    |      | B16 | 信息收集机制的完整性   |  |  | B32 | 服务提供的精准性  |
| A5 | 道德责任 | B37 | 服务组织的责任文化    |  |  | B33 | 服务对象的满意度  |
|    |      | B38 | 工作人员的服务态度    |  |  | B34 | 服务质量的改善水平 |
|    |      | B39 | 工作人员的敬业精神    |  |  | B35 | 政策目标的达成度  |
|    |      | B40 | 购买方工作人员的责任意识 |  |  | B36 | 权力寻租状况    |
|    |      | B41 | 承接方工作人员的责任意识 |  |  |     |           |
|    |      | B42 | 购买方工作人员的服务意识 |  |  |     |           |
|    |      | B43 | 承接方工作人员的服务意识 |  |  |     |           |
|    |      | B44 | 承接方承诺完成的及时性  |  |  |     |           |
|    |      | B45 | 购买方掌握公众需求的程度 |  |  |     |           |

#### 四、公共责任视角下政府购买公共服务绩效评估指标体系的权重确定

由于本研究的变量较多，且变量之间存在复杂的结构关系，因此需要对这些变量进行一个明确的估值。层次分析法（AHP）是由美国学者马斯·赛蒂提出的一项评价方法，一般用于绩效评价的过程中。因此本研究选择采用层次分析法来进行指标体系的权重确定，通过“迈实 AHP”软件实现指标体系的构建、判断矩阵的生成以及问卷数据的分析。

##### （一）构建层级结构模型

层次分析法（AHP）的使用需要构造层次结构模型。首先是目标层，也就是基于公共责任视角的政府购买公共服务绩效评估指标体系权重（T）的确定。其次是准则层，也就是指标体系的一级指标，包括：行政责任（A1）、法律责任（A2）、专业责任（A3）、政治责任（A4）和道德责任（A5）五个维度。最后是要素层，主要是指一级指标分类下的二级指标（B1-B45），在本研究中具体映射 45 个指标。最终，通过三个层次的结构模型建设，构建出基于公共责任视角的政府购买公共服务绩效评估指标体系权重确定的结构模型（见图 1）。

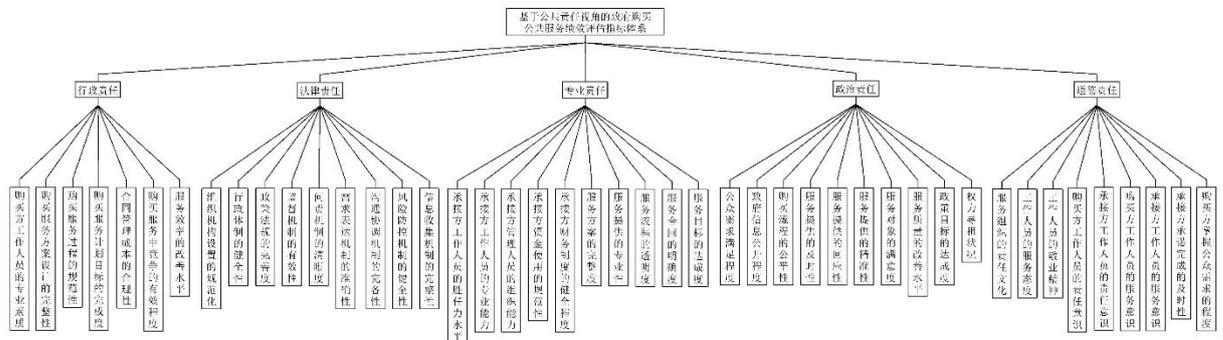


图 1 基于公共责任视角的政府购买公共服务绩效评价层次结构模型

##### （二）构造判断矩阵

层次分析法（AHP）在进行层级结构模型的构建后，需要在目标层以外的层次上展开指标重要性的两两比较分析，而通过两两比较构建出的判断矩阵，将通过调查问卷的形式发放给相关领域内的权威专家。本研究在同一层级上进行了指标重要性的两两比较，将指标的重要性程度衡量尺度划分为1-9之间的9个等级，并取“1、3、5、7、9”作为问卷的选项，用以说明一个指标对于另一个指标来说其重要性从一致到绝对重要的趋势，并在“其他”一栏增加“2、4、6、8”的下拉菜单为专家学者提供介于两个重要性等级之间的重要性选择。在调查问卷中，“1”表示两个指标同样重要，“3”表示一个指标比另一个指标稍微重要，“5”表示一个指标比另一个指标比较重要，“7”表示一个指标比另一个指标非常重要，而“9”表示一个指标比另一个指标绝对重要。根据判断矩阵的构建，本文最终制作了调查问卷并予以发放。

为了保证问卷的科学性和专业性，在问卷的发放对象上与第一轮的群体选择基本一致，主要发放给从事购买公共服务相关工作的政府人员、社会组织人员以及专家学者。最终问卷发放150份，回收问卷90份，其中专家学者问卷31份，政府工作人员问卷30份，社会组织人员29份。将专家问卷录入软件后，需要对问卷的判断矩阵进行一致性检验，共有8份问卷没有通过一致性检验。这也意味着最终分析的结果由25位专家学者、29位政府工作人员以及28位社会组织人员填写的问卷组成。

### （三）层次分析法的计算结果

由于本研究采用的是AHP层次分析法，其中构造判断矩阵需要专家对其进行打分，因此需要对90位专家的打分进行加权平均数计算最终得到一个群决策的打分值。

首先，从基于公共责任视角的政府购买公共服务绩效评估的总体目标来看，5个评估维度之间的重要性比较如下（见表5）：

表5 判断矩阵 T-A 的群决策结果

| T  | A1     | A2     | A3     | A4     | A5     | 权重(wi) |
|----|--------|--------|--------|--------|--------|--------|
| A1 | 1      | 1.2091 | 1.7721 | 1.5769 | 2.3911 | 0.2816 |
| A2 | 0.8271 | 1      | 2.2983 | 1.8176 | 2.7019 | 0.2899 |
| A3 | 0.5643 | 0.4351 | 1      | 1.0899 | 1.712  | 0.1586 |
| A4 | 0.6342 | 0.5502 | 0.9175 | 1      | 2.0479 | 0.1704 |
| A5 | 0.4182 | 0.3701 | 0.5841 | 0.4883 | 1      | 0.0994 |

$\lambda_{\max}=5.0404$ ,  $CR=0.0091$ ,  $CI=0.0102$

其次，在行政责任维度下，7个维度之间的重要性比较结果如下（见表6）：

表6 判断矩阵 A1-Bi 的群决策结果 (i=1-7)

| A1 | B1     | B2     | B3     | B4     | B5     | B6     | B7     | 权重     |
|----|--------|--------|--------|--------|--------|--------|--------|--------|
| B1 | 1      | 1.2673 | 1.3567 | 1.5657 | 1.7121 | 2.0625 | 1.8883 | 0.2066 |
| B2 | 0.7891 | 1      | 1.0774 | 1.2827 | 1.4409 | 1.6421 | 1.6618 | 0.1683 |
| B3 | 0.7371 | 0.9282 | 1      | 1.5677 | 1.7603 | 1.9053 | 1.8445 | 0.1792 |
| B4 | 0.6379 | 0.7796 | 0.6379 | 1      | 1.5165 | 1.7184 | 1.8933 | 0.1458 |
| B5 | 0.5841 | 0.694  | 0.5681 | 0.6594 | 1      | 1.467  | 1.3603 | 0.1153 |
| B6 | 0.4848 | 0.609  | 0.5249 | 0.582  | 0.6817 | 1      | 1.2448 | 0.0947 |
| B7 | 0.5296 | 0.6018 | 0.5421 | 0.5282 | 0.7351 | 0.8033 | 1      | 0.0901 |

$\lambda_{\max}=7.0498$ ,  $CR=0.0061$ ,  $CI=0.0083$

第三，在法律责任的维度下，9个指标之间的重要性比较情况如下（见表7）：

表7 判断矩阵 A2-Bi 的群决策结果 (i=8-16)

| A2  | B8     | B9     | B10    | B11    | B12    | B13    | B14    | B15    | B16    | 权重     |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| B8  | 1      | 1.1649 | 1.1092 | 1.199  | 1.305  | 1.6918 | 1.7542 | 1.6195 | 1.9786 | 0.174  |
| B9  | 0.8585 | 1      | 1.0968 | 1.234  | 1.4359 | 1.783  | 1.8364 | 1.7315 | 1.9379 | 0.1463 |
| B10 | 0.9015 | 0.9117 | 1      | 1.1287 | 1.5564 | 1.6337 | 2.0226 | 1.8798 | 2.1459 | 0.1471 |
| B11 | 0.834  | 0.8104 | 0.886  | 1      | 1.325  | 1.8412 | 1.9432 | 1.8167 | 2.2809 | 0.1392 |
| B12 | 0.7663 | 0.6964 | 0.6425 | 0.7547 | 1      | 1.5518 | 1.7366 | 1.5442 | 1.9744 | 0.1151 |
| B13 | 0.5911 | 0.5608 | 0.6121 | 0.5431 | 0.6444 | 1      | 1.4731 | 1.4612 | 1.5332 | 0.0901 |
| B14 | 0.5701 | 0.5446 | 0.4944 | 0.5146 | 0.5758 | 0.6788 | 1      | 1.025  | 1.1736 | 0.0734 |
| B15 | 0.6175 | 0.5775 | 0.532  | 0.5505 | 0.6476 | 0.6844 | 0.9756 | 1      | 1.3097 | 0.0773 |
| B16 | 0.5054 | 0.516  | 0.466  | 0.4384 | 0.5065 | 0.6522 | 0.8521 | 0.7636 | 1      | 0.0644 |

$\lambda_{\max}=9.0523$ ,  $CR=0.0045$ ,  $CI=0.0065$

第四，在专业责任的维度下，10个指标之间的重要性比较情况如下（见表8）：

表8 判断矩阵 A3-Bi 的群决策结果 (i=17-26)

| A3  | B17    | B18    | B19    | B20    | B21    | B22    | B23    | B24    | B25    | B26    | 权重     |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| B17 | 1      | 1.3819 | 1.9416 | 1.3789 | 1.4677 | 1.7466 | 1.4592 | 2.0267 | 1.9159 | 1.5184 | 0.1502 |
| B18 | 0.7236 | 1      | 1.596  | 1.3447 | 1.5542 | 1.7769 | 1.6082 | 1.8637 | 1.8535 | 1.4093 | 0.1374 |
| B19 | 0.515  | 0.6366 | 1      | 1.0677 | 0.9982 | 1.2986 | 1.2328 | 1.692  | 1.6075 | 1.3125 | 0.1035 |
| B20 | 0.7252 | 0.7437 | 0.9366 | 1      | 1.1528 | 1.6036 | 1.4329 | 1.8574 | 1.5068 | 1.4539 | 0.1146 |
| B21 | 0.6813 | 0.6434 | 1.0018 | 0.8647 | 1      | 1.3403 | 1.34   | 1.4907 | 1.5561 | 1.3532 | 0.1044 |
| B22 | 0.5725 | 0.5628 | 0.77   | 0.6236 | 0.7461 | 1      | 0.9057 | 1.2724 | 1.3107 | 1.1856 | 0.0826 |
| B23 | 0.6853 | 0.6218 | 0.8112 | 0.6979 | 0.7462 | 1.1041 | 1      | 1.4666 | 1.4027 | 1.1589 | 0.0898 |
| B24 | 0.4634 | 0.5366 | 0.591  | 0.5384 | 0.6708 | 0.7859 | 0.6819 | 1      | 1.0255 | 0.8788 | 0.0675 |
| B25 | 0.5219 | 0.5395 | 0.6221 | 0.6637 | 0.6426 | 0.763  | 0.7129 | 0.9751 | 1      | 0.9033 | 0.0693 |
| B26 | 0.6586 | 0.7096 | 0.7519 | 0.6878 | 0.739  | 0.8435 | 0.8629 | 1.1379 | 1.1071 | 1      | 0.0808 |

$\lambda_{\max}=10.0549$ ,  $CR=0.0041$ ,  $CI=0.0061$

第五，在政治责任的维度下，10个指标之间的重要性比较情况如下（见表9）：

表9 判断矩阵 A4-Bi 的群决策结果 (i=27-36)

| A4  | B27    | B28    | B29    | B30    | B31    | B32    | B33    | B34    | B35    | B36    | 权重     |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| B27 | 1      | 1.4748 | 1.4425 | 1.8301 | 1.8918 | 1.9531 | 1.8355 | 1.9139 | 2.1682 | 3.0354 | 0.1702 |
| B28 | 0.6781 | 1      | 1.2467 | 1.4081 | 1.4243 | 1.4656 | 1.2828 | 1.3282 | 1.3018 | 2.1879 | 0.1222 |
| B29 | 0.6932 | 0.8021 | 1      | 1.5463 | 1.5411 | 1.4456 | 1.4386 | 1.4253 | 1.6095 | 2.1879 | 0.1242 |
| B30 | 0.5464 | 0.7102 | 0.6467 | 1      | 1.1171 | 1.3824 | 1.268  | 1.3906 | 1.299  | 2.26   | 0.1012 |
| B31 | 0.5286 | 0.7021 | 0.6489 | 0.8951 | 1      | 1.169  | 1.2351 | 1.3375 | 1.3186 | 2.0685 | 0.096  |
| B31 | 0.512  | 0.6823 | 0.6918 | 0.7234 | 0.8554 | 1      | 1.063  | 1.2489 | 1.301  | 1.9671 | 0.0891 |
| B33 | 0.5448 | 0.7795 | 0.6951 | 0.7886 | 0.8096 | 0.9407 | 1      | 1.2648 | 1.4063 | 1.9914 | 0.0905 |
| B34 | 0.5225 | 0.7529 | 0.7016 | 0.7191 | 0.7476 | 0.8007 | 0.7907 | 1      | 1.321  | 1.8983 | 0.0824 |

|     |        |        |        |        |        |        |        |        |        |        |        |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| B35 | 0.4612 | 0.7682 | 0.6213 | 0.7698 | 0.7584 | 0.7686 | 0.7111 | 0.757  | 1      | 1.8856 | 0.0754 |
| B36 | 0.3294 | 0.4571 | 0.4425 | 0.4834 | 0.5084 | 0.5022 | 0.5268 | 0.5303 | 0.5511 | 1.8145 | 0.0489 |

$\lambda_{\max}=10.0617$ ,  $CR=0.0046$ ,  $CI=0.0069$

第六，在专业责任的维度下，9个指标之间的重要性比较情况如下（见表10）：

表10 判断矩阵 A5-Bi 的群决策结果 (i=37-45)

| A5  | B37    | B38    | B39    | B40    | B41    | B42    | B43    | B44    | B45    | 权重     |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| B37 | 1      | 1.2369 | 1.1909 | 1.2369 | 1.2007 | 1.2787 | 1.5218 | 1.1841 | 1.3149 | 0.1326 |
| B38 | 0.9492 | 1      | 1.1258 | 1.1817 | 1.1251 | 1.5079 | 1.5559 | 1.1253 | 1.3262 | 0.1307 |
| B39 | 0.8397 | 0.8883 | 1      | 1.1734 | 1.2543 | 1.4133 | 1.3314 | 1.2826 | 1.3714 | 0.1262 |
| B40 | 0.8085 | 0.8462 | 0.8523 | 1      | 1.0768 | 1.473  | 1.652  | 1.3645 | 1.4612 | 0.1238 |
| B41 | 0.8328 | 0.8888 | 0.7973 | 0.9287 | 1      | 1.4854 | 1.5601 | 1.5927 | 1.5841 | 0.1245 |
| B42 | 0.782  | 0.6632 | 0.7076 | 0.6789 | 0.6732 | 1      | 1.2449 | 1.2468 | 1.382  | 0.0976 |
| B43 | 0.6571 | 0.6427 | 0.7511 | 0.6053 | 0.641  | 0.8033 | 1      | 1.0609 | 1.2205 | 0.0871 |
| B44 | 0.8445 | 0.8887 | 0.7797 | 0.7329 | 0.6279 | 0.802  | 0.9426 | 1      | 1.4592 | 0.0956 |
| B45 | 0.7605 | 0.754  | 0.7292 | 0.6843 | 0.6313 | 0.7236 | 0.8193 | 0.6853 | 1      | 0.0818 |

$\lambda_{\max}=9.0695$ ,  $CR=0.006$ ,  $CI=0.0087$

#### （四）计算全局权重

由于本研究指标体系的内容较多，因此以准则层（T-Ai）的判断矩阵指标权重计算为例。在计算权重前，首先要对判断矩阵进行向量归一化处理。

对矩阵 T 以列进行归一化，得到新的矩阵 T1：

$$T1 = \begin{bmatrix} 0.2904 & 0.3392 & 0.2696 & 0.2640 & 0.2426 \\ 0.2402 & 0.2805 & 0.3497 & 0.3043 & 0.2742 \\ 0.1639 & 0.1221 & 0.1522 & 0.1824 & 0.1738 \\ 0.1842 & 0.1544 & 0.1396 & 0.1674 & 0.2078 \\ 0.1214 & 0.1038 & 0.0889 & 0.0818 & 0.1015 \end{bmatrix}$$

然后再将矩阵 T1 按行求和，并再次归一化后得到矩阵 T2：

$$T2 = \begin{bmatrix} 0.281 \\ 0.290 \\ 0.159 \\ 0.170 \\ 0.099 \end{bmatrix}$$

T-Ai 的权重向量为 (0.281、0.290、0.159、0.170、0.099)。

在判断矩阵的权重向量计算后，考虑到不同专家对问卷内容的认识存在偏好，不能保证每份问卷都有良好的一致性，因此需要通过矩阵进行一致性检验来保证权重之间不存在矛盾。判断矩阵一致性检验的公式为：

$$CR = \frac{CI}{RI}$$

$$\text{其中：} CI = \frac{\lambda_{\max} - n}{(n-1)}, \lambda_{\max} = \frac{\sum(Aw/w)}{n}$$

最后，为了验证总排序结果的一致性，也需要进行一次一致性检验。

$CR = (0.281 \times 0.0083 + 0.290 \times 0.0065 + 0.159 \times 0.0061 + 0.170 \times 0.0069 + 0.099 \times 0.0087) / (0.281 \times 1.36 + 0.290 \times 1.46 + 0.159 \times 1.49 + 0.170 \times 1.49 + 0.099 \times 1.46) = 0.005 < 1$

经过 AHP 层次分析法的计算分析后，本研究建立的指标体系及权重配置如下：

| 编号  | 准则层        | 编号     | 要素层           | 全局权重   | 同级权重   | 排名 |
|-----|------------|--------|---------------|--------|--------|----|
| A1  | 行政责任       | B1     | 购买方工作人员的专业素质  | 0.0582 | 0.2066 | 1  |
|     |            | B2     | 购买服务方案设计的完整性  | 0.0474 | 0.1683 | 3  |
|     |            | B3     | 购买服务过程的规范性    | 0.0505 | 0.1792 | 2  |
|     |            | B4     | 购买服务计划目标的完成度  | 0.0411 | 0.1458 | 7  |
|     |            | B5     | 合同管理成本的合理性    | 0.0325 | 0.1153 | 10 |
|     |            | B6     | 购买服务中竞争的有效程度  | 0.0267 | 0.0947 | 12 |
|     |            | B7     | 服务效率的改善水平     | 0.0254 | 0.0901 | 14 |
| A2  | 法律责任       | B8     | 组织机构设置的规范化    | 0.0426 | 0.174  | 4  |
|     |            | B9     | 行政体制的健全性      | 0.0424 | 0.1463 | 6  |
|     |            | B10    | 政策法规的完善度      | 0.0426 | 0.1471 | 5  |
|     |            | B11    | 监督机制的有效性      | 0.0404 | 0.1392 | 8  |
|     |            | B12    | 问责机制的清晰度      | 0.0334 | 0.1151 | 9  |
|     |            | B13    | 需求表达机制的准确性    | 0.0261 | 0.0901 | 13 |
|     |            | B14    | 沟通协调机制的完备性    | 0.0213 | 0.0734 | 18 |
|     |            | B15    | 风险防控机制的健全性    | 0.0224 | 0.0773 | 16 |
| B16 | 信息收集机制的完整性 | 0.0187 | 0.0644        | 21     |        |    |
| A3  | 专业责任       | B17    | 承接方工作人员的胜任力水平 | 0.0238 | 0.1502 | 15 |
|     |            | B18    | 承接方工作人员的专业能力  | 0.0218 | 0.1374 | 17 |
|     |            | B19    | 承接方管理人员的组织能力  | 0.0164 | 0.1035 | 26 |
|     |            | B20    | 承接方资金使用的规范性   | 0.0182 | 0.1146 | 22 |
|     |            | B21    | 承接方财务制度的健全程度  | 0.0166 | 0.1044 | 24 |
|     |            | B22    | 服务方案的完整度      | 0.0131 | 0.0826 | 32 |
|     |            | B23    | 服务提供的专业性      | 0.0142 | 0.0898 | 29 |
|     |            | B24    | 服务流程的透明度      | 0.0107 | 0.0675 | 40 |
|     |            | B25    | 服务合同的明确度      | 0.011  | 0.0693 | 39 |
|     |            | B26    | 服务目标的达成度      | 0.0128 | 0.0808 | 35 |
| A4  | 政治责任       | B27    | 公众需求满足程度      | 0.029  | 0.1702 | 11 |
|     |            | B28    | 政府信息公开程度      | 0.0208 | 0.1222 | 20 |
|     |            | B29    | 购买流程的公平性      | 0.0212 | 0.1242 | 19 |
|     |            | B30    | 服务提供的及时性      | 0.0172 | 0.1012 | 23 |
|     |            | B31    | 服务提供的回应性      | 0.0164 | 0.096  | 25 |
|     |            | B32    | 服务提供的精准性      | 0.0152 | 0.0891 | 28 |
|     |            | B33    | 服务对象的满意度      | 0.0154 | 0.0905 | 27 |
|     |            | B34    | 服务质量的改善水平     | 0.014  | 0.0824 | 30 |
|     |            | B35    | 政策目标的达成度      | 0.0128 | 0.0754 | 34 |
|     |            | B36    | 权力寻租状况        | 0.0083 | 0.0489 | 44 |
| A5  | 道德责任       | B37    | 服务组织的责任文化     | 0.0132 | 0.1326 | 31 |
|     |            | B38    | 工作人员的服务态度     | 0.013  | 0.1307 | 33 |
|     |            | B39    | 工作人员的敬业精神     | 0.0125 | 0.1262 | 36 |

|   |     |              |        |        |    |
|---|-----|--------------|--------|--------|----|
| 任 | B40 | 购买方工作人员的责任意识 | 0.0123 | 0.1238 | 38 |
|   | B41 | 承接方工作人员的责任意识 | 0.0124 | 0.1245 | 37 |
|   | B42 | 购买方工作人员的服务意识 | 0.0097 | 0.0976 | 41 |
|   | B43 | 承接方工作人员的服务意识 | 0.0086 | 0.0871 | 43 |
|   | B44 | 承接方承诺完成的及时性  | 0.0095 | 0.0956 | 42 |
|   | B45 | 购买方掌握公众需求的程度 | 0.0081 | 0.0818 | 45 |

## 五、基于公共责任视角的政府购买公共服务绩效评价指标体系的结果分析

本研究结合层次分析法（AHP）得到的研究表明，基于公共责任视角下政府购买公共服务绩效评估指标体系的五个维度中，法律责任维度对整个购买公共服务绩效评估的影响最大，其次是行政责任、政治责任、专业责任和道德责任。从准则层的权重分配结果来看，健全法律规制、规范行政职能、回应社会需求在政府购买公共服务绩效评估中的重要程度相对要高。这也意味着，虽然说政府购买公共服务实际上已经串联起“政府-社会”的多元主体结构，但实际上在政府购买公共服务中政府部门的制度安排等角色仍然起到关键作用。本研究在公共责任视角下构建理论分析框架之初，政府购买公共服务绩效评估注重从法律向道德维度由刚性向柔性发展的过程，试图寻找政府购买公共服务中道德责任增加的重要意义。然而，最终研究结果显示，公共责任视角下政府购买公共服务绩效评估中法律责任仍是最重要的评价指标。这说明政府购买公共服务中公共责任的实现依然需要刚性的底线约束，而道德责任对政府购买公共服务绩效评估的影响最小，这一维度还未引起足够重视的结果值得深思。

在法律责任（0.290）维度下，组织机构设置的规范化、政策法规的完善度、行政体制的健全性占据重要性的前三名。从该维度的整个情况来看，前三名重要性指标主要考察的是政府自身的机构体制和政策法规建设，反而对那些体现出政府和社会互动的需求表达、沟通协商、风险防控等机制不够重视。这也从侧面说明公共责任视角下政府购买公共服务行为对制度范畴内的政策规章更为关注。

在行政责任（0.281）维度下，购买方工作人员的专业素质、购买服务过程的规范性以及购买服务方案设计的完整性三个指标分别位列重要性的前三名。表明从行政责任的角度而言，政府购买公共服务绩效评估更加侧重于购买公共服务项目前期准备和过程管理，反而对购买公共服务中的竞争、效率问题没有过度重视。可见，公共责任视角下政府购买公共服务需要从购买服务程序端上加以重视，以保证购买公共服务行为的规范性发展。

从政治责任（0.170）维度来看，该维度下的指标基本处于重要性排名的中游，但其中公众需求满足程度这一指标的重要性十分突出，在所有的指标权重排名一位，排名第二和第三的分别是购买流程的公平性、政府信息公开程度。可以说，从政治责任的角度来看，公众需求的满意度在政府购买公共服务中的重要性较高，这与公共责任的本质直接相关。从政府的性质来看，为人民服务是政府的根本宗旨，而政府购买公共服务其根本目标也在于满足公众对公共服务需求。此外，购买流程的公平性、政府信息的公开程度则体现出政府对社会组织、对社会成员的责任担当。从政治责任角度讲，政府购买公共服务实际上体现出“政府内部-政府与社会-政府与社会组织”的全方位结构，这意味着虽然从表面理解政治责任归属于政府自身的范畴，但实际上从政治责任体现出一定的流动状态，在政府和其他主体之间传导。而其他关于服务提供过程的政治责任反而因为责任传导给非政府领域则显得不那么重要。

从专业责任（0.159）维度来看，专业责任并不是十分受重视，其中的指标权重排名基本处于整个指标体系的中下游。在该维度中，承接方工作人员的胜任力水平、承接方工作人员的专业能力和承接方资金使用的合理性是最重要的三个指标。而承接方的制度设计、与服

务对象互动等方面显然不是公共责任视角下政府购买公共服务的重点对象。从结果上来看,这一点与法律责任形成了各具鲜明的特征:法律责任评估维度更侧重政府购买公共服务制度建设及政策保障环节,专业责任的绩效指标则侧重于对承接方如社会组织承接能力等方面的考察。这表明在公共责任视角下政府购买公共服务过程中,政府部门的公共责任主要体现在制度设计层面,而承接方的公共责任围绕着如何“把服务做好”而展开。

最后,从道德责任(0.099)维度来看,该维度下的指标在政府购买公共服务绩效评估中重要程度不高。该维度中最重要的三个指标是服务组织的责任文化、工作人员的服务态度、承接方工作人员的责任意识。但由于指标的绝对权重比较小,所以即使这三个指标在相对权重中较大,但对整个公共责任视角下的政府购买公共服务影响较小。为何责任的道德维度没有得到理论意义上的重视,何种原因导致道德责任维度的权重值最低?这一结果分析也意味着政府购买公共服务中责任文化、服务态度及责任意识等方面一直被人们所忽视。然而,实践中政府购买公共服务存在的道德问题,或由其引发的委托代理等问题已造成诸多不良后果。因此,进一步说明政府购买公共服务中政府及政府官员不能只是按照法规做事,不违背法规也并不是其责任的界限。

基于上述研究结果,试图挖掘这些研究结果的潜在效用,可以从以下两方面进行阐述和解释说明:第一,公共责任视角构建的评估体系应和现有绩效评估体系整合应用。在给定的评估系统的范围和目标以内,一套绩效指标应该是全面的和综合的。本研究中一个全面的和综合的绩效评估指标系统,应该囊括所有的公共责任绩效维度,以及前文所讨论政府购买公共服务绩效评估现有其他指标类型,例如,包括公众的参与程度和满意度,还应包括服务效率和成本等等。所以,研究结论实践应用最重要的是,绩效评估系统在强调某些预期结果的同时,也不应忽视其他重要的结果。第二,本研究最终的落脚点是所构建的体系如何在实践中应用从而体现其价值。政府购买公共服务绩效评估是一种有效的监督管理工具,是落实公共责任的重要手段。因此,从公共责任视角构建的评估体系,政府购买公共服务项目在实践中应用时需强调制度建设的重要性,例如,建设公众参与机制,公众参与绩效评估将改善公共服务质量,强化对公众需求负责的责任意识,并且通过公民的外部监督增强评估结果的有效性;设计一套科学的奖惩机制是加强问责的重要方面,在一定程度上可以调动相关组织的积极主动性,并防范道德风险,形成有效的激励与约束,保障公共责任。

#### 参考文献

- [1] Behn, R. D. Rethinking Democratic Accountability[M], The Brookings Institution Press, Washington, 2001.
- [2] Arie Halachmi. Performance Measurement, Accountability, and Improved Performance[J]. Public Performance & Management Review, 2002, 25(4).
- [3] Mark Bovens, Thomas Schillemans, Paul' T Hart. Does Public Accountability Work? An Assessment Tool[J]. Public Administration, 2008, 86(1).
- [4] Anna A. Amirkhanyan. What is the Effect of Performance Measurement on Perceived Accountability Effectiveness in State and Local Government Contracts?[J]. Public Performance & Management Review, 2011, 35(2).
- [5] 李洪佳. 政府购买公共服务的公共责任流失及其治理[J]. 理论导刊, 2015(12): 87-91.
- [6] 叶托, 薛婉桦. 政府购买公共服务的责任风险与问责模式[J]. 地方财政研究, 2018(4): 17.
- [7] 徐国冲, 赵晓雯. 政府购买公共服务的“公共性拆解”风险及其规制[J]. 天津社会科学, 2020(03): 83-88.
- [8] 叶托, 胡税根. 政府购买社会服务的绩效评估指标体系研究——基于德尔菲法和层次分析法的应用[J]. 广东行政学院学报, 2015, 27(02): 5-13+45.

- [9] 王春婷, 李帆, 林志刚. 政府购买公共服务绩效结构模型建构与实证检测——基于深圳市与南京市的问卷调查与分析[J]. 江苏师范大学学报(哲学社会科学版), 2013, 39(1):109-115.
- [10] 范炜烽, 许燕. 政府向社会力量购买公共服务评估指标体系构建研究[J]. 科学决策, 2020(05):1-24.
- [11] 徐家良, 许源. 合法性理论下政府购买社会组织服务的绩效评估研究[J]. 经济社会体制比较, 2015(06):187-195.
- [12] 姜爱华, 杨琼. 政府购买公共服务“全过程”绩效评价探究[J]. 中央财经大学学报, 2020(03):3-9+43.
- [13] Barbara S. Romzek. Dynamics of Public Sector Accountability in an Era of Reform. *International Review of Administration Sciences*, Vol. 66 2000, 21-44.
- [14][15] Graeme A. Hodge, Ken Coghill. Accountability in the Privatized State. *Governance: An International Journal of Policy, Administration, and Institutions*, Vol. 20, No. 4, October 2007, 695-696;685.

# **Research on the Performance Evaluation Index System of Government Purchase of Public Services from the Perspective of Public Accountability**

**Dong Yang**

**Abstract:** How to effectively construct a performance evaluation index system to help managers make correct decisions and improve performance management issues is a challenging problem. Based on the unique perspective of public accountability, this paper systematically builds the performance evaluation index system of government procurement of public services, and establishes a comprehensive analysis framework, trying to provide a basis for the performance evaluation indicators of government procurement of public services under public accountability. This paper finally determines five evaluation dimensions of administrative accountability, legal accountability, professional accountability, political accountability, and moral accountability, uses the Delphi method to screen performance indicators in two rounds, and obtains 45 formal performance indicators. High, and through the consistent test and analysis, it effectively responds to the extent to which the performance evaluation of government procurement of public services enhances or reduces public accountability. The research results further show that the importance of perfecting legal regulations, standardizing administrative functions, and providing professional services in the performance evaluation of government procurement of public services. The final destination of this research is that the index system constructed will reflect its actual effect and value in practical application. Therefore, from the research results, emphasizing the importance of system construction will be the focus of future research.

**Key Words:** public accountability; government purchase of public services; evaluation dimensions; performance indicators