

## 摘要

随着全球公共组织管理模式的变化，公共选择和新管理主义思想作为一种思潮开始注入到公共管理领域。在这一过程中，绩效评价作为一种评价和改进政府绩效管理的实用工具，逐渐在公共管理的实践中得到普遍应用。在我国政府行政改革不断深入的同时，英美等发达国家的一些行政改革经验开始引起国内政府部门以及行政学界的关注。因此，借鉴国外的先进经验和总结自身的实践，建立一个科学的、符合中国国情的政府部门绩效评价体系显得十分重要。

本文对我国政府部门的绩效评价做出客观分析，其目的是为了寻找适用于我国政府部门绩效评价的方法；在促进部门绩效评价的同时，调动部门人员个人的积极性；为我国政府树立顾客至上的服务意识提供支撑，以实现国家的长期战略目标。

本文从政府部门绩效评价的产生开始，回顾了它的发展历程，总结其已有的成绩和存在的问题。根据我国政府部门的特点，本文采纳和改进平衡计分卡，结合德尔菲法、层次分析法和模糊综合评价法等方法，通过调查问卷，对政府部门绩效进行客观的综合评价，希望借此达到抛砖引玉的作用，使政府部门从目标、制度、标准、方法等方面更加深入透彻地改革其管理模式，提高管理水平。具体的文章结构如下：

第一部分首先对政府绩效评价的相关背景进行介绍，指出其改革的迫切性和艰巨性；然后对国内外政府绩效研究现状进行概括，提炼出前沿的理论和方法；最后介绍本文的研究框架和研究方法。

第二部分首先介绍政府部门绩效评价的概念、方法、产生、发展历程、业绩和问题，在总结美国、英国、澳大利亚等发达国家政府绩效评价实践的基础上，提炼出几点经验。

第三部分介绍了平衡计分卡的产生、内容和优缺点，将其引入政府部门绩效评价，在对其进行可行性分析和优化后，建立了相应的四维指标体系。

第四部分融合了层次分析法和模糊综合评价法的优点，完成了整个政府部门绩效评价的过程，并做无量纲和无极性处理，对评价结果进行分析。

第五部分结合地方税务机关的实际情况，对三、四部分进行实证检验。

第六部分对整篇论文进行总结，指出本文的创新点，并确定日后进一步进行深入研究的方向。

**关键字：**政府部门绩效评价，平衡计分卡，层次分析法，模糊综合评价

## Abstract

With the variety of management models in global public organizations, as a trend of thought, public choice and new managerialism are beginning to infuse into the field of public management, in which performance evaluation as the practical means is applied widely to evaluate and improve government performance management gradually. While reforming the government administration of China, some related experience in developed countries like America and UK takes our government and academia's attention. So it is highly important to build scientific evaluation system for government performance through utilizing these advanced experiences and combining with our situation.

The thesis objectively analyzed our government performance evaluation in order to find the appropriate methods to evaluate government performance, inspire everyone's enthusiasm in the same time, and sustain the idea of serving the public to realize long-term strategy targets.

From the beginning of the government performance's creation, the author reviewed the whole process, summarized the achievement and deficiency and gave the comprehensive evaluation by the methods of Delphi, BSC, AHP, fuzzy comprehensive evaluation and questionnaires, according to sectors' characters. The author made use of it to encourage reforming the government management models drastically at the aspects of target, system, standard and method. The content of this thesis is following:

Part one introduced the background of government performance evaluation to reflect the reform's urgency and hardship, summarized and assessed related literature to abstract the advanced theories and methods, supplied the thesis's research frame and means. Part two introduced the concept, means, creation, development, achievement and deficiency of government performance evaluation and abstracted some practical experience of developed countries such as America, UK, and Australia. Part three introduced BSC to government performance evaluation and built the corresponding index system from four dimensions after analyzing its feasibility. Part four recommended AHP and fuzzy comprehensive evaluation combined with the index system above to accomplish the evaluating process and analyzed evaluating result with the non-dimension and non-polarity. Part five took the revenue sector as example to prove academic result. Part six reviewed the thesis, pointed out its innovation points and ascertained research orientation in the future.

**Key words:** government performance evaluation, balanced scorecard(BSC), analytical hierarchy process(AHP), fuzzy comprehensive evaluation

## 独创性声明

本人生命，所呈交的论文是本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得武汉理工大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签名: 赵璐 日期: 2007.11.24

## 关于论文使用授权的说明

本人完全了解武汉理工大学有关保留、使用学位论文的规定，即学校有权保留、送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

（保密的论文在解密后应遵守此规定）

签名: 赵璐 导师签名:  日期: 2007.11.24

## 第1章 绪论

### 1.1 研究背景和意义

#### 1.1.1 研究背景

随着经济全球化的迅猛发展，传统的官僚制模式的弊端日益凸显，导致各国政府面临严重的财政危机、管理危机、信任危机以及合法性危机。西方各国为了摆脱政府管理的困境，纷纷开展了大规模的政府改革运动，延续近百年的传统行政模式面临着极大的挑战。

公共组织管理模式（包括组织结构、管理方式等）已经发生了重大的变化，公共选择以及新管理主义思想作为一种思潮以全新的面貌开始注入到公共管理领域。在这一过程中，绩效管理和绩效评价作为一种评价和改进政府绩效的管理实用工具，逐渐在公共管理的实践中得到广泛应用。

在我国，温家宝总理在十届全国人大三次会议所做的政府报告中明确提出：“抓紧研究建立科学的政府绩效评价体系”。党的十六大报告指出深化行政管理体制改革，进一步转变政府职能，改进管理方式，形成行为规范、运转协调、公正透明、廉洁高效的行政管理体制。党的“十一五”规划强调要深化政府机构改革，优化组织结构，减少行政层级，理顺职责分工，推进电子政务，提高行政效率，降低行政成本。

这一系列的改革目标，客观上对现行的一些绩效评价方式和方法提出了更高的要求。同时，以往对公共组织单一的评价方式越来越不符合当今社会的发展，因此在实践中迫切需要寻找一种适合我国国情的绩效评价方法。政府绩效考核不同于企业组织——企业可以通过工作分析来进行考核指标的选取；但政府部门的工作岗位众多，并且没有一个全国统一的岗位职责。随着我国改革开放的力度不断加大，以及政府行政改革实践的不断深入，英美等西方发达国家的一些行政改革经验开始引起国内政府部门以及行政学界的关注。因此，借鉴国外的先进经验和总结自身的实践，建立一个科学的、符合中国国情的政府部门绩效评价体系才是改革的目标。

从目前的情况看，绩效管理在公共管理领域的应用还存在着诸多困难。这除了与政府本身的制度特征复杂性以及我国现行的行政体制有关以外，其工作

性质的模糊性和难以量化也是绩效考核存在困境的重要原因。但是，经过公共管理专家以及公共组织实践者们的不断探索和努力，其应用范围日趋扩展，实施的效果也日趋显著。正如戴维·奥斯本和彼得·普拉斯特里克在他们所著的《摒弃官僚制：政府再造的五项战略》中指出：“尽管比起其它途径而言，绩效管理改进绩效的过程通常更为缓慢，他的事实也需要更长的时间（通常需要三年或四年），但是世界范围内的政府改革者都发现，使用绩效管理可以使政府的生产效率获得显著、持续的增长<sup>[1]</sup>。”

### 1.1.2 研究意义

本文结合发达国家的经验，运用平衡计分卡、德尔菲法、层次分析法和模糊综合评价法等方法对政府部门的绩效评价做出尽量客观的分析，寻找适用于我国的政府绩效评价方法，为我国政府改善管理、提高服务水平提供一点建议。具体来说，本文的研究具有以下几个方面的意义：

第一，有利于建立适用于我国的政府部门绩效评价体系。虽然我国政府经过了几次职能转变，管理水平有了很大的提高，并且也采用了一系列先进的管理方法和管理手段，但是，目前我国政府的绩效评价体系仍然存在许多困难，本文正是致力于解决其中的一些客观问题。

第二，有利于我国实现长期的战略目标。本文倡导把平衡计分卡运用于政府绩效评价体系，它本身的功能体系不仅可以考察我国政府部门当前的绩效，还注重考察其长远绩效，以实现我国政府及其所引导下的我国社会、经济与环境的可持续发展。这也是对科学发展观的最好诠释。

第三，有利于政府部门树立顾客至上的服务意识。从计划经济向市场经济转型的过程中，随着政府职能的转变，其在国家经济生活中的角色不再是领导者，而是服务者。然而，某些政府部门服务意识仍然欠缺，没有将其服务对象——社会公众的地位提高到应有的水平。本文通过在政府绩效评价的过程中引入平衡计分卡系统，增加了社会公众维度的内容，有利于政府部门树立顾客至上的服务意识。

第四，有利于调动我国政府部门人员的积极性。本文虽然是对部门绩效的评价，但是，它的观念和方法同样适用于行政人员个人绩效的评价。同时，对部门绩效指标的明确，能够更为全面、更为科学地衡量个人的工作绩效，从而也就能够使其明确自己的工作职责，调动其工作积极性，使我国政府在社会政治、经济生活中发挥更为重要的作用。

## 1.2 国内外研究现状

绩效包括组织绩效和个人绩效两个层面。相应地，绩效评价也包含组织绩效评价和个人绩效评价两个层面。本文主要针对组织绩效评价尤其是政府部门绩效评价进行研究。

公共部门绩效评价是经济学、管理学和公共行政学等领域的学者们所关注的焦点。它之所以成为关注的焦点，在布兰克（Jos L.T. Blank）等人认为主要是因为“在绝大多数国家中公共部门的绝对规模意味着它的绩效会对经济的整体状况产生深远的影响。这种影响通过两种方式表现出来，以最终产出的形式提供给消费部门，如教育和卫生；以中间产出的形式用作私有生产部门的投入，如能源和基础设施”。下面分别就国内外政府绩效评价研究现状进行总结。

### 1.2.1 国外研究现状

在评价理论上，凯瑟琳·纽科默等论述了对业绩进行测评与报告来提高公共组织的绩效。对业绩的测评与报告是通过识别项目的后果和测评成就来完成的，它能将公共机构的侧重点从投入与产出转移到后果上<sup>[3]</sup>。戴维·奥斯本和彼得·普拉斯特里克认为，凡是取得了高绩效的组织都是那些高度重视使命、组织目标及绩效等问题的组织。只有明确了组织使命、目标等问题，组织的其他努力才有正确的方向，提高效率的行动才有意义<sup>[1]</sup>。

在评价方法上，平衡计分卡是组织绩效评价的趋势。1992年，卡普兰和诺顿总结这项研究的成果，在《哈佛工商评论》上发表了关于平衡计分卡的第一篇论文《平衡计分卡——提升绩效的衡量方法》<sup>[4]</sup>；1993年又在该刊发表了《将平衡计分卡付诸实践》，揭示了平衡计分卡的指标必须与组织的战略紧密结合起来<sup>[5]</sup>；1996年又发表了《把平衡计分卡作为战略管理的基石》，总结了平衡计分卡在实务中的运作，许多采用平衡计分卡的公司将其当作重要管理流程的主要框架<sup>[6]</sup>。在此基础上，两人于1996年和2001年分别出版了《平衡计分卡——化战略为行动》和《战略核心组织——以平衡计分卡帮助企业成长》两本专著来详细论述平衡计分卡的内涵<sup>[7]</sup>。

1996年，非盈利组织和政府领域引入平衡计分卡还处于萌芽时期。经过多年的发展，它已经演化成为一种组织战略管理的有效工具。美国交通运输部（DOT）下的采购部是最早采用平衡计分卡的政府机构之一<sup>[8]</sup>。美国各州也根据各自特点，设置了不同的评价指标和标准，北卡罗纳州夏洛特市则是运用平衡

计分卡获得成功典型例子<sup>[9]</sup>。美国著名绩效研究机构 Compel 研究所建立了一套包括财政管理、人事管理、信息管理、领导目标管理和基础设施管理五个方面内容的政府绩效评价体系<sup>[10]</sup>。平衡计分卡在澳大利亚、新加坡等国的政府部门中也得到了认可和应用。澳大利亚 Cockburn 和 Melville 市的平衡计分卡绩效衡量系统获得了奖励。新加坡地区法院系统建立了平衡计分卡，是在世界司法领域中最先被使用的计分卡。

在评价实践上，20世纪70年代，受经济影响，世界各国政府出现了财政和信用危机，如何提高公共部门绩效成为人们关注的问题<sup>[11]</sup>。在此背景下，美英等国作为这场运动的领导者，其政府十分重视并以积极的态度来开展绩效评价改革。

实际上，美国在20世纪上半叶就已经开始了绩效评价活动。多数学者将1949年联邦政府推行绩效预算作为政府绩效评价的开端<sup>[12]</sup>。美国自1978年的卡特政府以来，历任总统都十分关注政府部门的绩效管理问题。到克林顿时期，这种重视程度已达到了顶峰。美国联邦政府制定了《国家绩效评价报告》，国会通过了《政府绩效和结果法》，将政府绩效评价制度化、法定化，使其不因行政首长的更迭而发生变化<sup>[12]</sup>。布什政府于2003年开始推行部门绩效“报告卡”制度，设立绩效基准和等级评价标准，以一种类似于危机管理的方式展示公共责任，激发公众的兴趣和关注。

英国政府绩效评价始于1979年的“雷纳评审”，它是指对政府部门工作特定的调查、研究、审视和评价活动，评审的重点是政府机构的经济和效率水平。后来，政府机构内部的评审发展到有社会参与的评价，评价内容侧重公共服务和服务质量，评价结果公开并直接向公众负责。1986年，英国政府各部门为评价拟定的绩效指标总数为1220个；1987年，这一数字上升到1810个；到1989年飞速增长到2327个。此后，指标总数基本上稳定在这一水平<sup>[13]</sup>。

在美英两国的带动和示范下，政府组织绩效评价在其他国家也得到广泛应用，并发挥了重要的作用。但是由于绩效评价是行政管理公认的难题之一，所以各国政府绩效评价的理论研究和实践探索仍在继续进行之中。

### 1.2.2 国内研究现状

改革开放以后，绩效评价的理论和方法才逐渐引入我国。我国对政府部门绩效的评价起始于干部人事制度。因此，很多相关研究是针对政府及领导干部政绩的考核评价<sup>[14]</sup>。我国学者关于政府及领导干部绩效评价的研究主要涉及评

价指标体系与评价方法两个方面。韩强就如何构建和完善党政领导干部考核评价指标体系的有关问题进行了深入探讨，并依据《党政领导干部选拔任用工作条例》，提出以德、能、勤、绩、廉为基本内涵，建立有中国特色的党政领导干部考核评价指标体系<sup>[15]</sup>。倪星、余凯根据效率与公平并重、效率与民主兼顾、经济增长与社会发展同步等价值标准构建了涉及国民经济、人民生活、科教文卫、生态环境、社会治安等多方面因素的政绩评价指标体系<sup>[16]</sup>。韩兆洲构建了由2大类中7类共30项指标组成的政府政绩综合评价指标体系，并在广东省政府部门开展了实证研究<sup>[17]</sup>。范柏乃、朱华以“科学发展观”和“执政为民”为指导原则，从行政管理、经济发展、社会稳定、教育科技、生活质量和生态环境等领域，构建了地方政府绩效评价体系<sup>[18]</sup>。姜晓萍、马凯利对“360度考核”模型予以高度的评价，强调实行公开考核，促进考核主体与被考核者之间的信息互动<sup>[19]</sup>。

平衡计分卡也被用于我国的政府部门绩效评价<sup>[20]</sup>。其中比较有代表性的文章有中国行政管理学会张定安博士的《平衡计分卡与公共部门绩效管理》，文章探讨平衡计分卡的特点，引入公共部门绩效管理的可行性和可能产生的影响<sup>[21]</sup>；北京大学张国庆、曹堂哲的《“平衡计分卡”与公共行政执行的有效性》概括了平衡计分卡内容、发展，并对在我国公共行政中的有效性进行了探讨<sup>[22]</sup>；西南财经大学文薇的《基于平衡计分卡的政府绩效管理系统构造研究》总结了我国政府绩效发展的历程和国外运用平衡计分卡的先进经验<sup>[23]</sup>。

另外，各种定量分析方法也被应用于政府绩效评价。刘晓辉、王红艳提出以定量和定性相结合的方法来考核政府部门，确定科学的考核指标应从以下几个方面做起：首先，考核指标的制定应由考核专家和考核对象共同参与；其次，考核指标的制定应有科学依据；再者，考核指标的设立应实事求是，量化的均用数量化表示<sup>[24]</sup>。彭国甫等人还应用层次分析法确定了政府绩效评价的指标权重，应用数据包络分析进行了政府绩效相对有效性评价<sup>[25]</sup>。模糊综合评价方法也同样被应用于政府绩效评价<sup>[26]</sup>。随着研究的不断发展，我国学者对政府绩效评价的理念和方法的研究也更加深入<sup>[27]</sup>。

我国在评价实践方面起步较晚<sup>[28]</sup>，但各类比较先进的绩效评价系统在政府及公共部门都得到了应用。2002年，厦门市思明区政府与厦门大学共同开发了“公共部门绩效评价系统”，该系统由评价主体、评价维度和评价指标三方面内容构成。2004年，国家人事部《中国政府绩效评价研究》课题组推出了一套适用于中国地方政府部门的绩效评价指标体系，社会保障、社会稳定、廉洁状况

和行政效率等多个指标被纳入考核范围<sup>[27]</sup>。国家行政学院在研究欧盟成员国使用的多种绩效评价模型的基础上，结合我国国情，构建了中国通用绩效评价框架（CAF）。该模型包括了促进和结果两大要素，共大 9 标准，27 个次级指标。模型在哈尔滨铁路检察院和厦门市思明区政府进行了试点，已初步取得成效<sup>[28]</sup>。2004 年底，中国第一家由民间发起的地方政府绩效评价机构——中国地方政府绩效评价中心在兰州大学成立。2005 年 3 月，该中心提交了《甘肃省非公有制企业评议政府工作报告》<sup>[29]</sup>。

### 1.3 研究内容和研究方法

本文的研究从政府部门绩效评价的产生开始，回顾它的发展历程，总结其已有的成绩和存在的问题。根据政府部门的特点和国外政府部门的成功经验，采纳和改进平衡计分卡，结合德尔菲法、层次分析法和模糊综合评价法等方法，对政府部门绩效进行综合评价，希望借此达到抛砖引玉的作用，使政府从目标、制度、标准、方法等方面更加深入地改革其管理模式，提高管理水平。

具体说来，本文通过六个部分对如何建立我国政府部门绩效评价系统进行论述和介绍：

第一部分首先对政府绩效评价的相关背景进行介绍，指出其改革的迫切性和艰巨性；然后对国内外政府绩效研究现状进行总结，提炼出前沿的理论和方法；最后介绍本文的研究框架和运用的方法。

第二部分对政府绩效评价的相关概念和内容进行介绍，包括政府部门绩效评价的概念、方法、产生、发展历程、业绩和问题；随后，在总结美国、英国、澳大利亚等发达国家政府绩效评价实践的基础上，提炼几点经验。

第三、四部分是文章的主体，主要运用平衡计分卡、德尔菲法、层次分析法和模糊综合评价等方法构建了政府部门绩效综合评价体系。第三部分介绍平衡计分卡的产生、内容和优缺点，将其引入政府部门绩效评价，建立相应的四维指标体系。第四部分首先介绍层次分析法和模糊综合评价法，然后将其与前面的指标体系相结合，完成整个政府绩效评价的过程，最后经过无量纲和无极性处理，对评价结果进行分析。

第五部分结合河北省沧州市国税局 2006 年度绩效状况的实例，对第三、四部分进行实证检验。

第六部分对整篇论文进行总结，指出本论文的创新点，并确定日后进一步进行深入研究的方向。

## 第2章 政府部门绩效评价的方法与问题分析

### 2.1 政府部门绩效评价的相关概念

#### (1) 政府部门绩效评价的含义

由于对绩效评价的研究广泛地涵盖在管理学、组织行为学、人力资源管理中，学者们的视角不同，从而也就造成了对绩效评价含义的不同理解。人们对绩效评价的理解主要有以下两种：

一是绩效评价是评价雇员绩效的系统。如“绩效考核是对员工一个既定时期内对组织的贡献做出评价的过程，从数量两方面对其工作的优缺点进行系统的描述，绩效评价是一个复杂的过程<sup>[16]</sup>。”

二是绩效评价是评价组织绩效的系统。如“绩效评价是评价达到预定目标的过程，包括信息资源转化为物品和服务输出的效率、输出的质量，提供给顾客的效果，顾客满意程度和结果与期望目的相比的后果，政府在对项目目标特定贡献方面运作的有效性<sup>[31]</sup>。”

一个完整的绩效评价体系应该兼顾不同的层面，既包括雇员层面微观的绩效评价，也包括组织层面宏观的绩效评价。因此，政府绩效评价也应该包括微观、宏观两个层面。在微观层面上，如“绩效评价是基于事实，有组织地、客观地评价组织内每个人的特征、资格、习惯和态度的相对价值，确定其能力、业务状态和工作适应性的过程<sup>[8]</sup>。”宏观层面上，如“绩效评价是通过系统的方法、原理来评定和测量行政部门整体的工作行为和工作成果，它是管理者与员工之间的一项管理沟通活动<sup>[32]</sup>。”

由于政府活动的特殊性，在宏观上又常有几个部门临时联合起来或是临时从一个部门抽调部分人员完成某一项目的情况出现。因此，政府绩效评价可以分为三种类型：组织部门的绩效评价、项目的绩效评价和个人的绩效评价<sup>[33]</sup>，本文主要研究第一种情况。

#### (2) 政府部门绩效评价的比较分析

政府部门绩效评价的特点可以从与两个特定对象的比较中体现出来，一个是跟传统的政府财务审计相比，一个是与企业的绩效评价相比。

##### ① 与传统政府财务审计相比

政府绩效评价从根本上说是政府绩效审计的一个组成部分，具有政府绩效审计的许多特征，与同属于政府审计范畴的传统政府财务审计相比，它具有以

下一些特点：

一是评价目的。

政府部门绩效评价的目的是评价政府的受托管理责任，确定评价客体是否进行有效的管理，把重点放在经济性、效率性、效果性等方面；而传统的财务审计主要是审查受托财务责任，重点审查财务收支的准确性、合规性和合法性。

二是评价程序与方法。

传统审计采用通用的财务报表审计的方法来完成，而政府绩效评价根据每一评价项目的特点，分别制定不同的绩效评价方案和方法。另外，由于绩效评价的滞后性，又势必要求进行追踪评价，以取得较为全面而准确的绩效评价结果。例如，我国审计署就对以前年度的被评单位的绩效进行了追踪评价，并在年度的审计报告中公布了各部门对以前年度审计出的问题的整改结果<sup>[14]</sup>。

三是评价报告。

政府绩效评价报告的类别较多，有的是可以向社会公众公布的绩效评价结果，也有的是只能向政府、人大或某些特殊人员提交的专题报告或专项报告。在报告的形式上，政府绩效评价也更为灵活，它不仅可以采用文字叙述的形式，也可以根据实际需要采用图表、图片等形式来对绩效评价结果进行报告。另外，这些报告一般包括更多的讨论、分析、合理的论证以及对被评价单位在管理方面的建议等。

② 与企业绩效评价相比

随着新公共管理理论（其创始者奥斯本认为，只要政府部门鼓励公务员的企业家精神，那么公务员的行为就会像企业家一样的有效率和低成本，从而实现政府“少花钱多办事”的理想，实现公众利益最大化<sup>[15]</sup>）的发展，人们越来越要求政府像企业绩效评价一样对自身进行评价。但是，政府绩效评价与企业相比大有不同：

一是评价主客体。

企业绩效评价的评价主体可以是董事会，也可以是利益相关者。而政府绩效评价的主体只能是政府绩效审计机关。对于政府工作的绩效，不同的利益相关者可能会有不同的要求，而且这些要求的差别往往很大。政府绩效评价的客体是政府部门、政府项目和政府个人，他们更加注重社会效应，这与企业绩效评价的客体要求是不同的。

二是评价指标。

企业以追求利润为目标，在企业的绩效评价中就设计了诸如投资回报率、经济附加值之类的指标来评价其绩效。而政府是一个非营利性的组织，其基本职能是维护国家安全和社会稳定、促进经济社会发展、管理公共社会事务等。因此，其评价指标就不以财务指标为主，而更多地考虑了一些非财务性的指标，其定性的指标也更多一些。

## 2.2 政府部门绩效评价的方法

19世纪以来，企业绩效评价的方式和方法深深地影响着政府绩效评价的发展。下面介绍一些各个不同时期影响政府绩效评价的方法。

### 2.2.1 财务评价方法

在20世纪初的生产管理阶段，企业的目标主要是通过提高生产效率，来追求利润最大化。由于不断地通过外部融资扩大生产规模，庞大的投资使企业最为关心以投资报酬率为为核心的财务指标。这一时期出现了“投资报酬率”法和“杜邦系统图”。应用最为广泛的评价指标有投资报酬率、权益报酬率和利润等财务指标。这些方法也为早期公共组织所利用<sup>[15]</sup>。

#### (1) 投资报酬率

投资报酬率是会计上的一个衡量指标，被看作是代表部门投资的真实经济报酬率。但是，除了极特别的情况外，会计上的投资报酬率并不等于部门内在的资产报酬率<sup>[20]</sup>。因此，以会计投资报酬率来代替实际资产报酬率，以会计投资报酬率来衡量部门的经营业绩和评价部门的经济效益是不合适的。

#### (2) 杜邦系统图

杜邦分析体系最大的特点是把众多的财务指标有机地结合在一起，突出各指标之间的联系。因此，可以深入地分析企业在具体财务活动中的优势与不足。而我国的综合财务指标分析体系则着重于综合评价企业财务活动，反映企业在盈利、营运和偿债等方面的情况，供相关单位和利害关系人参考。

杜邦分析体系有三大分支：一是从销售利润率开始，一直往下层去，这条分支反映组织在盈利能力方面的相关财务指标；二是从总资产周转率开始，一直往下层去，这条分支反映组织在营运能力（资产利用能力）方面的相关财务指标；三是从权益乘数开始，一直往下层去，这条分支反映组织的偿债能力，同时也是资本结构方面的相关财务指标<sup>[20]</sup>（如图2-1所示）。

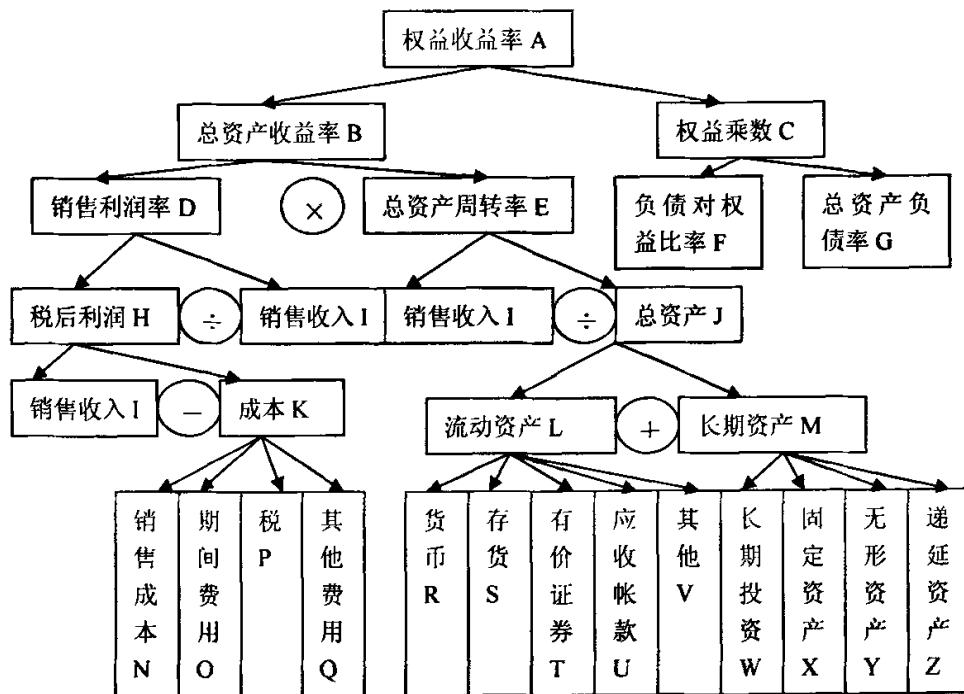


图 2-1 杜邦系统图

[资料来源：文薇. 基于平衡计分卡的政府绩效管理系统构造研究：[硕士学位论文]. 成都：西南财经大学行政管理专业, 2007: 24]

## 2.2.2 价值评价方法

站在股东的角度来评价企业业绩的价值模式，能够有效地将企业战略与日常业务决策与激励机制有机地联系在一起，最终为股东创造财富。它所使用的评价指标主要是经过调整的财务指标，或根据未来现金流量得到的贴现类指标。经济增加值（EVA）是最具代表性的。

经济增加值是经过调整的税后净经营利润减去该公司现有净资产使用的资本费用后的余额。它是扣除了所有使用的资本成本后的经济利润<sup>[36]</sup>。经济增加值是一个致力于创造股东价值财富的财务管理系统，他要求企业各级管理人员充分注重资本的使用效益，全力去寻找增加企业经济增加值的各种途径和方法，不为财务报表上一时的利润所迷惑。

但是，它无法摆脱财务指标所固有的缺陷——短期导向，体现在它无法有效地反映顾客、雇员、供应商等利益相关者对促进企业将来绩效的重要作用。在提高经济增加值的同时可能损害了利益相关者的利益，从而最终破坏企业的

长期价值，因而迫切需要做出进一步地改进，即融合非财务业绩评价方法。目前，虽然大量的公共组织已经引用经济增加值评价模式，但仍然摆脱不掉它是一种基于财务和价值增值的评价方式，对于非营利性占主要地位的政府部门仍有一些不适应性。鉴于经济增加值的优点，可以把财务作为一个维度，将它运用其中。

### 2.2.3 平衡评价方法

20世纪70年代以后，市场竞争日益激烈，企业面临的环境日益复杂。企业注意到必须将各职能部门协调统一。本阶段最大的突破就是引入了非财务指标，通过指标间的各种平衡关系以及关键指标的选取来体现企业不同利益相关者的期望，从而实现企业价值最大化的目标。这一时期出现了霍尔的四尺度论、业绩金字塔模型、“3E”评价法、管理标杆法、平衡计分卡（第4章种介绍）等评价方法。

#### （1）四尺度论

罗伯特·霍尔（Robert Hall）认为，评价公司的业绩需以四个尺度为标准，即质量、作业时间、资源利用和人力资源的开发。

霍尔把质量分为外部质量、内部质量和质量改进程序三种。外部质量包括顾客调查情况、服务效率、保修及可靠性等。内部质量包括总产量、生产能力、检验比率以及残品和返工率等。质量改进程序是公司组织采用的确保高水平的内在和外在质量的程序或一系列的公式化的步骤。霍尔认为作业时间是把原材料变为成品的时间段。其中包括：工具检修时间、设备维修时间、改变产品和工序设计的时间、项目变更时间、工具设计时间和工具建造时间等。资源使用尺度用以计量特定资源的消耗和与此相关的成本，如直接人工、原材料消耗、时间利用和机器利用情况。前两项指标是制造产品和提供劳务的直接成本，后两项既包括直接成本因素<sup>[37]</sup>。霍尔提出公司需要有一定的人力资源贮备和能恰当评价和奖励雇员的管理系统。

霍尔把质量、时间和人力资源等非财务指标导入公司的业绩评价系统，并认为企业组织可以通过对上述四个尺度的改进，减少竞争中的风险。他同时承认，要求公司做出全方位的改变是困难的，只能在一段时间内取得四个方面的逐渐改进。需要注意的是，任何指标的改进不应以牺牲其他指标为代价，如作业时间的改进不应该以降低质量为代价，质量方面的改进也不应该以牺牲资源为代价。

## (2) 业绩金字塔模型

凯文·克罗斯 (Kelvin Cross) 和理查德·林奇 (Richard Lynch) 提出了一个把公司总体战略与财务和非财务信息结合起来的业绩评价系统。为了强调总体战略与业绩指标的重要联系，他们列出了一个业绩金字塔<sup>[38]</sup>。在业绩金字塔中，公司总体战略位于最高层，由此产生具体战略目标，并向企业组织下层逐级传递，直到最基层的作业中心。有了合理的战略目标，作业中心就可以开始建立合理的经营效率指标，以满足战略目标的要求（如图 2—2 所示）。

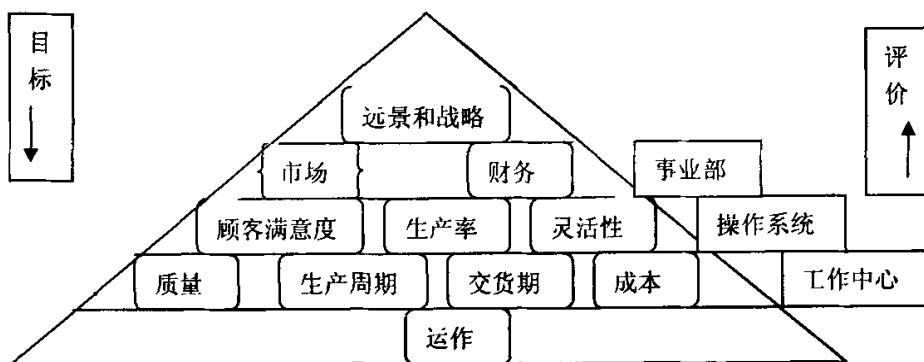


图 2—2 绩效金字塔

[资料来源：白蔚秋. 正确评价企业的经营业绩. 工业企业管理. 2001(5): 75]

克罗斯和林奇的业绩金字塔的循环过程揭示了公司持续发展的能力，为正确评价公司业绩做出了意义深远的重要贡献。业绩金字塔的主要缺点是在确认组织学习的重要性上是失败的。而在竞争日趋激烈的今天，对组织学习能力的正确评价尤为重要。因此，虽然这个模型在理论上比较成型，但实际工作中采用率较低。

## (3) “3E”评价法

20世纪80年代初，英国由雷纳爵士负责的效率小组提出“3E”评价模型。不久，英国审计委员会将“3E”指标纳入到绩效审计框架中，并运用于地方政府以及国家健康服务系统的管理实践。此后，“3E”模型逐渐成为西方国家政府部门绩效评价的基本框架。

所谓“3E”指的是经济(Economy)、效率(Efficiency)和效益(Effectiveness)。其中，“经济”是指以最低可能的成本供应与采购维持既定服务品质的公共服务项目；“效率”是指投入与产出之比，通常包括每项服务的单位成本、服务水平的提供、活动的执行等；“效益”则是指公共服务实现目标的程度<sup>[39]</sup>。

但是，政府部门不是一个盈利组织，在经济目标之外，还有许多更为重要

的目标，比如公平、民主、社会进步等，而这些目标是“3E”评价法所不能反映的。这就导致在经济获得前所未有的发展的同时，还普遍存在着失业、贫穷、疾病、无知和绝望，从而引发各种社会危机，构成了对现有政治制度的根本威胁。为了弥补这一不足，学者们后来加入了“公平”（Equity）指标，从而发展为“4E”评价法。但是这些改进和完善都无法从根本上消除“3E”评价法所存在的不足。

#### （4）标杆管理法

标杆管理法是20世纪80年代发展起来的一种新型的绩效评价方法，它是指不断寻找和研究业内外一流的组织的最佳实践，并以此为标杆，将本组织的产品、服务和管理与之进行定量化的考核和比较，分析这些标杆组织达到优秀水平的原因，结合自身实际加以创造性地学习、借鉴，并选取最佳策略，从而赶超一流组织或创造高绩效的不断循环提高的过程<sup>[40]</sup>。

标杆管理法是一种直接的、片段式的、渐进的管理方法。因为组织的业务、流程、环节都可以解剖、分解和细化，因此，组织不仅可以寻找整体最佳实践来作为标杆，也可以寻找优秀“片断”来作为标杆。同时，标杆管理法所具有的渐进性可使组织从初级到高级，分阶段确立不同的标杆，循序渐进地进行绩效改善<sup>[41]</sup>。指标体系的内容在一定程度上引导着政府部门努力的方向，使政府部门全面考虑自身在社会中应该承担的责任，从而对社会的全面发展起到领导作用。

## 2.3 政府部门绩效评价的发展、业绩和问题

### 2.3.1 政府部门绩效评价的发展历程

我国学术界对公共组织绩效评价的研究始于20世纪30年代，最初是作为行政学理论框架的一个部分来介绍和研究的。后来一些地方政府设立了行政效率促进委员会，促使学术界对行政效率进行较为深入的专门研究，并取得了一些积极的成果。从我国地方政府绩效评价的初衷来看，其发展历程大体可分成四个阶段：

第一阶段：以提高政府机关工作效率为目的的政府绩效评价。这一阶段，主要是根据精简、统一、效能的原则，转变政府职能，建立办事高效、运转协调、行为规范的政府管理体制。为实现这一目标，从20世纪80年代初期开始，我国就开始在全国范围内推广政府部门的目标责任制，并以绩效考核作为政府

内部管理控制的手段<sup>[12]</sup>。

第二阶段：以改善政府及行业服务质量、提高公众满意度为目的的政府绩效评价。随着我国体制改革的深化，要求各级政府部门不断更新思想观念和工作方法，不断提高服务质量和水平。在这一背景下，许多地方政府相继开展了以改善政府服务质量、提高公众满意度为目的的绩效评价活动。沈阳、珠海、南京、杭州等地纷纷开展了万人评议政府的活动，山东、江苏等地政府部门也纷纷开展了社会服务承诺活动。

第三阶段：以建设效能政府、全面提高政府管理质量和能力为目的的政府绩效评价。2003年2月，中共十六届二中全会审议通过了《关于深化行政管理体制和机构改革的意见》，指出通过机构改革建立全新的行为规范、运转协调、公正透明、廉洁高效”的行政管理体制，更好地为改革开放和社会主义现代化建设服务<sup>[42]</sup>。各级地方政府围绕着这一要求，对政府及其部门进行改革，全面提高政府部门的管理质量和能力。

第四阶段：根据广义政府绩效概念内涵，以强化公共治理为目的的政府绩效评价活动。广义的政府绩效不仅仅表现为政府的行政结果与行政投入之比，还应当包括公民满意度和地方发展战略机制两个重要方面。从2004年的全国“两会”和新一轮的五年计划编制中，大家已经能够清晰地看到“以人为本、可持续发展、社会经济协调发展、统筹城乡一体化进程”等信息<sup>[42]</sup>。这就要求改变以往那种以GDP增长为主要指标的政府绩效评价方法，对“绿色GDP”的追求逐渐成为衡量政府绩效的重要指标。2006年9月，首份《中国环境经济核算2004》绿皮书正式公布<sup>[42]</sup>。

### 2.3.2 政府部门绩效评价的业绩和问题

#### (1) 政府部门绩效评价的业绩

通过积极探索和实践，我国目前在政府绩效评价的理论和实践上都取得了一定的成果。在理论研究上，影响较大的是国家人事部《中国政府绩效评价研究》课题组在总结国内外相关指标体系设计思想和技术方法的基础上，提出了一套适用于我国地方政府的绩效评价指标体系。

这是我国学者在政府绩效评价领域的第一次完整而全面深入的研究，是我国政府绩效评价研究的最新成果。该体系共分三层，包括了职能指标、影响指标和潜力指标这3个一级指标，11个二级指标和33个三级指标构成<sup>[43]</sup>（如表2-1所示）。

表 2-1 我国政府部门绩效评价指标框架

一级	二级	三级
影响指标	经济	人均 GDP、劳动生产率、外来投资占 GDP 比重
	社会	恩格尔系数、人均预期寿命、平均受教育程度
	人口与环境	非农业人口比重、人口自然增长率、环境与生态
职能指标	经济调节	GDP 增长率、财政收支状况、城镇登记失业率
	市场监督	法律法规完善程度、执法状况、企业满意度
	社会管理	贫困人口占总人口比例、刑事案件发案率、生产与交通事故死亡率
	公共服务	基础服务建设、信息公开程度、公众满意度
潜力指标	国有资产管 理	国企资产保值增值率、国有资产占 GDP 比重、国企利润增长
	行政效率	行政经费占财政支出比重、行政人员占人口比重、信息管理水平
	人力资源	领导班子团队建设、行政人员受教育程度、人力资源开发战略规划
	廉洁	机关工作作风、腐败案件涉案人员占行政人员比重、公众评议结果

[资料来源：王伟. 基于平衡计分卡的公共组织绩效评价研究：[硕士学位论文]. 天津：天津大学行政管理专业，2005：18]

在实践方面，我国政府部门的绩效评价实践主要呈现出以下一些特点：

一是评价指标以经济指标为主。GDP、财政收入、招商引资、项目投资等经济性评价指标都在我国政府的绩效指标体系中占据着重要地位。

二是与经济社会全面发展有关的指标逐渐受到重视。近年来，我国部门地方政府逐渐在绩效评价指标中引入了对资源控制、环境保护、社会保障等方面的评价。

三是评价指标与当前党委政府中心工作密切相关。评价形式及方法一般以党委或政府文件的形式下发，考核内容每年都会有所变化。许多地方把当前政府中心工作纳入考核之中，一些重点工程、为民办实事项目等成为当年考核的主要内容。

四是评价主体逐渐多元化，部分地方政府开始尝试建立群众参与机制。我国的政府评价一般是由上级政府对下一级政府的目标完成情况进行评价，有些

政府职能部门还组织对各自负责的专项工作进行评价，同时党委组织部门按照干部管理权限对政府领导班子及领导成员进行以定性为主的年度评价。

## （2）政府部门绩效评价的问题

虽然近年来我国在政府绩效评价方面取得了一定的成绩，但是总的来说，还存在着许多不足和问题。归结起来主要表现在以下几个方面：

一是政府绩效评价的理论和实践发展滞后。首先，在理论研究方面，政府绩效评价的理论体系准备严重不足：一方面是由于政府绩效评价理论涉及了公共管理学、行政管理学、企业管理学、社会学、经济学、统计学等多种学科知识，而我国对公共管理学的研究刚刚起步，具备以上各方面专业知识的综合性人才非常缺乏；另一方面是受到技术、外语能力等多方面条件的限制，我国理论界对国外先进经验的系统性介绍不足。其次，长期以来我国政府部门的绩效评价主要是以对公务员个人的评价为主，而缺乏对政府组织层面的评价。

二是政府绩效评价缺乏政策指导和法律依据。当前我国的政府评价多半还处于自发状态，没有像美国、英国那样建立起专门的政府绩效评价的政策和法律法规，导致了各级政府的绩效评价往往带有一定的盲目性和随意性，各自的绩效评价指标体系所包含的指标互不统一，难以进行比较和经验交流。

三是政府绩效评价标准难以确定。在绩效评价理论中，有人主张采用标杆定位法，以行业中的最佳实践者的业绩为标准进行绩效评价；有人主张采用经济附加值法，以经济附加值来评价政府部门的绩效；有人主张采用利益相关者评价法，以政府对各利益相关者利益的满足程度来评价政府部门的绩效；有人主张采用平衡记分卡，从财务、顾客、内部经营过程和学习与成长四个方面对政府的绩效进行评价。以上所列示的几种绩效评价方法的产生都是缘于对企业或公司绩效进行评价的需求，并没有专门针对政府绩效评价方法的理论研究。因此，在对政府绩效进行评价时，就需要对以上方法进行借鉴与发展。

四是政府绩效评价资料难以获得。我国政府部门尚未形成评价而专门收集数据资料的习惯，更不用说建立信息收集中心。部分部门有自己的信息系统，可能对政府服务所提供的来源，如某部门和数量较易于得到和分析，但对这些服务的质量、产出的效果、公众的满意程度这些材料不但难以计量也不易得知，甚至有些效果在短期内并不显现，难免会鼓励政府的短视行为。

五是政府绩效评价的指标结构不合理。这体现在三个方面：重经济增长指标，轻社会发展指标；重短期产出，对长期效益关注不够；重正向指标，轻负向指标。当前我国在选择政府部门的评价指标时，只有加分项而没有减分项<sup>[44]</sup>，

使政府不能清楚的知道有哪些事情是自己不能或不必做的，容易造成精力的分散和行政的低效率。

政府绩效评价中存在以上问题，给当前政府管理带来了诸多弊端。评价中对经济指标的偏重导致了对政府绩效的片面理解，把政府职能归结为经济职能，这可能会在短时间内比较明显地促进经济发展，但是将导致政府职能扭曲变形，形成经济与社会发展的不平衡。仅靠经济的发展并不能自动解决贫困、道德、教育、文化、治安、社会公平、社会保障等问题，因此需要积极地探索新的绩效管理方法。

## 2.4 国外政府部门绩效评价的经验

绩效评价在西方国家政府管理中的孕育和产生是西方国家社会发展和政府改革推动的结果。在垄断资本主义条件下，从维护资本主义生产关系和正常社会秩序的客观需要出发，西方国家不得不采用对社会公共事务的积极干预主义。这种管理的集权化和官僚化形成了政府垄断，推行“福利国家”、“人民社会主义”和“混合经济国家”的政策导致了政府管理失控、低效率、财政赤字，导致了公众对政府的不信任和民权运动高涨<sup>[45]</sup>。面对这种社会现实，以美国为代表的西方国家掀起了一场政府改革运动。以绩效为本的政府绩效评价措施正是这场改革所普遍采取的一项重要措施。

但是，西方国家的政府绩效评价被更多地作为政府绩效审计的一个部分，很少被当成一个独立的体系进行研究。这是因为政府绩效审计有着更为完备的步骤与功能，如绩效报告与后续跟踪等，能保证政府绩效评价的结果与建议受到更多的重视和更好的执行。因此，要研究国外政府绩效评价的理论与实践，就离不开研究国外政府绩效审计的理论与实践<sup>[46]</sup>。通过对多国政府绩效审计与评价实践的比较，得到以下一些启示：

① 进行政府绩效评价是必要的，也是有益的。从美国政府绩效评价带来的收益来看，政府绩效评价的成本效益比是令人满意的。虽然我国与美国在政治体制、民主法制、经济生活习惯等方面存在着巨大的差异，但是，大家有理由期待对同样采用现代管理理论的我国政府进行绩效评价也会取得不俗的成效。

② 要有立法的保证。将政府绩效评价以法律的形式确定下来，不仅可以引起人们对政府绩效评价的重视，而且可以使评价主体在进行绩效评价时有据可依，也有利于扫除评价主体在绩效评价过程中的一些不必要的障碍。我国是法治社会，要开展政府绩效评价首先应该对其进行立法。

③ 评价主体应该是国家审计机关。发达国家的评价主体中都是国家审计机关，加拿大的民间审计组织也可以参与评价政府绩效。在政府绩效评价的过程中，国家审计机关和民间审计组织各有其优劣之处。一般来说，国家审计机关由于具有较大的人才优势和法律赋予的权威性而更适合于评价情况复杂的政府绩效，而民间审计组织则由于其独立性而备受人们的青睐。在我国，由于民间审计组织力量的过于单薄而使其难以胜任，因此，我国的政府绩效评价工作就应该由我国的审计机关进行。但是，我国审计机关并不象西方议会制国家那样隶属于国家的最高权力机关，而是隶属于政府，这就难以保证其在进行绩效评价时的独立性。要解决这个问题，改变其隶属关系使其直接对人大负责势在必行。

④ 评价内容应该偏重于结果，同时注意保持持久的绩效。多国的评价内容总体上都包含了“3E”的基本内容，即经济性、效率性和效果性，这三个方面较为全面地概括出了政府当前绩效的全部内容。但是，正如经济发展提倡可持续发展观那样，政府也应该保证持续的绩效。加拿大审计中的“环境审计”和“公平性审计”就在一定程度上体现了持续绩效的一些因素。本研究则认为，平衡计分卡可以更有效地解决这一问题。

⑤ 应该加大政府绩效评价在政府审计机关工作中的比重。美国的这一比重达到了 90%。值得注意的是，美国政府绩效评价得到特别的重视和普遍地开展的时期也是其近年国内经济最快的时期——克林顿时期。

## 第3章 基于平衡计分卡的政府部门绩效评价指标体系

从发展趋势看，无论是西方发达国家还是中国，政府绩效评价正逐步由侧重成本和效率，追求投入产出比的最大化的政府行为转向侧重质量和“客户满意”的企业化运作，尤其注意公共服务的质量和效益<sup>[47]</sup>。所以，选择有效的政府绩效评价方法显得尤为重要。本章引入平衡计分卡，对政府部门绩效考核进行评价。

### 3.1 平衡计分卡基本原理

#### 3.1.1 平衡计分卡的发展

平衡记分卡至今已经经历了十几年的发展历程，大致可以把它分为萌芽、理论研究与应用三个时期。从这三个时期可以大致看出平衡记分卡是如何产生、发展并成为全球企业界广泛接受的一个重要的管理工具的。

##### (1) 萌芽期（1987—1989）

在卡普兰和诺顿研究平衡记分卡之前，Analog Device（简称：ADI）公司最早于1987年就进行了平衡记分卡的实践尝试。这是一家半导体公司，之前每年修改的战略方案被束之高阁，并不能在公司的生产经营工作中得以执行。这次的战略方案制定，公司更加注意战略的实施<sup>[48]</sup>。此次公司将全部的战略文档资料精简到几页纸。他们希望通过面对面与公司员工的交流与沟通，使他们充分理解并认同公司战略。同时，公司继续将战略目标的实现的关键要素转化为年度经营绩效计划，由此衍生出了世界上第一张平衡记分卡的雏形。

此次方案制定过程中，公司为了推行作业成本法邀请了一部分管理学者参与，哈佛大学的教授罗伯特·卡普兰（Robert S. Kaplan）是其中的一位。在帮助公司推行战略方案的过程中，他发现了平衡记分卡，并认识到它的重要价值。

##### (2) 理论研究期（1990—1993）

在卡普兰发现第一张平衡记分卡后，他与复兴全球战略集团（Nolan-Norton）总裁大卫·诺顿（David P. Norton）开始了平衡记分卡的理论研究。

1990年，毕马威（KPMG）的研究机构诺兰诺顿研究所赞助的研究计划，叫做“未来的组织绩效衡量方法”。该研究基于过分依赖财务绩效衡量，会妨碍企业创造未来经济价值的能力的理念。这项计划由卡普兰和诺顿主持，12家来

自制造、服务、重工业和高科技产业的企业参与了这项计划<sup>[3]</sup>。经过研究小组反复讨论，新的衡量系统围绕着四个层面：财务、客户、内部流程、学习与成长而建立，称为平衡计分卡。

1992年初，卡普兰和诺顿将平衡记分卡的研究结果发表在《哈佛商业评论》上，名称为《平衡记分卡——驱动绩效指标》<sup>[4]</sup>。文章详细阐述了1990年参加最初研究项目采用平衡记分卡进行公司绩效考核所获得的益处。该论文发表后，卡普兰和诺顿很快就受到了几家公司的邀请，平衡记分卡开始得到企业界的关注。以后的十多年中，两人又发表了多篇论文和专著（见第1章中的综述），为平衡积分卡的完善做出巨大贡献。

### （3）应用期（1994—今）

卡普兰和诺顿将平衡记分卡延伸到企业的战略管理系统之后，开始广泛得到全球企业界的接受与认同，同时平衡记分卡还延伸到非盈利性的组织机构中。2003年Glanced Scorecard Collaborative Pty Ltd的调查统计显示：在全世界范围内有73%的受访企业正在或计划在不久的将来实施平衡记分卡；21%的企业对平衡记分卡保持观望态度；只有6%的企业不打算实施平衡记分卡。其在非盈利机构的运用在第一章中已经有详细的介绍。在《财富》杂志公布的世界前100位公司中有的公司采用了平衡计分卡系统，以上的北美企业已采用它作为企业内绩效评价的方法。并且平衡计分卡所揭示的非财务的评价方法在这些公司中被广泛运用于员工奖金计划的设计与实施中。

## 3.1.2 平衡计分卡的内容

平衡记分卡是一项与公司策略相结合，协助企业将组织策略加以落实的管理制度，是一个综合评价企业长期战略目标，把企业及其内部各部门的任务和决策转化为多样的、相互联系的目标，然后再把目标分解成多项指标的多元业绩评价系统<sup>[48]</sup>。它由下面四个方面组成：一是客户角度，客户如何看我们；二是内部流程角度，如何改善流程；三是创新和学习角度，如何提高应变能力和创新能力；四是财务角度，我们怎样满足股东。以下为平衡计分卡基本思路示意图（如图3-1所示）。

### （1）客户方面

企业要把向客户提供一流的产品和优质的服务作为其主要价值之一。现代企业的竞争力在于产品和服务能满足客户需求，并从客户的角度开发新产品，避免产品需求少、没有市场的局面。此外，根据客户情况可以对市场进行划分，

针对不同客户制定不同的销售和定价战略。这样，满足客户的需求成为企业的战略目标。平衡计分卡在客户方面的内容应该真实地反映企业在市场中为客户提供了什么样的价值，可用市场占有率、客户获得率、客户满意程度、退货率等指标表示<sup>[49]</sup>。企业要把客户服务的概念贯彻到企业的每一个环节当中，通过不断细分，把客户服务的重要性明确到每一个部门和流程，转化为企业内部的指标。使产品生产、产品销售以及财务、人事等职能部门都围绕客户这一中心，确保为客户提供更优质的服务。

#### （2）内部流程方面

平衡记分卡的内部测量指标，应当来自对客户满意度有最大影响的业务程序，包括影响循环期、产品质量、雇员技能和生产率的各种因素。公司还应努力确定和测量自己的核心能力，即为保证持久的市场领先地位所需的关键技术。由于经济环境和企业战略目标的改变，需要对传统的业务流程进行改造。新的业务流程必须达到以下目的：企业能够迅速预知客户的需求，更好地为客户创造价值，从而最终实现企业的战略目标。

#### （3）学习与成长方面

只有持续不断地开发新产品，为客户提供更多价值并提高经营效率，公司才能打入新市场，增加收入和毛利，才能发展壮大，从而增加股东价值。创新和学习反映了公司和人员素质的提高，这是企业长期稳定发展的基础，也是企业在长期竞争中的核心竞争优势。只有通过创新与学习，才能将客户、内部流程和财务指标整合起来，才能增加公司的价值，使公司具备前进的动力。

#### （4）财务方面

财务指标历来是绩效评价系统中必不可少的一部分，是绩效评价中被最广泛应用的一项指标。在平衡计分卡中，财务指标也是很重要的一部分，因为客户、学习、内部流程等方面的业绩会通过财务表现出来，企业的经营成果也是通过财务指标显现出来。平衡计分卡的财务指标以战略为基础，与传统的财务指标有本质上的区别。

平衡计分卡的四个方面并不是相互独立的，而是可以通过一条因果链，展示业绩和业绩动因之间的关系。首先，企业的战略目标反映在这四大指标中。其次，四大指标存在着相互联系，即企业经营效果的好坏通过相关财务指标来表现；为提高经营成果，必须使产品或服务赢得客户的信赖；要使客户信赖，必须提供客户满意的产品，为此就需要改进内部经营管理过程；而要改进内部经营管理过程，必须对职工进行培训，提升企业的创新和学习能力。

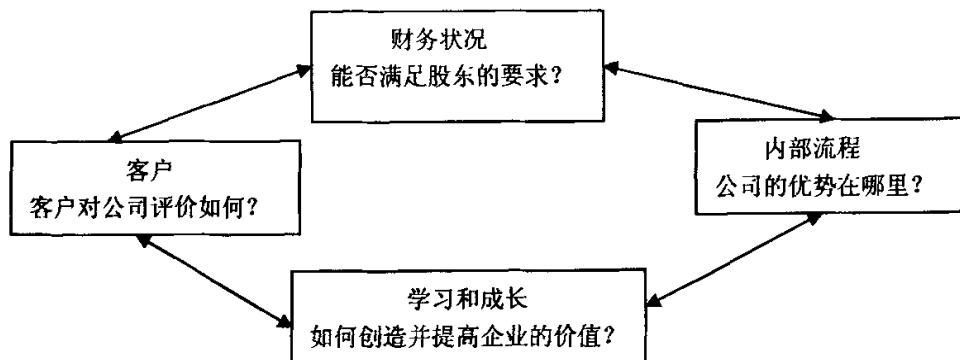


图 3-1 平衡计分卡基本思路示意图

[资料来源：Kaplan R.S, David P Norton. The balanced scorecard-measures that drive performance. Harvard Business Review, 1992(Jan/Feb): 771]

### 3.1.3 平衡计分卡的优点和不足

#### (1) 平衡计分卡的优点

平衡计分卡打破了传统绩效评价方法财务指标一统天下的局面，实行财务指标与非财务指标的有机结合，是一个具有多维度的绩效评价体系<sup>[50]</sup>。

- ① 平衡计分卡实现了内部考核与外部考核之间的平衡。
- ② 平衡计分卡使整个组织行动一致，服务于战略目标，克服财务评价方法的短期行为，实现了长期目标与短期目标之间的平衡。
- ③ 平衡计分卡实现了成果与成果动因之间的平衡，各指标间的因果关系有利地保证了企业目标的实现。
- ④ 平衡计分卡能有效地将组织的战略转化为组织各层的绩效指标和行动。
- ⑤ 平衡计分卡有助于各级员工对组织目标和战略的沟通和理解，有利于组织和员工的学习成长和核心能力的培养。
- ⑥ 平衡计分卡具有长远的战略意义，可以实现组织长远发展。

#### (2) 平衡计分卡的不足

- ① 指标难以量化。平衡计分卡中的评价指标在创建时要考虑与原有的企业日常绩效指标系统的关系，它们之间有无重叠和矛盾的地方，该如何取舍，如何分配权重。有些指标是很难量化的，如员工满意程度、员工受激励程度等方面指标。难以量化便降低了考核的准确性和实用性。
- ② 实施成本高。平衡计分卡要求企业从财务、客户、内部流程、学习和成长四个方面考虑战略目标的实施，并为各个方面制定详细而明确的目标和指标。

这是一项复杂而艰巨的工作，需要为此投入大量的人力、财力和物力。较高的实施成本使一些企业，尤其是中小企业难以接受，限制了平衡计分卡的推广。

③ 不利于全员参与管理目标的实现。平衡计分卡是建立在高层管理者已经明确制定出的战略基础之上的，之后的工作只是将战略逐层分解、下传、落实。这种自上而下将战略转化为行动的过程使员工只能被动接受，不能参与到企业战略决策的制定中，不利于调动员工的积极性和激发他们的创造性。

### 3.1.4 平衡计分卡运用于政府部门绩效评价的意义

卡普兰和诺顿（1996）认为，平衡计分卡除了适用一般营利机构外，也非常适合非营利事业机关，尤其是政府机关<sup>[6]</sup>。平衡计分卡注重各方面的平衡，将战略很清楚的显现出来，并转化为实际行动，此可协助政府机构了解在预算的收支之外，透过绩效动因的设计与引导，真正的将目标达成。“平衡计分卡的观念和方法逐渐在全世界的政府部门和非盈利组织中得到了比较广泛的采纳与使用<sup>[6]</sup>。”具体来说：

① 平衡计分卡引入政府改革，体现了公民为本的理念。它专门将客户（公众）划分为政府绩效管理的一个层面，强调政府应从质量、服务和成本几方面关注客户的需求和满意度，设立了相应的关键评价指标，同时也给客户提供了监督政府的方法和可能。

② 平衡计分卡引入政府改革之中，体现了市场理念。市场理念的内容被有关学者总结为：一方面市场化是改善政府的要径和社会治理的基本战略的信念；另一方面市场理念不仅要求政府部门接受市场管理的“技术”，而且政府改革接受市场“价值——竞争的价值、社会选择中对市场机制的偏好、对企业家精神的尊崇”。

③ 平衡计分卡引入政府改革之中体现了结果导向。结果导向(result-oriented)是针对传统行政管理实践中的“规则导向”和“投入导向”而提出的，结果导向关注使命和组织目标的实现，政府管理应着眼于终极产品和实际的社会效果<sup>[51]</sup>。平衡计分卡的产生就是对单一的财务评价难以完整地体现组织使命和目标的否定。它纳入到政府的战略管理体制之后，把政府抽象的使命和战略转变为明晰的目标，并用具体的评价手段和指标加以衡量；对战略目标和衡量方法进行宣传，把激励机制同业绩评价手段及指标联系起来；制定计划，确立明确的目标；强化战略性的反馈和学习。

## 3.2 平衡计分卡运用于政府部门绩效评价的可行性分析

### 3.2.1 平衡计分卡运用于政府部门绩效评价的适用性

Attkission 等人曾经指出：“为了使绩效评价发挥作用，必须对组织的目标达到共识。一个组织必须把其目标具体化，以使它们所取得的成效的质量和程度能够得以测量<sup>[51]</sup>。”所以，管理者要为组织的每一个员工设定岗位目标，把组织目标分解到每个部门，进而分解到个人。

政府部门的绩效管理也是通过管理目标来实现的。政府绩效评价的一个主要功能就是详细地表明组织的目标或使命，以至于这些目标和使命可以被评价。而从平衡计分卡的概念来看，平衡计分卡表明了企业员工需要什么样的知识、技能和系统（学习与成长角度）才能创新和建立适当的战略优势和效率（内部流程角度），使公司能够把特定的价值带给市场（客户角度），从而最终实现更高的股东价值（财务角度）<sup>[52]</sup>。而当今政府战略目标也是围绕整体规划发展理念进行的。由于政府绩效评价目标的战略导向性和平衡计分卡的战略目标导向性正好契合，所以平衡计分卡在战略意义上可以运用于政府绩效评价。

从客户价值角度看，政府需要满足的人群是公众，为他们提供满意的政府服务。客户价值可以由客户感知价值转变为公众感知价值，客户至上的理念就可转变为公众至上。所以，平衡计分卡应用于政府是可行的。

政府为了向公众提供更好的服务，也必须规范其内部的管理，建立相应的规章制度以保证政府部门的正常运作。政府部门实施内部管理的目的也是为了保证政府部门与外界的信息沟通渠道的畅通，使政府部门能快速、准确地了解公众的意见和反馈，实时对政府行为进行调整，提高政府的工作效率。

为了满足公众不断变化的服务需求，政府部门除了需要强化服务意识外，还需不断开发出新的服务内容，这就要政府工作人员的创新意识和能力，而这种创新意识和能力是需要通过培训获得的。所以，政府部门需不断对其工作人员进行培训，提升其学习和创新能力，并给予他们发挥潜能的机会。

政府部门的经费虽然是财政划拨，但从内部管理的角度看，也需对相关费用进行合理预算，对各种资源的利用也需用效率指标来衡量，即要求用尽可能少的投入产生尽可能多的产出。

综上所述，平衡计分卡用于政府绩效评价无论从战略、客户价值上考虑，还是从内部管理、学习与成长、财务的角度考虑都是适用的。

### 3.2.2 政府部门绩效评价的特殊性

将考核企业绩效的平衡计分卡运用于政府绩效评价时，必须要考虑到企业与政府在绩效评价本质特征上的差别，并针对这些差别对平衡计分卡进行相应的调整<sup>[19]</sup>。

从企业的先天条件来看，绩效评价系统在处于市场竞争的环境中产生。绩效评价在实施过程中，企业具有政府部门无法比拟的先天优势。首先，有着强烈的市场导向性；其次，企业绩效评价中有明确的组织目标，而且这个目标是管理者与员工双方认同的；再次，绩效管理更多的是强调组织的整合过程，整合一个组织、部门、群体和组织内个人的目标，并与人力资源的开发密切联系；最后，由于产出的商品性，企业的绩效指标很容易通过市场机制反映出来，从而便于进行成本收益分析<sup>[53]</sup>。

由于政府具有自身的特殊属性——公共性，与生产企业相比，政府绩效评价具有特殊性：

① “产品”的非商品性<sup>[54]</sup>。由于政府特有的公共性，公共产品只能由政府来生产，政府产出不具有商品性，没有市场竞争，所以政府绩效也没有相应的市场产品标准进行衡量检验。

② 目标任务的政治性。政府机构的目标往往取决于外部因素，比如执政党的政治纲领、利益集团的要求、领导人的变换等等。行政机构领导人自主性较弱。一个具体的政府部门除了上级部门的命令外，就是工作程序所要求的任务，并没有更多地考虑到其产出的效益和对社会环境的影响。

③ 目标的多样性和抽象性。目标的多样性表现为不同的政府部门追求政治、经济、军事、社会、文化等不同性质的目标，甚至同一部门追求多重目标且经常发生变化。目标的抽象性即目标表述笼统、抽象且难以量化为硬性指标，如提高人们的素质和道德水平，保障社会公正等，这都为政府绩效评价的指标设立带来了一定的技术难度。

④ 投入产出难以量化。政府部门的产品是非商品性的，经济学家称其为“非市场产出”。所以，产出不能像企业产品那样最终进入市场的交易体系，形成一个货币价格。而且，政府部门产出涉及范围的综合性、涵盖地区的差异性、标准指标的主观片面性等特点，政府部门的绩效指标难以确定。由于政府部门的生产具有较强的劳动密集型特征，劳动投入是关键投入，行政效率的高低主要取决于劳动的边际生产率，但劳动的投入往往难以量化<sup>[54]</sup>。

⑤ 强调公平胜于效益。政府所追求的是每个公民的平等，主要是收益的平等。而企业绩效管理意味着要找到有效的激励手段，把个人的投入与收益联系起来。企业员工创造的成果越多，收益也就越大，奖赏就越高，这就是企业绩效管理的“公平”。但这与政府部门理念是相违背的。

⑥ 封闭式的财务核算体系<sup>[34]</sup>。政府部门制定财务预算或进行年末核算时，根本就没有经过市场检验。国家税款靠强制性预先征收获得，行政收入无需经受市场检验，无需政府自行垫付，行政成本与收入之间的逻辑关系与企业生产正好相反。因此，政府生产缺乏亏损底线那样的硬约束。在盈利的情况下，政府部门不仅不能把利润占为已有，而且还要担心盈利多的情况下，下一年度的预算会因此减少。这就鼓励政府以追求规模、消费扩张作为最大化目标，而非以成本最小化为目标。

### 3.2.3 平衡计分卡运用于政府部门绩效评价的修正

考虑到政府的上述特殊性，必须对平衡计分卡进行适当的调整。从本质来说，政府目标具有多元性，而且是抽象性的、笼统的和难以量化的。可见，在组织目标是评价组织绩效的主要依据的前提下，这些就在无形中加大了政府绩效评价的难度。但是，无论是哪一级政府在一定时期内都会有一个总目标，政府各部门在这个总目标的指导下进行各种工作。可见，目标导向的组织安排有助于政府绩效管理的实施（如图 3-2 所示）。所以，可以把政府战略目标作为政府绩效评价的切入点，这也符合平衡记分卡的基本思想<sup>[18]</sup>。

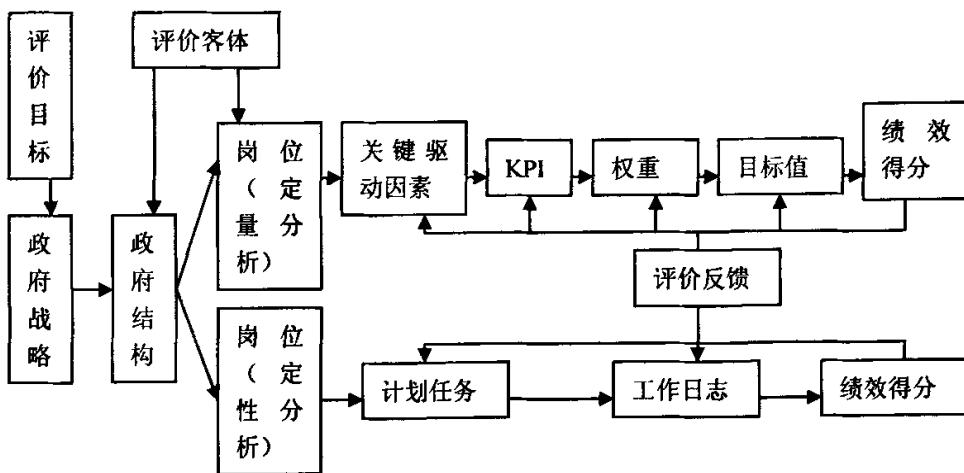


图 3-2 基于平衡记分卡的政府绩效评价示意图

然后再来分析一下平衡计分卡的四个维度。

在企业中，财务方面成为企业业绩考核的核心内容，而由于政府财务核算的封闭性，政府原则中强调的公平性，政府绩效评价中财务方面只作为辅助平衡计分卡中其他三方面的协助部分。但又考虑到政府考核的经济原则，财务维度必须作为绩效评价的重要一块，否则就会造成政府垄断条件下的浪费。

由于政府的非营利性、特有的公共性，政府绩效的落脚点应该在客户上面。政府的客户是公众。怎样为公众服务成为了考核的标准，公众满意原则也成为了政府绩效评价的基本原则。然而，公众客户很难调查。并且，政府的工作绩效由于其特殊的垄断性，无法向公众公开，公众在不透明的情况下，填写的有关客户价值的调查问卷很难科学地反映客观事实。因为客户价值等于感知到的收益减去感知到的成本。感知到的成本公众是无法衡量的。加之政府产出难以量化，政府不能也没有能力公开提供公众服务所消耗的成本，所以公众无法衡量感知到的成本。鉴于这种情况，只能暂且用提供公共服务来衡量客户价值。

在学习和成长方面，企业强调的是人员的创新能力和积极性、信息系统的完善、企业程序的完备。而转化到政府部门时，对应的应该是人力资源状况、电子政务、创新、改革能力等，即政府的自身建设。

在内部管理方面，企业表现出来的是为了满足客户需要，同时满足预算约束。由于政府的业务有着一定的公共性，内部业务就是对公众能发挥什么作用，内部业务外部体现的职能是什么，这就是政府应该满足公众的最基本外部职能，如经济调节、市场监督等。

由此可以得出以下结论（如图 3—3 所示）：

第一，政府的战略目标是政府绩效评价的起点，全部的考核内容都要根据这个目标来确定。

第二，由于政府的非营利性，财务方面成为了其他考核维度的辅助部分，其具体内容表现为公共事业管理成本、政府自身建设产生的内部管理成本、国有资产管理和政府职能成本<sup>[55]</sup>。政府的学习和成长表现为政府部门的自身建设，内部管理表现为市场监督和经济调节职能，客户价值表现为政府公共事业服务。

第三，公众服务是政府部门绩效评价的核心，政府的自身建设是为市场监督和经济调节服务的，也是政府部门为公众提供公共事业服务的，在保证为公众提供良好服务、发挥市场监督和管理作用的基础上，尽可能地有效利用各种资源。

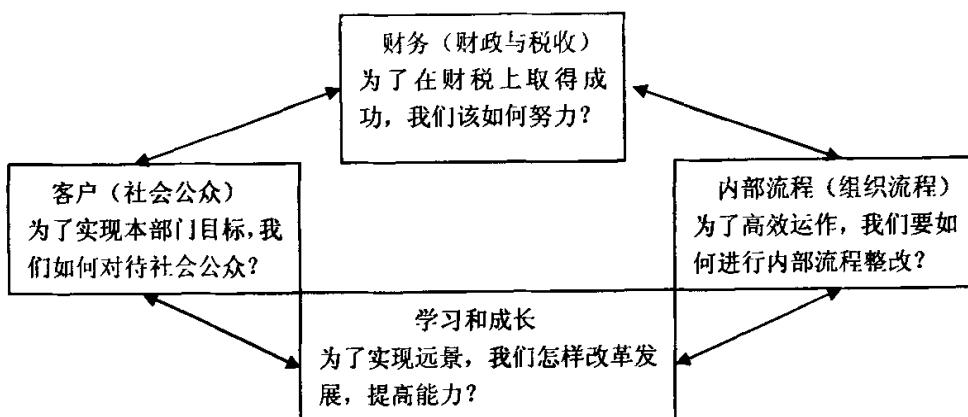


图 3-3 中国模式的政府部门平衡计分卡示意图

### 3.3 基于平衡计分卡的政府部门绩效评价指标体系

#### 3.3.1 指标影响因素的分析

运用平衡计分卡构建政府部门绩效评价指标体系是一个受环境和诸多外部条件影响和制约的系统。在政府部门绩效评价过程中，绩效评价目的、政府部门职能、政府部门发展目标、政府部门主客观条件、发展趋势及可获取信息数量和质量等六大因素对政府部门绩效评价指标体系及其构建有重大的影响。

##### （1）绩效评价目的

致力于绩效改进的政府部门绩效评价在指标体系构建上比单纯的绩效评价更应注意影响绩效的潜在因素和改进方面的分析<sup>[42]</sup>。事实上，能够用于评价地方政府或者地方政府某一个部门、某一个岗位的绩效的指标是多种多样、数量可观的，但进行地方政府公共事业管理绩效评价时，不可能也没有必要面面俱到，只能选择那些以绩效评价目的为依据的评价指标进行评价，否则，绩效评价不仅没有可操作性，而且也会失去意义。

##### （2）政府部门职能

从宏观上分析，一级地方政府具有教育、科技、文化、卫生、体育、基础设施建设、环境保护等公共事业管理职能，履行这些职能会产生政府相应的教育事业管理绩效、文化事业管理绩效、体育事业管理绩效、基础设施建设管理绩效、环境保护管理绩效等<sup>[42]</sup>。各级各类政府部门的管理绩效具有相关性和整体性。如果不从系统的高度全面地制定相应的绩效评价指标体系，顾此失彼、厚此薄彼，就会误导政府行为。同时，不同的地区由于受自然条件和历史因素

的影响，导致各地政府部门在管理公共事务上存在着显著的差异。因此，在构建指标体系时，应充分考虑不同职能对指标设置的决定性作用和影响，既要设置一些通用指标，也要考虑一些特殊指标，或是运用数量分析技术，尽量缩小地区间职能差异给绩效评价指标体系构建与绩效评价带来的误差。

### （3）政府部门发展目标

发展目标是政府部门发展战略构成的一个基本内容，它表明政府部门履行其使命所要达到的结果。发展目标表明了政府部门的使命、价值观，体现了公众对它的具体期望和要求。只有明确了发展目标才能使绩效评价，进而绩效管理有了源头。

### （4）政府部门主客观条件

政府部门的主、客观条件包括三个方面的内容：一是上级要求，即上级制定的总目标，下达的任务、指令、计划及有关政策、法规、条例等；二是本身的实际情況，包括经济、社会发展的水平和质量及上期目标的实现程度、人员素质、管理水平等情况；三是外部环境，即相关区域、部门、单位的发展速度、发展质量等因素<sup>[47]</sup>。只有全面系统地分析探究和掌握运用这些主、客观条件，才能确保政府部门绩效评价指标体系的科学性和可行性。

### （5）发展的趋势

绩效评价指标是标志和指向未来的，具有导向性。构建政府部门绩效评价指标体系时，不能只依据事物存在的现状，还必须考虑未来的因素，以事物发展趋势为依据。当某些指标已经不适应变化的情况或者根本没有战略意义时，应该被及时删除。

### （6）政府部门信息收集

从信息管理的角度来讲，政府部门绩效评价的过程就是一个信息的搜集、处理过程。信息是否充足、准确、真实、及时，直接影响到绩效评价指标能否具有可测量性及在多大程度上生效，影响到绩效评价能否进行及其成败。

## 3.3.2 指标选择的方法

### （1）德尔菲法的概念

德尔菲依据系统的程序，采用匿名发表意见的方式，即专家之间不得互相讨论，只能与调查人员发生关系，通过多轮次调查专家对问卷所提问题的看法，经过反复征询、归纳、修改，最后汇总成专家基本一致的看法，作为预测的结果。这个方法主要用于难以量化或者数据收集困难的情况。

德尔菲法的具体实施步骤如下：首先，组成专家小组。专家人数的多少，可根据预测课题的大小和涉及面的宽窄而定，一般不超过 20 人。其次，向所有专家提出所要预测的问题及有关要求，并附上有关这个问题的所有背景材料，同时请专家提出还需要什么材料。然后，由专家做书面答复。再次，各个专家根据他们所收到的材料，提出自己的预测意见，并说明自己是怎样利用这些材料并提出预测值的。最后，将各位专家第一次判断意见汇总，列成图表，进行对比，再分发给各位专家，让专家比较自己同他人的不同意见，修改自己的意见和判断。将所有专家的修改意见汇总，再次分发给各位专家，以便做第二次修改。逐轮收集意见并为专家反馈信息是德尔菲法的主要环节。收集意见和信息反馈一般要经过三、四轮，在向专家进行反馈的时候，只给出各种意见，但并不说明发表各种意见的专家的具体姓名。这一过程重复进行，直到每一个专家不再改变自己的意见为止。

### (2) 德尔菲法的应用

在绩效评价过程中，用德尔菲法统计的内容一般有三个：

一是评价指标体系的构建中，指标的选择和从属关系。

二是每一个测评指标权重的平均值，计算公式为：

$$M_i = \frac{1}{n_i} \sum_{j=1}^m W_{ij} \quad (i=1, 2, \dots, m) \quad (3-1)$$

其中， $M_i$  是专家值的加权平均， $W_{ij}$  是第  $j$  个专家对第  $i$  指标权重的估计值。

三是每一个专家的估计值与平均估计值的偏差分析，其分析公式为：

$$\Delta_j = w_{ij} - M_i \quad (i=1, 2, \dots, m) \quad (3-2)$$

其中， $\Delta_j$  为每位专家值与平均值的标准差。当进行第二轮咨询时，要将第一轮的统计分析情况反馈给专家，并提示  $\Delta_j$  较大的专家尽量做出新的判断。通过  $n$  轮比较反复以后，一般可以得到较为一致的权重系数。

### 3.3.3 一般性指标体系的建立

由于在自然条件、经济发展水平、社会结构等方面存在着差异性，不同地区的政府部门绩效管理的重点不同，难以建立一套适合于我国所有政府部门的指标体系。但是，可以选取一些具有普遍适用性的绩效指标，以便为政府的实

践提供参考。根据上文对政府平衡计分卡的分析，结合前人的研究成果，本文尝试对政府平衡计分卡的评价指标进行一般性分析，建立四维指标体系（如表 3-1 所示）：

表 3-1 基于平衡计分卡的政府部门绩效评价指标体系

一级	二级	三级	四级
客户指标	经济	基础经济	GDP 增长率、人均 GDP、劳动生产率
		产业经济	三产占 GDP 比重、外商投资额、国有资本保值增值率
	社会稳定	社会安全	刑事案件发案率、生产与交通事故死亡率
		社会保障	社会保险总资金、下岗人员安置率、城镇登记失业率
		社会公平	女性劳动者就业率、残疾人就业率
		社会财富	人均可支配收入、恩格尔系数、基尼系数
	公共事业	科技	科研经费投入、科技进步贡献率、专利批准量
		教育	教育投入、教育质量、师资队伍学历、人均受教育程度
		文体	文艺事业指标、体育事业指标
		公共卫生	医疗投入、投入产出、公众健康
财务指标	公众监督	城市建设	城市规划、基础设施建设、社区建设、公共产品提供
		立法执法情况	法律法规完善程度、执法程度、企业满意度
		政府声誉	公众对政府决策的支持率、信息公开程度、公众满意度
	人口与环境	资源	单位 GDP 能耗降低率、可持续开发程度
		环境	森林覆盖率、污染指数、三废利用率、市民居住舒适感
		人口	自然人口增长率、人均预期寿命、非农业人口比重
	支出	会计成本	公务员工资总额、例行的公共财政支出、公共项目投资
		社会成本	社会、经济、生态损失
	收入	税	年税收总额、税收管理效率
		利	国企资产保值增值率、国企利润增长率
		债	国际债务结算、年发行国债总额
		费	各种费用征收总额

表 3—1 基于平衡计分卡的政府部门绩效评价指标体系（续表）

内 部 流 程 指 标	行 政 效 率	行政执行力	行政经费占财政支出比重、信息管理水平
		行政质量	投诉率
		行政决策	政府投资项目回报率
学 习 与 成 员 长 指 标	行 政 廉 洁	行政人员违纪	违纪行政人员占行政人员比重、违纪处理结果
		腐败案件	腐败案件涉案金额、腐败案件人数占行政人数比重
		作风评议	机关工作作风、公众评议结果
	行 政 资 源	人力资源	公务员人数、公务员素质、行政人员占人口比重
		物资资源	车辆配置、现代通讯设施、办公用品
	公 务	职业培训	学历再教育、专业培训、增强管理行技能、提高咨询能力、
		工作满意度	合理意见采纳
		信息系统的建设	信息的准确及时性、反馈速度、系统的稳定性和安全性
	政 府	改革能力	机构设置和行政适应度、政府行为公平性和法制化程度
		部门文化建设	重塑政府文化

上表已经把政府绩效考核的各方面指标全部囊括进去，在进行具体部门的分析时，应该从中挑选 20—25 个最有代表性和可获得性的指标，不可太概括含糊，否则将无法取数。

## 第4章 政府部门绩效的模糊综合评价

关于人力资源绩效评价的研究，国内外已经研究出比较系统的定性与定量的评价方法。定性的方法包括比较法、分类法、特性法、结果法、质量法等。但是，人力资源绩效具有复杂性、多因素性、模糊性等特点，这就决定了定性分析难以对被评价者作出科学、客观的评价。定量评价的方法主要有目标管理法（MBO）、360度评价法、平衡计分卡（平衡计分卡）和关键指标法（KPI）。这几种方法虽然能够在一定程度上弥补定性分析带来的不足，但不足以体现评价的公正性和客观性。模糊教学的问世，为人力资源绩效评价定性与定量相结合的研究提供了一个有力的工具<sup>[56]</sup>。

### 4.1 政府部门绩效模糊综合评价方法

#### 4.1.1 层次分析法

层次分析法是20世纪70年代由美国运筹学教授T. L. Saaty提出的（原名为The Analytical Hierarchy Process），简称AHP法。层次分析法是对系统工程中非定量事件一种测评分析方法，用来确定权向量，并对确定目标进行排序，从中选出最优者。该法是从定性分析到定量分析综合集成的一种典型的系统工程方法，它将人们对复杂系统的思维过程数学化，将人的主观判断为主的定性分析进行定量化，将各种判断要素之间的差异数值化，帮助人们保持思维过程的一致性，适用于复杂的模糊综合测评系统，是目前一种被广泛应用的确定权重的方法<sup>[22]</sup>。运用层次分析法建模，大体上可按下面4个步骤进行：建立递阶层次结构模型；构造判断矩阵；计算权向量；一致性检验。

#### 4.1.2 模糊综合评价法

美国著名控制论专家扎德（Zadeh）于1965年创立了描述模糊现象的方法——模糊集合论。这一方法把考察的模糊对象及反映其模糊观念作为一定的模糊集合，建立适当的隶属函数，并通过模糊集合论的有关运算和变换，对模糊对象进行定量分析<sup>[57]</sup>。

模糊综合评价法就是以模糊集合论为基础，应用模糊关系合成的原理，将一些边界不清、不易定量的因素定量化，通过构造等级模糊子集把反映被评事

物的模糊指标进行量化（确定隶属度），然后利用模糊变换对各指标进行综合评价的一种方法<sup>[58]</sup>。模糊综合评价法把定性描述和定量分析紧密地结合起来，应用面广。它最大的优点是不但能处理现象的模糊性，综合各个因素对总体的影响作用，而且能用数字反映人的经验。凡是涉及到多因素的综合判断问题，都可以用模糊综合评价方法来解决。

## 4.2 政府部门绩效模糊综合评价模型

德尔菲法、层次分析法和模糊综合评价三者各有其优缺点。将三种方法综合起来，扬长避短，建立一种新的综合测评，应用于政府部门绩效考核评价体系中，势必将进一步提高评价的信度、效度和精度，目前这一设想在这个领域中的应用很少。

评价指标体系的合理与否，直接影响评价结果的信度和效度。层次分析法提供了按支配关系分组而形成有序递阶层次结构的指标体系设计方法，但用该方法设计的指标体系是否合理，在此基础上进行进一步的优化设计的方法需要进一步讨论。因个人偏好、主观臆断、以点带面等人为因素给评价结果带来的偏差，这里用灰色理论或者模糊评价都太过客观。所以，本文先运用德尔菲法对指标体系进行筛选，可以有效地减少主观成份的影响。将层次分析法引入模糊综合评价的权重集的建立中，加大科学含量，从而结合模糊综合评判理论与层次分析法建立模糊综合评价方案。具体的步骤如下：

① 要素分解：即确定需要测评的特征要素，建立指标初选集。用德尔菲法和层次分析法思想设计并发放调查问卷，得到最终的政府部门绩效评价指标体系。

② 权重确定：部门不同，各特征要素不同，其重要性也不同。在确定平衡计分卡四个维度及其各指标权重的过程中，本文按德尔菲法和层次分析法思想设计并发放调查问卷，用层次分析法根据影响的主次关系确定它们的加权系数建立权重集。

③ 评定计分：由测评人员对测评对象的各要素进行评定计分，建立评价集。在这里，可以运用德尔菲法和历史经验。

④ 分值计量：对测评人员的评定结果进行等级分配率统计，并运用模糊矩阵运算，计算出其分值。

⑤ 等级划分：把定性和定量指标值进行无量纲化和极性变换，使其可以加权平均，得到综合评价结果。如果需要，还可以给分数划分等级。

#### 4.2.1 指标权重的确定

##### (1) 建立层次结构模型

根据问题和要达到的目标，将问题分解为不同的组成因素，并按照因素间的相互关联以及隶属关系将因素按不同层次聚集组合，形成一个多层次的分析结构模型。一般按照目标层、准则层和方案措施层的形式排列，采用框图表示各层次的递阶结构与各因数之间的从属关系（如图 4-1 所示）。对层次划分的数目及定义并没有严格的规定，可以划分为三层或多于三层，各层次的名称也有具体问题而定。通常每一上层因素一般不与九个以上的下层因素相联系，否则，将会给因素间的两两比较带来困难。

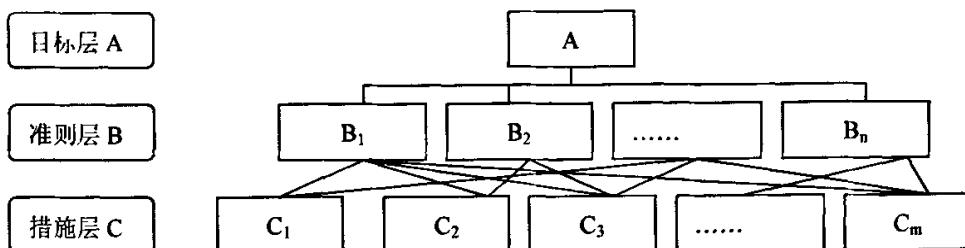


图 4-1 层次结构图（以三层为例）

##### (2) 构造判断矩阵

判断矩阵表示针对上一层次某因素而言，本层次与之有关的各因素之间的相对重要性。根据建立的层次结构模型，以针对 A 层的 B 层两指标的重要性进行比较（如表 4-1 所示）。判断矩阵  $B = [b_{ij}]$  应满足下列条件： $b_{ij} = b_i / b_j$  表示指标  $b_i$  对指标  $b_j$  相对重要性的判断值； $b_{ii} = 1 / b_{jj}$ ， $i \neq j$  ( $i, j = 1, 2, \dots, n$ )； $b_{ii} = 1$ ，即每个元素相对于自身的重要性为 1。

表 4-1 评价指标因子判断矩阵 B

		评价指标				
		$b_1$	$b_2$	.....	$b_n$	
评价指标	$b_1$	1	$b_{12}$	.....	$b_{1n}$	
	$b_2$	$b_{21}$	1	.....	$b_{2n}$	
	.....	.....	.....	.....	.....	
	$b_n$	$b_{n1}$	$b_{n2}$	.....	1	

通过专家评价或由历史（经验）数据分别对每一个评价指标的相对重要性

进行定性描述，并用准确的数字进行量化表示（1—9 分制），数字的取值所代表意义见表 4—2。

表 4—2 数值意义

数值	意义
1	i 指标与 j 指标相同重要
3	i 指标比 j 指标稍微重要
5	i 指标比 j 指标较为重要
7	i 指标比 j 指标非常重要
9	i 指标比 j 指标极端重要
2, 4, 6, 8	表示需要在上述两个标准之间折中时的标度
上列各数的倒数	两个指标的反比较

### (3) 计算权向量

就是计算每一个判断矩阵各因素针对其准则的相对权重。判断矩阵  $B$  对应于最大特征值  $e_{\max}$  的特征向量  $w$ , 经归一化后即为同一层次相应因素对于上一层次某因素相对重要性的排序权值。

① 将判断矩阵  $B$  按列归一。先将每行元素乘积:  $m_i = \prod_{j=1}^n b_{ij}, i = 1, 2, \dots, n$ ; 接着, 求  $m_i$  的  $n$  次方根  $\bar{w}_i = \sqrt[n]{m_i}$ ; 最后, 对  $\bar{w}_i$  进行标准化, 求特征向量各分量

$$w_i = \bar{w}_i / \sum_{j=1}^n \bar{w}_j \quad (4-1)$$

其中,  $i = 1, 2, \dots, n$ ,  $w = [w_1 \ w_2 \ \dots \ w_n]^T$  是  $B$  的特征向量。

②  $B$  的最大特征值为:

$$e_{\max} = \frac{1}{n} \sum_{i=1}^n \left( \frac{(bw)_i}{w_i} \right) \quad (4-2)$$

### (4) 一致性检验

为避免其他因素对判断矩阵的干扰, 在实际中要求判断矩阵满足大体上的一致性, 需进行一致性检验。只有通过检验, 才能说明判断矩阵在逻辑上是合理的, 并继续对结果进行分析。为检验层次单排序的一致性, 需要计算以下指标:

① 一致性指标  $C.I.$  (consistency index) 为:

$$C.I. = \frac{e_{\max} - n}{n - 1} \quad (4-3)$$

式中,  $e_{\max}$  为判断矩阵的最大特征根;  $n$  为成对比较因子的个数。当判断矩阵具有完全一致性,  $C.I.=0$ 。 $C.I.$  越大, 矩阵一致性越差。

② 为了检验判断矩阵是否具有满意的一致性, 需要将  $C.I.$  与平均随机一致性指标  $R.I.$  比较。对于 1—9 阶矩阵,  $R.I.$  分别如下表所示 (如表 4-3 所示)。

表 4-3 平均随机一致性指标  $R.I.$  表

阶	1	2	3	4	5	6	7	8	9	10	11	12
R.I.	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54

③ 判断矩阵的一致性指标  $C.I.$  与同阶平均随机一致性指标  $R.I.$  之比称为判断矩阵的随机一致性比例, 记为  $C.R.$ 。

$$C.R. = \frac{C.I.}{R.I.} \quad (4-4)$$

当  $C.R. < 0.10$  时, 可以认为判断矩阵具有满意的一致性, 否则就需要重新调整判断矩阵。

④ 层次总排序。利用同一层次中所有层次单排序的结果, 就可以计算针对上一层次而言本层次所有因素重要性的权值, 这就是层次总排序。层次总排序需要从上到下逐层顺序进行, 对于最高层下面的第二层, 其层次单排序即为总排序。假定上一层次所有因素  $A_1, A_2, \dots, A_k$  的总排序已完成, 得到的权值分别为  $a_1, a_2, \dots, a_k$ , 它们关于  $A_i$  层次单排序权重分别为  $b_{1i}, b_{2i}, \dots, b_{ni}$ 。现求  $B$  层中各因素关于总目标的权重  $b_1, b_2, \dots, b_n$ , 即

$$b_i = \sum_{j=1}^n b_{ij} a_j, i = 1, 2, \dots, k \quad (4-5)$$

#### 4.2.2 隶属度函数的确定

隶属于度函数 (或隶属函数) 是模糊数学与模糊系统的奠基石, 用来描述差异的中间过渡, 这是精确性对模糊性的一种逼近, 其数学定义为: 设  $A$  是论域  $U$  上的一个模糊子集, 对于任意  $x \in U$ , 都指定一个实数  $\mu(x) \in (0, 1)$  与之对应, 这个  $\mu(x)$  就叫做  $x$  对  $A$  的隶属度 (degree of membership)<sup>[58]</sup>。这实际上构造成了一个映射。具体方法如下所述:

设评价指标集  $U = \{u_1, u_2, \dots, u_m\}$  的定量集合  $X = \{x_1, x_2, \dots, x_m\}$ , 评价等级标准  $V = \{v_1, v_2, \dots, v_n\}$ , 设  $v_j$  和  $v_{j+1}$  为相邻两级标准, 且  $v_{j+1} > v_j$ , 则  $v_j$  级隶属度函数为:

$$r_j = \begin{cases} 1 - r_{j-1} & v_{j-1} \leq x_i \leq v_j \\ \frac{v_{j+1} - x_i}{v_{j+1} - v_j} & v_j \leq x_i \leq v_{j+1} \\ 0 & x_i \geq v_{j+1} \text{ 或 } x_i \leq v_j \end{cases} \quad (4-6)$$

根据上式, 计算第  $i$  个评价指标  $u_i$  隶属于评价等级  $v_j$  的隶属度  $r_{ij}$ , 生成隶属矩阵  $R$ :

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2j} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} & \cdots & r_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mj} & \cdots & r_{mn} \end{bmatrix} \quad (4-7)$$

#### 4.2.3 单级模糊综合评价

- ① 确定评价指标集合  $U = \{u_1, u_2, \dots, u_m\}$ 。其中,  $u_i, i=1, 2, \dots, m$  为评价指标,  $m$  是评价指标的个数。
- ② 确定评语集合  $V = \{v_1, v_2, \dots, v_n\}$ 。其中  $v_j, j=1, 2, \dots, n$  是评价等级,  $n$  是定性的评价等级的个数。这一集合规定了某一个评价指标的评价结果的选择范围。
- ③ 确定隶属度矩阵。假设对每个评价指标  $u_i$ , 在每个评价标准  $v_j$  之上进行单指标评价得到一个隶属度矩阵  $R$  (如公式 4-7 所示)。其中  $r_{ij}$  为指标  $u_i$  具有  $v_j$  的程度,  $0 < r_{ij} < 1$ 。显然, 隶属度矩阵中的每一行是对每一个单指标  $u_i$  的评价结果, 整个矩阵涵盖了按评价等级集合  $V$  对评价指标集合  $U$  进行评价所获得的全部结果。
- ④ 确定评价指标的模糊权重向量  $W = (w_1, w_2, \dots, w_m)$ 。一般情况下,  $m$  个评价指标对事务评价的重要性不可能一致, 因此在合成之前要确定模糊权向量。在模糊综合评价中, 权向量  $A$  中的元素  $a_i$  本质上是指标  $u_i$  在评价指标集中权重, 因而需要做归一化处理:

$$W = \frac{a_1}{u_1} + \frac{a_2}{u_2} + \cdots + \frac{a_m}{u_m} \quad (4-8)$$

其中,  $a_j \geq 0$ ,  $\sum a_j = 1$ 。

⑤ 模糊合成。用模糊向量  $W$  将不同的行进行综合就可以得到该被评价事物从总体上来看对各等级模糊子集的隶属程度, 即模糊综合评价结果向量  $B$ 。模糊综合评价的模型为:

$$B = [b_1 \ b_2 \ \cdots \ b_j \ \cdots \ b_n] = [w_1 \ w_2 \ \cdots \ w_i \ \cdots \ w_m] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ r_{i1} & r_{i2} & \cdots & r_{in} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \quad (4-9)$$

其中  $b_j$  是由  $W$  与  $R$  的第  $j$  列运算得到的。

⑥ 评价得分。为了定量地表示评价的结果, 可以对于等级因素赋值。设赋值之后得到矩阵  $V'$ , 即  $V' = [p_1 \ p_2 \ \cdots \ p_n]'$ 。最后可以得到综合评价的结果分数  $B'$ , 即:

$$B' = [b_1 \ b_2 \ \cdots \ b_n] \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} \quad (4-10)$$

#### 4.2.4 多级模糊综合评价

以上的模型给出了评价因素只有一级时的综合评价法, 本研究以此作为基础, 进行多级评价。

① 确定评价指标集。将所有指标分成  $S$  个子集, 记为  $U_1, U_2, \dots, U_s$ , 并满足条件  $U = \{U_1, U_2, \dots, U_s\}, U_i \cap U_j = \emptyset (i \neq j)$ 。

其中  $U_i = \{X_{i1}, X_{i2}, \dots, X_{in}\}, i=1, 2, \dots, s$ ;  $n_i$  表示  $U_i$  的元素个数, 这个工作上面已经完成。做出所有评价指标的统一评语集,  $V = \{v_1, v_2, \dots, v_m\}$ 。

② 对每一个评价指标  $U_i$  进行单指标评价, 得出单指标评价矩

$$R_i = (r_{ijk})_{n_i \times m} \quad (4-11)$$

其中  $i=1,2,\dots,s$ ;  $j=1,2,\dots,n_i$ ;  $k=1,2,\dots,m$ 。这里  $r_{jk}$  表示指标  $X_j$  对评语  $v_k$  的隶属度。进行综合评价时, 可以根据实际意义确定单指标评价矩阵  $R_i$ , 在此采用比较容易的一种方法, 即专家评分法。具体过程是: 每位专家给每项指标打分, 打分范围在区间[0—1]之内。例如, 每位专家给  $X_j$  项指标打分时, 应满足:  $\sum_{k=1}^m r_{jk} = 1$  打分后, 对每项指标的得分取平均值, 得出最终得分, 并以其作为对应的隶属度。

③ 给出  $U_i$  中各评价指标的权重,  $\omega_i = \{a_{i1}, a_{i2}, \dots, a_{im}\}$ , 应有  $\sum_{j=1}^n a_{ij} = 1$ 。

④ 得出  $U_i$  的最终评语为:  $B_i = \omega_i \cdot R_i = (b_{i1}, b_{i2}, \dots, b_{im})$ ,  $i=1,2,\dots,s$ 。将  $U_i$  作为一个单独元素, 用  $B_i$  作为  $U_i$  的单指标评价向量, 可构成  $U$  到  $V$  的模糊评价矩阵:

$$R = \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_s \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ b_{s1} & b_{s2} & \cdots & b_{sm} \end{pmatrix} \quad (4-12)$$

按照  $U_i$  在  $U$  中的重要程度给出的权重,  $\omega = \{a_1, a_2, \dots, a_s\}$ , 于是得出  $U$  的最终评语向量

$$B = \omega \cdot R = (b_1, b_2, \dots, b_m) \quad (4-13)$$

计算最后的评价得分:

$$G = B \cdot V^T \quad (4-14)$$

#### 4.2.5 评价值无量纲化和极性变换

一般来说, 不同的评价指标具有不同的量纲, 为了消除由于指标的量纲不同所带来的不可公度性, 可将评价指标进行标准化处理。目前对评价指标进行标准化处理有多种方法, 如线性插值法、目标打分法等。考虑到评价指标体系的定量指标中既有正指标(即指标数值越大越好), 又有逆指标(即指标数值越小越好), 因此在进行指标标准化时, 对定量指标可采用线性插值法<sup>[58]</sup>。

线性插值法是将某个指标下的最优值定为 100 分, 最差值定为 1 分, 其他数值通过代入直线方程求得。设某项指标实际发生数值中的最大数为  $f_{\max}$ 、最小数分别为  $f_{\min}$ , 则带量纲的评价值可通过线性插值公式转化为无量纲的评价值

$v_y$ 。对于正指标，线性插值公式为：

$$v_y = 1 + (f_y - f_{\min}) / (f_{\max} - f_g) \times 99 \quad (4-15)$$

而对于逆指标，线性插值公式则为：

$$v_y = 1 + (f_{\max} - f_y) / (f_{\max} - f_{\min}) \times 99 \quad (4-16)$$

在对各指标评价值经过无量纲化和极性变换后，采用加权平均法对绩效评价值进行量化综合评价。加权平均法是指标综合的基本方法，本文将采用此法。

通过以上介绍的方法可以计算出其评估值，但是对于计算出的评估值代表了什么样的绩效，还不好评判。因此，需要给出一个具体的分级标准，确定多大的值代表多高的绩效。根据指标标准值  $G_i$  和给出的权重  $B_i$ ，即可算出绩效标准值：

$$R = \sum_{i=1}^n G_i B_i \quad (4-17)$$

按级差为十的五级分级标准，本文得出如下绩效等级（如表 4-4 所示）：

表 4-4 绩效评估标准级差及评语

等级	绩效评估综合值	评语
第 1 级	R-100	绩效高
第 2 级	(R-10)-R	绩效较高
第 3 级	(R-20)-(R-10)	绩效一般
第 4 级	(R-30)-(R-20)	绩效较低
第 5 级	(R-40)-(R-30)	绩效低

## 第5章 政府税务部门绩效评价实证分析

鉴于整个政府系统范围过大，笔者结合实证分析，选取河北省沧州市国税局2006年（涉及到增减率的指标以2005年度为参照）度绩效状况，建立基于平衡计分卡的绩效评价指标体系。税务机关绩效评价体系包括财务、纳税人（相当于企业的顾客，是税务机关服务的对象，包括应该纳税的企业法人以及自然人）、内部业务流程、学习和成长四个方面。

### 5.1 评价指标体系的建立

根据卡普兰和有关权威咨询公司的调查，发现大多数平衡计分卡一般有20至25项衡量指标。四个维度的指标数量典型分配为：财务占5项衡量指标（24%），顾客占5项衡量指标（24%），内部业务流程占6项衡量指标（28%），学习与成长占5项衡量指标（24%）。本研究的指标数量基本上按照以上规律设计。

#### 5.1.1 指标体系的初选

事实上，大多数应用平衡计分卡的公司都将平衡计分卡作为公司战略的宣传和管理工具，平衡计分卡将公司的战略转化为公司日常的行动并将这些行动统一起来进行管理，以帮助公司早日实现战略。因而，建立平衡计分卡评价体系，首先要确定战略目标，然后将战略目标按四个维度进行分解，找出影响目标实现的关键因素，建立绩效评价指标体系。

##### （1）使命

税务机关依法征收各项税种，交给国家和地方政府，作为财政收入的主要来源，国家和地方利用财政收入进行市政工程建设、公益事业、政务管理等，推动经济发展，维护社会安定。另外，税务机关通过不同对象的税收差异，也可以直接调节经济分配，如：征收范围、减免退税等，从而起到经济杠杆作用，促进经济健康发展。税务机关还要承担社会责任，要体现为纳税人服务、为社会服务的宗旨，要体现社会公正、完善社会保障体系。所以，税务机关的工作目的是确保财政收入，调控经济发展，而它的价值既体现在对经济的贡献上，又体现在对社会的贡献中。

##### （2）机构设置

全国共设四级税务机构：国家税务总局为国务院直属正部级机构；省、自

治区、直辖市国家税务局、地方税务局为正厅级机构；地区、地级市(区)、自治州、盟的国家税务局、地方税务局为正处级机构；县、县级市(区)、旗的国家税务局、地方税务局为正科级机构。征收分局、税务所是县级税务局的派出机构，是直接从事税收征收管理的税务机关。

### (3) 战略目标

通过对基层税务机关使命、机构和职能的分析，可总结出其战略目标是在保证纳税人满意的前提下提高税收绩效，具体来说是：提高税收征收率，缩小税法和执法差距，减少税款流失，力求做到应收尽收；提高税收执法效率，优化税收征管流程，税收执法做到公开、透明和高效；提高纳税服务水平，为纳税人提供优质、便捷、全方位的纳税服务；降低税收成本，包括税务机关办税成本和纳税人依法纳税成本；提高决策水平，为国家宏观调控提供科学依据。

根据上述原则和要求，在分析税务机关主要职能的基础上，结合工作实际，拟定指标草案，财务、纳税人、内部业务流程、学习与成长的衡量指标数目分别为8项、9项、11项、10项（如表5-1所示）。

表5-1 沧州市国税局绩效指标体系初选集

财 务	收入	征收预计比率、税收增长率、税收计划完成率、新税源培植率、 税收计划准确率、
	成本	税收成本率、税收行政赔偿费用增减率、信息系统引进成本
纳 税 人	结果	纳税人满意度、纳税人投诉率、偷漏抗骗税增减率、纳税成本率
	过程	服务时间、纳税人意见采纳率、服务一次成功率、纳税人接受服 务成本降低率、税务机关形象与声誉
内 部 流 程	结果	纳税申报率、税务登记率、税收入库率、滞纳金加收率、欠税增 减率、处罚率
	过程	反应时间、分析投诉情况、提高服务效率、税务人员违纪增减率、 电子化服务率
学 习 成 长	结果	员工满意度、员工离职率、员工劳动生产率、个人贡献率、个人 挖潜率
	过程	信息网络覆盖率、员工建议采纳率、员工收入、税收信息系统使 用率、员工培训率

### 5.1.2 指标体系的确定

(1) 问卷调查。设计《沧州市国税局绩效评价调查问卷表》(详见附录 1),选定纳税人 20 人、沧州市基层税务机关员工 15 人作为问卷调查对象,发放问卷调查表 35 份(回收合格问卷调查表 31 份),对沧州市国税局绩效指标初选集进行初步筛选(如表 5-2 所示)。

表 5-2 沧州市国税局绩效指标体系复选集

财务	税收增长率、税收计划完成率、新税源培植率、税收计划准确率、税收成本率、税收行政赔偿费用增减率
纳税人	纳税人满意度、纳税人投诉率、服务时间、纳税人意见采纳率、纳税人接受服务成本降低率、纳税成本率
内部流程	纳税申报率、税务登记率、税收入库率、滞纳金加收率、欠税增减率、处罚率、分析投诉情况、税务人员违法违纪案件增减率、电子化服务率
学习成长	员工满意度、员工离职率、员工劳动生产率、个人挖潜率、信息系统使用率、员工建议采纳率、员工培训率

(2) 德尔菲法筛选。根据初步筛选的结果编辑问卷发给由 5 位绩效专家组成专家小组(详见附录 2),进行可行性分析和复选,将各位专家第一次判断意见汇总,算出每一个测评指标权重的平均值(如公式 3-1 所示),再分发给各位专家,让专家比较自己同他人的不同意见,修改自己的意见和判断。两轮修改后,专家们的意见趋于一致,得出最终指标体系(如表 5-3 所示)。

表 5-3 沧州市国税局绩效指标体系

财务 B1	税收计划准确率 C1	定量指标, 越大越好
	新税源培植率 C2	定量指标, 越大越好
	税收增长率 C3	定量指标, 越大越好
	税收成本率 C4	定量指标, 越小越好
	税收行政赔偿费用增减率 C5	定量指标, 越小越好
纳税人 B2	纳税人满意度 C6	定性指标, 越大越好
	纳税成本率 C7	定量指标, 越小越好
	纳税人意见采纳率 C8	定量指标, 越大越好
	纳税人投诉率 C9	定量指标, 越小越好

表 5-3 沧州市国税局绩效指标体系（续表）

内部流程 B3	纳税申报率 C10 税收入库率 C11 欠税增减率 C12 处罚率 C13 税务人员违纪增减率 C14 电子化服务率 C15	定量指标，越大越好 定量指标，越大越好 定量指标，越小越好 定量指标，越小越好 定量指标，越小越好 定量指标，越大越好
学习成长 B4	员工满意度 C16 员工培训率 C17 员工个人贡献率 C18 信息系统使用率 C19 员工离职率 C20	定性指标，越大越好 定量指标，越大越好 定量指标，越大越好 定量指标，越大越好 定量指标，越小越好

## 5.2 指标的综合评价

### 5.2.1 定量指标评价

#### (1) 财务维度

① 税收计划准确率  $C1 = (\text{年度税收计划收入数} \div \text{实际完成的税收收入数}) \times 100\%$ 。说明：如果完成年度计划收入，则该项指标按上述方法计算；如果未完成年度计划收入，则按 $(\text{实际完成的税收收入数} \div \text{年度税收计划收入数}) \times 100\%$ 计算。

② 新税源培植率  $C2 = (\text{新税源税收收入} \div \text{全年税收收入}) \times 100\%$

③ 税收完成增长率  $C3 = (\text{考核年度税收收入} - \text{上年度税收收入}) / \text{上年税收收入} \times 100\%$

④ 税收成本率  $C4 = (\text{征税成本} \div \text{税收收入}) \times 100\%$

⑤ 税收行政赔偿费用增减率  $= [(\text{考核年度行政赔偿费用} - \text{上年度行政赔偿费用}) \div \text{上年度行政赔偿费用}] \times 100\%$

#### (2) 纳税人维度

① 纳税成本率  $C7$  是每百元税收的纳税费用。该项指标值可从纳税人的会计资料中分析得到。

② 纳税人意见采纳率  $C8 = (\text{被采纳的纳税人建议数} \div \text{纳税人建议总数}) \times 100\%$

③ 纳税人投诉率  $C9=(\text{投诉次数}-\text{纳税人总数}) \times 100\%$

#### (3) 内部流程维度

① 纳税申报率  $C10=(\text{按期申报户次} \div \text{应申报户次}) \times 100\%$ 。说明：按期申报户次是指在税法规定的期限内实际办理纳税申报的户次(分税种评价)。应申报户次=期初开业户数+本期复业户数+本期开业户数+本期迁入户数+本期非正常转正户数-本期批准延期申报户数-批准本期停业户数-本期认定的非正常户数-本期法院正式宣布破产户数-非独立纳税户数-审定不达起征点户数-其他。期初指税款所属期的期初；本期指税款所属期。

② 当期申报应纳税款入库率  $C11=(\text{当期申报实际征收入库的税款} \div \text{当期申报应缴纳税款}) \times 100\%$ 。说明：分子为按照税法规定的各税纳税期限的当月实际征收入库的税额(包括退税金、批准延期申报核定的预缴税款)；分母为纳税人申报的应纳税款+批准延期申报核定的预缴税款-批准缓缴的税款。

③ 欠税增减率  $C12=[(\text{期末欠税余额}-\text{期初欠税余额}) \div \text{期初欠税余额}] \times 100\%$ 。说明：欠税增减率只按年度计算。期初欠税指考核期起始月初的欠税余额；期末欠税指考核期截止月末的欠税余额。

④ 处罚率  $C13=(\text{实际涉税罚款金额} \div \text{查补税款金额}) \times 100\%$ 。说明：分子为税务稽查查补税款并处罚款的处罚户次(金额)；分母为税务稽查未予以处罚与予以处罚的查补税款户次(金额)之和。罚款数额中不包括违反税务规定的行为处罚。

⑤ 税务人员违纪增减率  $C14=[(\text{考核年度税务人员违法违纪案件数}-\text{上年度税务人员违法违纪案件数量}) \div \text{上年度税务人员违法违纪案件数}] \times 100\%$ 。说明：该指标所指案件包括税务人员该为而不为、未按规定时限作为、未按规定标准作为、未按规定程序作为、贪污受贿、以税谋私案件等。

⑥ 电子化服务率  $C15=(\text{采用的电子化服务方式} \div \text{当前税务系统通用的电子化服务方式}) \times 100\%$ 。说明：评价项目主要包括在线网络报税、电话申报、手机短信申报、IC 卡申报、财税库横向联网、发票网上认证、金融机构批量划缴税款、稽查案件管理电子化、综合行政办公自动化系统、税收管理员系统、执法责任系统、综合征管软件应用系统、客户服务中心、信息服务网站、纳税大厅纳税人自助服务系统。

#### (4) 学习与成长维度

① 员工培训率  $C17=(\text{考核年度接受培训的员工数} \div \text{考核年度员工总数}) \times 100\%$

② 员工个人贡献率  $C18 = (\text{考核年度超额完成工作任务员工人数} - \text{员工总数}) \times 100\%$

③ 信息系统使用率  $C19 = (\text{税务内部信息系统使用人数} \div \text{员工总数}) \times 100\%$

④ 员工离职率  $C20 = (\text{考核年度员工离职人数} \div \text{员工总数}) \times 100\%$

根据以上公式和沧州市国税局的统计资料，得出定量指标评价表（如表 5—4 所示）。

表 5—4 沧州市国税局绩效定量指标评价

指标		评价结果	权重
财 务 B1	税收计划准确率 C1	95.59%	0.020 D1
	新税源培植率 C2	12.42%	0.045 D2
	税收完成增长率 C3	12.55%	0.090 D3
	税收成本率 C4	13.22%	0.068 D4
	税收行政赔偿费用增减率 C5	-70.00%	0.032 D5
纳 税 人 B2	纳税人满意度 C6		0.210 D6
	纳税成本率 C7	3.01%	0.120 D7
	纳税人意见采纳率 C8	92.00%	0.046 D8
	纳税人投诉率 C9	0.17%	0.085 D9
内 部 流 程 B3	纳税申报率 C10	98.44%	0.038 D10
	税收入库率 C11	99.57%	0.038 D11
	欠税增减率 C12	20.00%	0.038 D12
	处罚率 C13	86.00%	0.037 D13
	税务人员违纪增减率 C14	-82.11%	0.020 D14
	电子化服务率 C15	93.33%	0.021 D15
学 习 与 成 长 B4	员工满意度 C16		0.021 D16
	员工培训率 C17	27.27%	0.032 D17
	员工个人贡献率 C18	98.00%	0.009 D18
	信息系统使用率 C19	98.35%	0.010 D19
	员工离职率 C20	0.06%	0.020 D20

## 5.2.2 定性指标评价

### (1) 纳税人满意度 C6

① 确定因素集  $U=\{$ 服务质量 D1, 服务态度 D2, 工作效率 D3, 税务公开 D4, 勤政廉洁 D5 $\}$ , 评语集  $V=\{\text{很好}, \text{较好}, \text{一般}, \text{较差}, \text{很差}\}$ 。

② 根据问卷调查(详见附录1), 列出指标因子判断矩阵(如表5-5所示)。

表5-5 沧州市税务部门纳税人满意度指标因子判断矩阵

满意度	D1	D2	D3	D4	D5
D1	1.000	3.139	2.440	3.425	2.418
D2	0.319	1.000	2.113	3.805	2.818
D3	0.410	0.473	1.000	3.950	2.342
D4	0.292	0.263	0.253	1.000	1.402
D5	0.414	0.355	0.427	0.713	1.000

根据第4章对矩阵进行归一化, 得到特征向量  $w=\{0.382, 0.248, 0.192, 0.086, 0.092\}$ , 最大特征值  $\lambda_{\max}=5.356$ ,  $C.I.=0.089$ ,  $R.I.=1.120$ ,  $C.R.=0.079 < 0.1$ 。所以, 判断矩阵一致性合格, 得出指标权重  $A=\{0.382, 0.248, 0.192, 0.086, 0.092\}$ 。

### ③ 纳税人满意度评分

根据调查问卷(详见附录1), 就每一项因素指标评价统计结果如下(如表5-6所示)。

表5-6 沧州市国税局纳税人满意度统计调查结果

	很好	较好	一般	较差	很差
服务质量	0.71	0.17	0.07	0.05	0
服务态度	0.79	0.10	0.10	0.01	0
工作效率	0.70	0.20	0.05	0.05	0
税务公开	0.86	0.10	0.02	0.02	0
勤政廉洁	0.81	0.06	0.10	0.03	0

### ④ 建立评价矩阵

$$R = \begin{bmatrix} 0.71 & 0.17 & 0.07 & 0.05 & 0 \\ 0.79 & 0.10 & 0.10 & 0.01 & 0 \\ 0.70 & 0.20 & 0.05 & 0.05 & 0 \\ 0.86 & 0.10 & 0.02 & 0.02 & 0 \\ 0.81 & 0.06 & 0.10 & 0.03 & 0 \end{bmatrix} \quad (5-1)$$

### ⑤ 进行综合评价

$$B = A \cdot R = \begin{pmatrix} 0.382 & 0.248 & 0.192 & 0.086 & 0.092 \end{pmatrix} \cdot \begin{bmatrix} 0.71 & 0.17 & 0.07 & 0.05 & 0 \\ 0.79 & 0.10 & 0.10 & 0.01 & 0 \\ 0.70 & 0.20 & 0.05 & 0.05 & 0 \\ 0.86 & 0.10 & 0.02 & 0.02 & 0 \\ 0.81 & 0.06 & 0.10 & 0.03 & 0 \end{bmatrix} \quad (5-2)$$

$$= \begin{pmatrix} 0.750 & 0.131 & 0.072 & 0.036 & 0 \end{pmatrix}$$

结果显示：有 75% 的纳税人认为沧州市国税很好，13.1% 的纳税人认为较好，7.2% 的纳税人认为一般，3.6% 的纳税人认为较差，没有纳税人认为很差。

### (2) 税务人员满意度 C16

① 确定因素集  $U = \{\text{工资待遇 } D_1, \text{ 工组环境 } D_2, \text{ 人际关系 } D_3, \text{ 个人实现 } D_4\}$ ，评语集  $V = \{\text{很好, 较好, 一般, 较差, 很差}\}$ 。

② 根据调查问卷（详见附录 1），列出指标因子判断矩阵（如表 5-7 所示）。

表 5-7 沧州市税务部门员工满意度指标因子判断矩阵

满意度	D1	D2	D3	D4
D1	1.000	2.645	2.187	1.794
D2	0.378	1.000	1.909	1.483
D3	0.457	0.524	1.000	1.930
D4	0.557	0.674	0.518	1.000

对矩阵进行归一化，得到特征向量  $w = \{0.411, 0.236, 0.197, 0.156\}$ ，最大特征值  $\lambda_{\max} = 4.183$ ， $C.I. = 0.061$ ， $R.I. = 0.900$ ， $C.R. = 0.068 < 0.1$ 。所以，判断矩阵一致性合格，得出指标权重  $A = \{0.411, 0.236, 0.197, 0.156\}$ 。

### ③ 员工满意度评分

每一项因素指标评价统计结果如下（如表 5-8 所示）。

表 5-8 海沧州市国税局员工满意度统计调查结果

	很好	较好	一般	较差	很差
工资待遇	0.10	0.20	0.48	0.14	0.08
工作环境	0.36	0.32	0.24	0.06	0.02
人际关系	0.26	0.46	0.22	0.04	0.02
个人实现	0.14	0.32	0.48	0.02	0.04

### ④ 建立评价矩阵

$$R = \begin{bmatrix} 0.10 & 0.20 & 0.48 & 0.14 & 0.08 \\ 0.36 & 0.32 & 0.24 & 0.06 & 0.02 \\ 0.26 & 0.46 & 0.22 & 0.04 & 0.02 \\ 0.14 & 0.32 & 0.48 & 0.02 & 0.04 \end{bmatrix} \quad (5-3)$$

#### ⑤ 进行综合评价

$$B = A \cdot R = (0.411 \ 0.236 \ 0.197 \ 0.156) \begin{bmatrix} 0.10 & 0.20 & 0.48 & 0.14 & 0.08 \\ 0.36 & 0.32 & 0.24 & 0.06 & 0.02 \\ 0.26 & 0.46 & 0.22 & 0.04 & 0.02 \\ 0.14 & 0.32 & 0.48 & 0.02 & 0.04 \end{bmatrix} = [0.199 \ 0.298 \ 0.372 \ 0.436 \ 0.018] \quad (5-4)$$

结果显示：有 19.9% 的沧州市国税员工认为单位很好，29.8% 的员工认为单位较好，37.2% 的员工认为单位一般，43.6% 的员工认为单位较差，还有 1.8% 员工认为单位很差。

对于这两种满意度，一般认为其值应该达到“一般”级次以上，即满意度比例为很好、较好和一般的比例之和，因此有如下结论：沧州市国税的纳税人满意度为 95.3%，员工满意度为 86.9%。

### 5.3 评价结果分析

#### 5.3.1 确立评价标准

参考沧州市国税局历年统计数据，求平均值，得出各评价指标的标准参考值（如表 5-9 所示）。

表 5-9 评价指标的标准参考值

指标	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
标准值	96%	12%	12%	11.12%	-80%	98.54%	1.75%	80%	0.2%	100%
指标	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
标准值	100%	-80%	100%	-80%	90%	93%	26%	83.5%	97%	-80%

采用前文介绍的线性插值法对各指标标准参考值进行无量纲和无极性处理，得到（如表 5-10 所示）。

表 5-10 评价指标的无量纲和无极性标准参考值

指标	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
标准值	84.61	80.20	80.28	80.04	83.50	80.82	80.15	80.20	80.75	100
指标	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
标准值	100	83.50	100	82.84	83.50	80.11	82.88	83.10	81.07	82.3

根据表 5-10, 用加权平均求出相对绩效标准值:

$$R = \sum_{i=1}^{20} G_i B_i = 89.05 \quad (5-5)$$

其中,  $C_i$  表示指标的标准值,  $D_i$  表示指标权重。按级差为十的五级分级标准, 本文得出如下绩效等级(如表 5-11 所示)。

表 5-11 沧州市国税局绩效评估标准级差及评语

等级	绩效评估综合值	评语
第 1 级	89.05—100	绩效高
第 2 级	79.05—89.05	绩效较高
第 3 级	69.05—79.05	绩效一般
第 4 级	59.05—69.05	绩效较低
第 5 级	49.05—59.05	绩效低

### 5.3.2 确定绩效等级

采用前文介绍的线性插值法对税务机关各绩效指标评价值进行无量纲和无极性处理, 得出各指标最终可比较得分:  $C1=87.56$ ;  $C2=82.96$ ;  $C3=84.48$ ;  $C4=62.47$ ;  $C5=75.25$ ;  $C6=100.00$ ;  $C7=60.25$ ;  $C8=92.08$ ;  $C9=83.94$ ;  $C10=69.11$ ;  $C11=92.16$ ;  $C12=1.00$ ;  $C13=65.35$ ;  $C14=84.65$ ;  $C15=89.00$ ;  $C16=97.16$ ;  $C17=88.33$ ;  $C18=80.20$ ;  $C19=89.57$ ;  $C20=1.00$ 。

根据无量纲后的值和权重, 可以得到最终的综合评价分数:

$$R = \sum_{i=1}^{20} G_i B_i = 77.829 \quad (5-6)$$

对照表 5-11 可以得出, 沧州市国税局 2006 年度的绩效总体一般。将每项指标得分和标准值对比, 对差距大的指标进行整改, 以加强绩效水平。

## 第6章 全文总结与研究展望

### 6.1 全文总结

全文共分六章从理论、方法到实践对政府部门绩效评价的过程进行了改革，一二章做了背景铺垫，三四章实现了方法的集成和优化，第五章提高了体系的实际操作性。文章首先对政府绩效评价的相关背景进行了介绍，对国内外政府绩效研究文献进行了综述，给出了本文的研究框架和研究方法；接着介绍了政府部门绩效评价的概念、方法、产生、发展历程、业绩和问题，总结了发达国家在政府绩效评价方面的几点经验；然后介绍了平衡计分卡的产生、内容和优缺点，将其引入政府部门绩效评价，在对其进行可行性分析和优化后，建立了相应的四维指标体系，并且融合层次分析法和模糊综合评价法，完成了政府部门绩效评价的过程，并进行了无量纲和无极性处理，对评价结果进行了分析；最后结合沧州市国税局的实际情况，对模型进行了实证检验。

本文的创新之处在于其将现代企业广泛应用的绩效评价方法运用到政府这种非营利性部门中来，根据政府工作的性质，将其演化和融合，形成了一个比较科学、客观的绩效评价体系。具体地说：

第一，文章突破了现有的以“德、能、勤、绩、廉”为主要维度展开的绩效评价，把现代企业中广泛应用的平衡计分卡（财务、客户、内部流程、学习和成长）运用于政府评价中。这样既隐含了政府部门应该着重强调的工作能力、服务态度、工作效率、作风正派等因素，又用可衡量的指标表现出来，使绩效评价更加客观有效。

第二，评价指标的选取具有一定的广泛性，有利于其在其他公共部门的推广；评价的方法步骤清晰、明确，被评价部门容易找出不足之处，及时改进。

### 6.2 研究展望

本文虽然构建了一套比较完整的基于平衡计分卡的政府组织绩效评价体系，但很多研究还刚刚起步，这些理论和方法尚需要进一步深化、完善和补充。

进一步探讨基于平衡计分卡在公共组织绩效评价中的数量化方法的研究，可以将模糊数学、概率论、常微分方程等数学工具进一步引入到绩效评价之中，比如可以将平衡计分卡和经济增加值（EVA）结合起来，将平衡计分卡与数据包

络分析（DEA）结合起来，将神经网络与平衡计分卡结合起来，引入到政府部门的绩效评价之中。这些大胆的尝试都是未来研究的方向。

## 参考文献

- [1] [美]戴维·奥斯本, 彼得·普拉斯特里克. 摒弃官僚制: 政府再造的五项战略. 北京: 中国人民大学出版社, 2002: 45-183
- [2] Jos L.T. Blank, C. A. Knox Lovell. Public provision and performance: contributions from efficiency and productivity measurement. The Netherlands: Elsevier Science B.V., 2000: 277-296
- [3] Kathryn E. Newcomer. Using performance management to improve public and nonprofit programs. San Francisco, California: Jossey-Bass Inc, 1997: 780-803
- [4] Kaplan R.S., David P Norton. The balanced scorecard-measures that drive performance. Harvard Business Review, 1992(Jan/Feb): 71-79
- [5] Kaplan R.S., David P Norton. Putting the balanced scorecard to work. Harvard Business Review, 1993(Sep/Oct): 134-147
- [6] Kaplan R.S., David P Norton. Using the balanced scorecard as a strategic management system. Harvard Business Review, 1996 (Jan/Feb): 474
- [7] Kaplan R.S., David P Norton. The strategy focused organization: how blanced scorecard thrive in the new business environment. Boston: Harvard Business School Press, 2001: 45-48
- [8] 马克·波波维奇. 创建高绩效政府组织. 北京: 中国人民大学出版社, 2002: 67
- [9] Washnis G.J.. Productivity improvement handbook for state and local government. New York: John Wiley and Sons, 1980: 38-50
- [10] 母天学. 对美国政府绩效考评活动的考察. 行政论坛, 2001(5): 79-81
- [11] 卓越. 公共部门绩效评价. 北京: 中国人民大学出版社, 2004: 205
- [12] Grifel S.S.. Performance measurement and budgetary decision making. Public Productivity and Management Review, 1993, 16(4): 403-407
- [13] 中国行政管理学会课题组. 政府部门绩效评价研究报告. 中国行政管理, 2006 (5): 15
- [14] 徐绍刚. 建立健全政府绩效评价体系的构想. 政治学研究, 2004(3): 76-83
- [15] 韩强. 对建立和完善党政领导干部考核评价指标体系的若干思考. 政治学研究, 2003(4): 37-45
- [16] 倪星, 余凯. 试论中国政府绩效评价制度的创新. 政治学研究, 2004(3): 84-92
- [17] 韩兆洲. 政府政绩综合评价指标体系及实证分析. 暨南学报(人文科学与社会科学版), 2004, 26(10): 37-42
- [18] 范柏乃, 朱华. 我国地方政府绩效评价体系的构建和实际测度. 政治学研究, 2005(1): 84-95
- [19] 姜晓萍, 马凯利. 我国公务员绩效考核的困境及其对策分析. 社会科学研究, 2005(1): 22-24

- [20] 吴建南, 郭雯著. 绩效目标实现的因果分析: 平衡计分卡在地方政府绩效管理中的应用. 管理评论, 2004(6): 22-27
- [21] 张定安. 平衡计分卡与公共部门绩效管理. 中国行政管理, 2004(6): 87-90
- [22] 张国庆, 曹堂哲. “平衡计分卡”与公共行政执行的有效性. 湖南社会科学, 2005(2): 30-37
- [23] 文薇. 基于平衡计分卡的政府绩效管理系统构造研究: [硕士学位论文]. 成都: 西南财经大学行政管理专业, 2007: 20-48
- [24] 刘晓辉, 王红艳. 论我国公务员考核制度的完善. 哈尔滨商业大学学报, 2004(5): 71-72
- [25] 彭国甫, 李树丞, 盛明科. 应用层次分析法确定政府绩效评价指标权重. 中国软科学, 2004(8): 20-25
- [26] 刘泽双, 甄翠, 苏大伟. 公务员绩效考核模糊层次综合评价模型及实例. 西安理工大学学报, 2006(6): 438-441
- [27] 蔡立辉. 政府绩效评价的理念与方法分析. 中国人民大学学报, 2002(5): 93-100
- [28] 陈燕, 邓旭. 平衡记分卡用于政府绩效考核的思考. JTG, 2007(3): 151-154
- [29] 张一弛. 人力资源管理绩效. 北京: 北京大学出版社, 1999: 4-11
- [30] 王刚. 谁来考评政府. 中国新闻周刊, 2005(11): 20
- [31] Benchmarking Study Report. Serving the American public: Best practice in performance management. 1997 (June): 12-15
- [32] 秦立春. 我国基层政府公务员绩效考评工作刍议. 中国行政管理, 2003(4): 87-89
- [33] 盛明科. 论当前我国政府绩效评价价值取向的重塑. 行政论坛, 2004(6): 92-95
- [34] Carl Cz T.. The evolution of performance measurement in government. Journal of Cost Management, 2000(May/June): 18-26
- [35] 王淑红, 龙立荣. 绩效管理综述. 中外管理导报, 2002(9): 40-44
- [36] 马克·霍哲, 张梦中. 公共部门业绩评价与改善. 中国行政管理, 2000(3): 38-42
- [37] [美]罗伯特·E·霍尔, 马可·利伯曼. 经济学: 原理与应用(第2版). 北京: 中信出版社, 2003: 655-670
- [38] 白蔚秋. 正确评价企业的经营业绩. 工业企业管理, 2001(5): 75
- [39] 张小玲. 国外政府绩效评价方法比较研究. 软科学, 2005(5): 36-40
- [40] 周志忍. 当代国外行政改革比较研究. 北京: 国家行政学院出版社, 1999: 39
- [41] 帕特里夏·英格拉姆, 国家行政学院国际合作交流部编译. 公共管理体制革新的模式. 西方国家行政改革评述, 1999: 151-156
- [42] 李文艳, 陈通. 政府绩效评价的价值取向及我国政府绩效评价的完善. 唯实, 2004(6): 33-36
- [43] 王伟. 基于平衡计分卡的公共组织绩效评价研究: [硕士学位论文]. 天津: 天津大学行

政管理专业, 2005: 31-33

- [44] Johnsen A.. Balanced scorecard: theoretical perspectives and public management implications. *Managerial Auditing Journal*.2001, Vol.16, Iss. 6: 119-130
- [45] 温森特·怀特. 欧洲公共行政现代化: 英国的个案分析. 北京: 国家行政学院出版社, 1998: 247-250
- [46] 尼古拉斯·亨利. 公共行政与公共事务. 北京: 中国人民大学出版社, 2002: 311-330
- [47] 刘旭涛. 政府绩效管理——制度、战略与方法. 北京: 机械工业出版社, 2003: 5-9, 251
- [48] Zanetti C..Managing change: focus on improving service quality and information management. *Australian Journal of Public Administration*, 1998, Vol.57(5): 3-13
- [49] 秦扬勇. 平衡计分卡与绩效管理. 北京: 中国经济出版社, 2005: 15-19
- [50] 赵新伟. 平衡计分卡与绩效管理. *企业研究*, 2007(3): 75-76
- [51] 毕意文, 孙永玲. 平衡计分卡——中国战略实践. 北京: 机械工业出版社, 2004: 12
- [52] [瑞典]尼尔斯·奥尔韦, 卡尔·彼得里等. 使平衡计分卡发挥效用——平衡战略与控制. 北京: 中国人民出版社, 2004: 289-299
- [53] 李林, 肖牧, 王永宁. 将平衡计分卡引入我国政府部门绩效管理的可行性分析. *中国科技论坛*, 2006(4): 114-118
- [54] Breton.A.. Competitive governments: an economic theory of politics and public finance. Cambridge: Cambridge University Press, 1996: 556-575
- [55] 邓国胜. 非营利组织评价. 北京: 社会科学文献出版社, 2001: 112-113
- [56] 蔡志强, 卢厚清. 模糊数学在人力资源管理绩效评价中的应用. *数学的实践与认识*, 2006(7): 212-217
- [57] 杨纶标, 高英仪. 模糊数学——原理及应用. 广州: 华南理工大学出版社, 2002: 1-5
- [58] 李伟. 模糊理论在人才测评中的应用研究: [硕士学位论文]. 昆明: 昆明理工大学工商管理专业, 2005: 5-11

## 致 谢

经过几个月的艰苦写作，这篇硕士学位论文终于完成了。在论文的写作过程中，我得到了很多人的帮助和指点，使我的论文得以较为顺利的完成。

在这里，我首先感谢我的父母和朋友，谢谢他们在物质和精神上给予我的支持。

衷心感谢我的导师聂规划教授。回想在武汉理工大学的岁月，聂老师给予了我严师般的教诲和鼓励、父亲般的关爱和帮助。每当遇到学术及其他困惑无法参悟，聂老师无不谆谆教诲，拳拳之情难以言表。从论文的选题、文章框架的搭建、论文的写作到最终定稿，无不凝结着导师的汗水和心血。在导师的教诲下，我得到的绝不仅仅是知识，他在我人生的攻坚阶段为我点亮了一盏前进的灯塔！同时，非常感谢高新亚、杨琦峰、陈冬林、刘平峰等各位老师，在百忙的工作中为论文开题和答辩给予的指导和帮助。感谢武汉理工大学经济学院的各位领导和老师，特别是我的代课老师和指导老师们，正是由于他们的孜孜不倦、授业解惑才有了我今天学业上的进步与提高。也要感谢经济学院研究生办公室的各位老师，感谢他们的辛勤工作。

还要感谢河北省沧州市国税局的领导为我提供了研究所需的大量资料数据，帮助我完成了调查问卷和专家分析的工作。

最后，感谢我硕士研究生班上的全体同学和协助我完成论文问卷调查的师弟师妹们，感谢你们对我的关心、支持和帮助，很怀念大家一起共同学习和生活的日子。

论文的完成意味着两年多学习生涯的结束，但是今后对人力资源管理知识的学习和研究将是永无止境的。此时此刻我的心中充满了对母校的无限依恋，我将满载着对这段美好时光的追忆，加倍努力，在今后工作、学习等方面更上一个新的台阶。

## 附录

### 附录 1 沧州市国税局绩效评价调查问卷

本问卷是针对面向大众的沧州市基层税务部门绩效考核方面的设计的。基层税务机关是税务部门的窗口，是建设服务型税务机关的关键。专业的绩效评价会促使其更加认真、准确地执行税收法律法规，为纳税人提供更加周到的服务。因此，我们希望借助这次调查来研究基层税务机关提高绩效的可行措施，以便为基层税务机关的绩效管理提供借鉴。

我们希望您认真填写以下问卷题目(在没有注明“多选”的情况下，均为单选题)，在您认为对的选项上打“√”或者填入相关内容。谢谢您的积极参与和支持。

(背景知识：我们对税务机关这种政府部门的绩效评价分为组织评价、项目评价和工作人员个人评价。本问卷是针对组织整体的，即绩效评价是通过系统的方法、原理来评定和测量行政部门整体的工作行为和工作成果，它是管理者与员工之间的一项管理沟通活动。)

1、您的身份是：

- A 机关行政人员      B 企事业单位人员      C 相关专业在校学生  
D 专家学者      E 人民群众      F 其他(请注明)

2、您的性别是：

- A 男      B 女

3、您所在的单位\_\_\_\_\_您所在的部门\_\_\_\_\_您所从事的业务\_\_\_\_\_

4、您对基层税务机关绩效评价了解程度：

- A.熟悉      B.了解      C.听说过，但不太清楚      D.不了解

5、您认为在基层税务机关开展绩效评价有必要吗？

- A.有必要      B.没有必要      C.说不清楚

6、您对当前您所在地的税务机关绩效的总体评价是：

- A.好      B.较好      C.一般      D.较差      E.差

7、您认为当前制约基层税务机关绩效提高的最主要因素是：

- A.干部素质不高、作风不实      B.部门职能界定不清晰  
C.缺乏科学、有效的考核措施      D.缺乏有力的监督机制  
E.与纳税人沟通、为纳税人服务不够      F.其他(请注明)

8、如果从纳税人角度考虑，您认为应该从以下哪些指标进行绩效评价?(多选)

- A 纳税人满意度      B 纳税人投诉率      C 偷漏抗骗税增减率  
D 纳税成本率      E 服务时间      F 纳税人意见采纳率  
G 服务一次成功率      H 纳税人接受服务成本降低率  
I 税务机关形象与声誉      J 其他

9、如果从财务角度考虑，您认为应该从以下哪些指标进行绩效评价?(多选)

- A 征收预计比率      B 税收增长率      C 税收计划完成率  
D 新税源培植率      E 税收计划准确率      F 税收成本率  
G 税收行政赔偿费用增减率      H 信息系统引进成本      I 其他

10、如果从内部业务流程角度考虑，您认为应该从以下哪些指标进行绩效评价?(多选)

- A 纳税申报率      B 税务登记率      C 税收入库率      D 滞纳金加收率  
E 欠税增减率      F 处罚率      G 反应时间      H 分析投诉情况  
I 提高服务效率      J 税务人员违纪增减率      K 电子化服务率      L 其他

11、如果从税务人员学习与成长角度考虑，您认为应该从以下哪些指标进行绩效评价?(多选)

- A 员工满意度      B 员工离职率      C 员工劳动生产率  
D 个人贡献率      E 个人挖潜率      F 信息网络覆盖率  
G 员工建议采纳率      H 员工收入      I 税收信息系统使用率  
J 员工培训率      K 其他

12、您如何认为从纳税人、财务、内部业务流程、学习与成长四个方面对基层税务机关进行绩效评价的合理性?

- A 合理      B 有一定的合理性      C 不合理

13、您认为基层税务机关绩效评价主体应该包括哪些组织或个人?(多选)

- A 直接领导      B 税务机关自身      C 当地人大      D 当地党委  
F 纳税人      G 绩效专家      H 从事评价的中介机构      I 其他

请根据 14 和 15 题的提示，两两比较(前者相对于后者重要性)打分，打分标准如下：1 分表示两者一样重要；3 分表示前者比后者重要一点(稍微重要)， $1/3$  表示后者比前者重要一点(稍微重要)；5 分表示前者比后者重要(明显重要)； $1/5$  表示后者比前者重要(明显重要)；7 分表示前者比后者重要得多(强烈重要)； $1/7$  表示后者比前者重要得多(强烈重要)；9 分表示前者比后者极端重要(绝对重要)； $1/9$

表示后者比前者极端重要(绝对重要)。它们之间的数, 2, 4, 6, 8 也具有相应的类似意义。

14、从纳税人满意度角度而言，其下各指标相对重要性比较：

满意度	服务质量	服务态度	工作效率	税务公开	勤政廉洁
服务质量	1				
服务态度		1			
工作效率			1		
税务公开				1	
勤政廉洁					1

15、从税务人员满意度角度而言，其下各指标相对重要性比较：

满意度	工资待遇	工作环境	人际关系	个人实现
工资待遇	1			
工作环境		1		
人际关系			1	
个人实现				1

16、请对纳税人、财务、内部业务流程、学习与成长四个方面，按重要性由低到高排序：

①           ②           ③           ④

17、您认为目前基层税务机关绩效评价的主要问题是什么？

---

18、您对基层税务机关绩效评价有什么建议？

---

问卷设计人：武汉理工大学经济学院 赵璐

Email:zhaolu830127@yahoo.com.cn

## 附录 2 沧州市国税局绩效评价绩效专家调查问卷

我们在向河北省沧州市国税局税务人员和纳税人问卷调查的基础上，确立了一系列的绩效评价指标。现在，我们向绩效专家进行深入调查，此次调查的目的是最终确定基层税务机关绩效评价所使用的指标。为此，我们希望您认真填写以下问卷题目。谢谢您的积极参与和支持。为此，我们希望您认真填写以下问卷题目(选择题均为多选)，在您认为对的选项上打“√”或者填入相关内容。谢谢您的积极参与和支持。

1、如果从纳税人角度考虑，您认为应该从以下哪些指标进行绩效评价？

- A 纳税人满意度    B 纳税人投诉率    C 服务时间    D 纳税人意见采纳率  
E 纳税人接受服务成本降低率    F 纳税成本率

2、如果从财务角度考虑，您认为应该从以下哪些指标进行绩效评价？

- A 税收增长率    B 税收计划完成率    C 新税源培植率  
D 税收计划准确率    E 税收成本率    F 税收行政赔偿费用增减率

3、如果从内部业务流程角度考虑，您认为应该从以下哪些指标进行绩效评价？

- A 纳税申报率    B 税务登记率    C 税收入库率    D 滞纳金加收率  
E 欠税增减率    F 处罚率    G 分析投诉情况  
H 税务人员违法违纪案件增减率    I 电子化服务率

4、如果从税务人员学习与成长角度考虑，您认为应该从以下哪些指标进行绩效评价？

- A 员工满意度    B 员工离职率    C 员工劳动生产率    D 个人挖潜率  
E 信息系统使用率    F 员工建议采纳率    G 员工培训率

5、您对以上绩效评价指标有什么补充？

---

问卷设计人：武汉理工大学经济学院 赵璐  
Email:zhaolu830127@yahoo.com.cn

### 附录 3 攻读硕士期间发表的论文

- [1] Zhao Lu, Nie Guihua. Performance comprehensive evaluation of virtual R&D organization in small and medium-sized enterprises. The 4<sup>th</sup> international symposium of corporate governmence, 2007 (submitted)
- [2] Zhao Lu, Nie Guihua. The fussy evaluation model of environmental performance in green supply chain. the 1st International Symposium on Technology Innovation, Risk Management and Supply Chain Management, 2007: 1134-1135
- [3] Zhao Lu, Nie Guihua. Performance Evaluation of Technical Talents Based on Fussy Comprehensive Theory. The 4th International Conference on Innovation and Management(ICIM07), 2007 (submitted)
- [4] 赵璐, 聂规划. 网络新产品最优广告费用的制定. 统计与决策 (已收录于 2008 第三期)

## 附录 4 攻读硕士期间参与的项目

- [1] 武汉钢铁集团公司绩效考核创新项目（2007.9—2007.12）
- [2] 武汉市城市建设项目管理创新研究(编号：200702)（2007.1—2007.12）
- [3] 湖北省科技厅攻关项目“IT企业员工心理契约违背及其干预研究”（编号：2006AA401C11）（2007.1—2008.12）
- [4] “十一五”国家科技支撑计划“电子商务与现代物流共性集成技术研究开发”子课题（编号 2006BAH02A08）（2007.1—2007.12）
- [5] 国家自然科学基金“基于知识网格的电子商务智能推荐系统研究”（编号：70572079）（2006.1—2008.12）