

我国大学评价方法研究

王 琴

(中国科学院科技政策与管理科学研究所, 北京 100080)

【摘 要】近年来,大学评价工作越来越受到社会的关注,其深层原因在于:高等教育市场化走势越来越明显,学生作为高等教育服务产品的购买者,对服务机构信息需求不断增加,大学评价的兴起自在情理之中。但是目前的大学评价在方法上还不是很完善,如在评价指标体系的科学性、指标体系赋权的合理性、数据的可靠性以及计算方法的准确性上都存在一定的问题。本文首先对我国的大学评价研究现状进行了综述,并在此基础上分析了这些评价方法目前存在的问题,最后对目前评价方法中存在的问题做出了改进,并提出了新的一套大学评价方法。

【关键词】大学评价; 排名; 指标体系; 赋权方法; 大学分类

【中图分类号】G40

A Study of the Method for Appraising Chinese Universities

WANG Qin

【Abstract】 The ranking of Chinese university has been paid more and more intention to recent years. The profundity reason lies in that the higher education become more and more marketing. The students, who act as the purchaser of the higher education service product, has the higher demand for the information of service institution. Then, the spring up of appraising university is in the reason.

But now the method of appraising university is not perfect, for example the index, evaluation, data and calculate method have some faults.

This paper firstly to introduce the current situation of domestic universities evaluation and the method. Then it analysis the main problems of them. In addition, it improve these method and a new method for appraising Chinese Universities is put forward at last.

【Key words】 university evaluation; ranking; index system; weighting method; category of universities

1 引言

我国的大学评价,始于1954年中国人民大学等6所重点大学确立,这项评价工作分批进行了几次,到1963年国家确立了68所学校为全国重点高校,“文革”后经过恢复及相应的扩大,到1981年,被确定为重点高校的有98所。这是我国建国后对大学评价的最早的一种形式。1984年到1985年,经国务院批准,国家计委、教育部对重点学校再次作出评估,先后确定16所高等学校为国家“七五”期间重点建设项目,被高教界称为“重中之重”,国家拨专款改善学校和科研条件。1987年和1988年,国家教委在全国评选出416个重点学科进行重点建设与资助,涉及107所学校。1984年和1986年,国务院批准了33所大学试办研究生院。1986年,国家教委首批确认32所大学拥有教授任职资格的审批权。1995年至1996年,国家教委审定一批大学先期进

入“211”工程。1996年开始,国家教委高等教育司每年对若干所大学进行教学工作的评估。这些被评出来的学校,正在财力、物力、人力上得到了国家很大的支持,也倍受社会关注^[1,2]。

大学评价工作体现的是一种高校能力实力的社会认可,而且评价的指标和权重实际上是一种价值观,能够引导高校不断改善工作,加强其创新、教学能力,提高其社会声望。大学评价对于具有参谋功能即能对高校的现状进行诊断评估;也能起到一种社会导向功能、学习功能;评价工作对高校还有管理功能和监督功能;在一定程度上,如果高校发展实在不能让人满意,可以根据评价的结果对高校的资源和其他的方面做一些调整,这就是制衡功能。

大学评价的重要性引导人们不断探索其方法。大学评价主要包含四个方面:指标体系的设立、权重的选取、数据的收集、计算方法的选择。目前我国大学评价中,这四方面都不同程度地存在一些需要进

一步讨论的空间。

2 大学评价的由来及我国大学评价研究现状

2.1 我国大学评价指标体系的研究现状

在影响大学排行的诸要素中,指标发挥着基础的功能。指标的科学与合理性与评价结果的信度与效度直接关联。缺乏科学合理的指标,评价就难以公正合理,也很难以引导被评价者走向良性发展的轨道。指标只有建立在自身科学、合理的基础上,才能够发挥其评价、引导的作用^[3]。虽然参与评价的机构众多,但是我国迄今尚无较为统一与权威的评价指标体系。

1987~1989年,中国管理科学研究院科学学研究所先后提出三种指标体系。一是采用SCI收录的论文数进行评价。二是采用SCI和EI收录的论文数进行评价,三是采用“国外及全国刊物论文”、“专利批准”、“国家级奖”三个指标进行评价。

1991年开始,中国科技信息研究所提出的排名方法中,所用的指标是发表的论文数。即“在国外刊物上发表的论文数”、“在国内刊物上发表的论文数”、“国际期刊发表论文引用数”。这与后来1997年和2000年,中国科学院文献情报中心和南京大学的研究方法比较类似。后两者也是根据发表论文数评价大学,即中国科学院文献情报中心评出了中国科学引文数据库CSCD收录论文最多的前20所大学,指标是CSCD收录论文数,而南京大学《中文社会科学引文索引》(CSSCI)课题组根据的指标则是CSSCI。

学位与研究生教育评估所公布的大学排名中主要涉及3个一级指标:(1)研究生培养及质量;(2)学科建设及成果;(3)研究生院机构建设。而教育部科技司在2000年6月,进行了高校科技实力排名,涉及3个指标:(1)科研经费;(2)校办产业销售收入;(3)高校科技产业年销售收入超过亿元的企业^[4,5]。

目前最具有影响力的是网大和广东管理科学研究院的武书连的大学排名。1993年开始,广东管理科学研究院的武书连先生对7所大学排名,包括综合排名、单项排名及学科、专业、研究生院等排名,有2个一级指标:(1)人才培养;(2)科学研究下设子类指标。根据专家调查法确定各项指标的权重。而后来的网大评出的中国大学排名榜,沿用《美国新闻与世界报道体系》,设6个指标:学术声誉、专家评分、论文、新生质量、师资、科

研经费。每次各做相应调整^[6,7]。

笔者观察到他们的指标体系中有些问题,还很少有人提及,所以把它单列出来,希望能分析一下。如笔者认为,他们两者的指标体系中没有考虑过要计算负指标,根本没有扣分项,这与社会的发展并不相符,近几年,高校频繁的出现招生等丑闻,如果指标中没有这一扣分项,只在权重上加以调整,则不能很好的引导高校注重声誉,防微杜渐;他们二者的指标体系也没有注意到反映教学科研的质量,“质量”这一概念在高校评估中是很重要的,如果只注重一些表面的东西而忽视这种本质的质量,笔者认为是不可取的;二者都存在应该列入的指标却没有列入的问题,硬指标不全面,软指标也不完善,造成在指标体系上的片面性^[8]。

2.2 我国大学评价指标体系赋权方法研究现状

对于同样的基础数据,采用不同的权重确定方法,所得到的大学排名波动比较大,个别大学甚至波动了20多位。因此权重确定问题是一个很重要的问题,采取客观赋权更为合理。但是我国大学排名所确定权重的方法都是主观赋权法,大体上是两类:专家咨询法和层次分析法^[9]。

在武书连的评价中,使用的就是专家咨询法,但是遭到了许多非议,如顾海兵认为其“粗暴地把高校教员的人力总投入分为人才培养投入(57.09%)与科学研究投入(42.91%)”、“粗暴地把理科师生与文科师生分为8:2之数量比”^[10],认为他这样设立权重是没有道理的。

所以有些评价机构就开始绕开专家咨询法这条路走,而寻求层次分析法的帮助。

AHP法可对原始观测数直接加权运算进行综合排序,未削弱原始信息量,使评价指标逻辑判断量化且保持判断思维全过程的一致性。但构造各层指标的权重判断矩阵时,一般采用的是分级定量赋值,这可能会造成同一类中同一指标是另一指标的5倍、7倍,甚至是9倍,从而影响了权重的合理性^[11]。

所以现阶段各大学排名榜的权重的确定都存在一定的的问题,使用专家咨询法对于权重的计算分配更容易带有决策者的主观个人色彩,但是现在的大多数的评价者都倾向使用这种方法,因为这种方法最简单。

有些评价机构即使能考虑使用AHP法,也没有考虑到各决策者(专家)权重,用以衡量各评审专家在本次评审过程中的重要程度,即各评审专家的决策不是均权。所以一般采用同等重要进行处理专家的权

重。这些问题都直接影响到最后排名的结果^[12]。

3 大学评价方法的一些改进

前面已经提到我国的大学评估在指标体系、权重确定和数据的准确性上都存在一定的问题,对此,笔者提出一些改进的意见。

3.1 对指标体系的改进

前面提到了现在的指标体系存在的一系列问题,所以,笔者提出了一个新的指标体系来对现存的指标体系进行改进。

(1) 首先,现存的指标体系有许多地方是值得借鉴的,这基本都在新的指标体系中体现出来了。新的指标体系考虑了声誉、投入、产出的多方位评价,定性与定量相结合,重视相对量;既重视现在,除评价公布年度前1~2年的成绩外,又把积累性的指标,如两院院士、长江学者、博士点、硕士点、国家重点学科、研究基地等列入评价指标。

(2) 新的指标体系一共是三级指标,对教育、科研、声誉等进行排名。考虑了数量与质量、规模与效益等方面的关系。考虑了负指标。以往评估的指标体系中没有考虑过负指标,虽然很多专家提到过这一点,后来的指标体系改进也最多是在权重上做了一些调整,这不能从根本上解决问题。本指标正式把负指标列入了指标体系中,对各方面的情况综合考虑。

(3) 新的指标把质量评估的这一概念引进指标体系,办学质量评估中教育质量与科研质量分别评估这一方面保证了高校具有基本的办学条件和基本教育质量;另一方面又鼓励少数著名大学在高层次人才培养上进行学术竞争,形成社会影响,且还在一定程度上发挥评估的导向作用,防止大学过分重科研功能轻教育功能。

在教育质量中增加了大学的学科建设。这种做法不仅增加了可比性,使评估结果更具说服力,而且还引导学校走内涵发展道路,办好自己的优势学科,同时一些名牌大学的劣势学科也不获得高于实际的声誉,这进一步维护了评估的客观、公正性,对非名牌大学起到了很好的激励作用^[13]。

教育质量其实不仅仅包含着老师对学生的教育,也包含学校的管理情况,而这种管理情况在很大程度上是与教育管理者有关的,如果学校机构重叠,干部臃肿,人浮于事,效率低下,干群关系紧张,教师积极性受压抑,改革步履维艰,发展十分缓慢,所以本指标体系加入了“管理队伍”的评价指标,从各职

能部门运行绩效、院系部门运行绩效;学生满意度、教师满意度;管理人员占有教职工人数比例;是否企事分开、后勤工作社会化;本届校长任职期间学校发展状况等五个方面加以评价。

(4) 大学评价不能忽视总结性评价的“特殊的功能”^[14],所以仍要进行必不可少的鉴定和教学核定,以检验教育实施的效果。但是大学评价更为注重的是形成性评价,即直接指向正在进行的活动的。正如 Catherine A Palomba 和 Trudy W Banta 所说“对结果进行评价意味着终结;而对过程进行评价则暗示着还有改进的时间和机会”^[15]。

科研条件这个指标里增加了技术转让合同数和当年成交金额总数,符合现在发展了的大学变化情况,这是一个形成性评价。前面的“管理队伍”中“本届校长任职期间学校发展状况”也是一个发展性的指标,属于对过程进行评价。

(5) 新的指标体系借鉴美国的大学排行榜中对美国大学本科排行的标准与权重中的“财力”这一项指标。这一指标考查花在每个学生身上的平均经费,包括在教学、研究、学生服务相关教育支出方面的投资。因此,新指标在教育经费中增加了“生均教育事业费”这样,更能反映为学生提供学习的物质条件,也便于与国际接轨^[16]。

(6) 本指标体系把声誉情况分三部分考虑,即学术声誉、国内声誉和国际声誉。学术声誉是一些定量的打分,而国内声誉分为社会认同度和校友满意度两个下级指标,社会认同度又分为社会人员的认同度、学生的认同度、国内同行的认同度三个下级指标,这样的软指标一进入指标体系就更加完善与充实。而国际声誉部分主要参考的是国际同行的评价和留学生、国外交换学生占全校学生比,以往的指标体系中忽略了这几点,以至于局限性很多。

(7) 在负指标中主要是一些减分的指标,这些年,随着高校的不断发 展,人们对高校的期望值越来越大,高校中的一些丑闻也受到了人们的关注。如果在指标体系中没有这方面的体现,则整个体系很不完整。因而,笔者在综合考虑各方面因素后,决定把负指标纳入其中。

本指标体系只是涉及国内的一些高校的情况,不包括港澳台等地方的高校以及国际上的高校,另外一些如研究生院的排名也没有列入其中,这些都有待下一步指标体系的扩充和完善。本指标体系,只是作为一个评价方面的探讨,随着我国教育统计数据库的完善,指标体系还有可以进一步的调整。

另外,在做完今年大学排名榜的工作后下一步需要做的是,进行有效的预测,结合实际情况和一些有用的模型,对高校一些方面的情况作一个比较科学的预测。

3.2 对大学评价指标体系赋权方法的改进

目前的大学评价常常用 Delphi 法、专家调查法、层次分析法等来确定权重系数,这几种方法都属于主观赋权法。这类方法人们研究较早,也较为成熟。它的优点是简便易行,它的缺点是主观性太大,由主观赋权法确定出的权重系数真实与否,很大程度上取决于专家的知识结构、工作经验及其偏好,权重的确定存在不少主观随意性,这将严重影响评价结果的客观性;同时主观赋权法在评价过程中透明性比较差,而且在一定的时间区间内,权重系数具有保序性(在一定时间不同时刻处各指标对整体而言的重要性保持不变)和可继承性。

在评价指标体系赋权上,笔者改进了几点。

第一点,笔者改进了目前大学评价单一使用主观赋权法或客观赋权法,而提出了使用组合赋权法。

第二点,在主观赋权法中,笔者改进了传统的 AHP 法,因为传统的 AHP 法没有考虑到专家的权重,而在笔者的方法中,把这一项考虑进去了。

第三点,在客观赋权法中,笔者考虑到了指标相关性的问题,把指标相关性检验引入进来,让指标的设置更规范和合理。

第四点,在进行组合赋权法时,笔者对传统的组合赋权的两种方法进行了综合,提出了一个新的组合赋权法的计算公式。

这一套指标体系的赋权方法,是针对大学评价中专家繁多、一些指标缺乏样本数据、指标体系中含有大量定性指标等问题提出来的。本方法综合考虑了这些问题,希望能对此有个好的解决。笔者认为,在大学评估中采用此方法确定各指标的权重,能够更为准确地计算出权重系数,并且其结果较之其他赋权方法更为完善和更具有灵活性、科学性和可信度。因而有着极强的应用和推广价值。

4 结束语

大学排名对于促进我国的高等教育向着健康的方向发展,具有积极的导向作用。中国的大学排行榜却引起了国内专家、学者们的异议。笔者通过定性与定量分析认为:由于中国大学评价工作中所使用方法的缺陷导致了中国大学评价的不准确。

研究大学评价方法还是很重要的,一些偏激的提

附表

高校综合排名指标体系

	一级指标	二级指标	三级指标
正 指 标	教育条件	教育资源	图书馆藏书量及生均藏书量
		校园硬件设施	生均校园建筑面积
			生均教室占有面积
		经费情况	总量教育经费,生均教育费用
		教师资源	教授比例
			长江学者比例
			院士比例
			正高级以上人数比例
			博导比例
			博士学历人数比例
	教育质量	学科建设	国家重点学科
			博士、硕士学位授权一级学科
			硕士学位授权学科、专业
		教育基地	理科基地、文科基地、工科基地
			教育部重点实验室
			教育部工程研究中心
			教育部社科重点研究基地
		人才培养	本科毕业生人数
			博士点、硕士点、博士后流动站
		教师质量	授权专利、专著情况
			高校青年教师奖
			国家杰出青年科学基金
			跨世纪优秀人才
		学生质量	学生就业率
			全国英语六级通过率
			国际级奖、国家级奖、省部级奖、发明情况、拥有专利、专著情况
			全国百篇优秀博士学位论文入选情况
	科研条件	经费情况	总量科研经费,人均科研经费
		科研人员	科研机构人员比例
负 指 标	科研质量	科研基地	国家重点实验室,重点研究基地,国家工程研究中心,国家实验室,国家技术转移中心,国家级大学科技园和文化科技园数量,国家批准的文科研究中心,人文社科重点研究基地
		重大项目	获国家自然科学基金资助的项目,获国家社会科学基金的项目,973项目,863项目
		技术转让	技术转让合同数
			当年成交金额总数
		学术成果	SCI, S & N, EI, SCICITA, CSCD, CSCDCITA, SSCI, A & HCI, CSSCI, 及新华文摘发表论文的质量
	声誉水平	学术声誉	专家打分
		社会声誉	社会人员的认同度,学生的认同度,国内同行的认同度,校友的满意度
		国际声誉	留学生、国外交换学生占全校学生比
			出席国际学术会议
			到国外进行访问研修,向国外派遣研究生
负 指 标	教育质量		补考率,违纪率,专业设置不合理
	科研质量		科研人员违纪
	声誉水平		严重有损声誉的事件,在国家权威媒体上有过负面报道

高校竞争力排名指标体系
投入表

一级指标	二级指标	三级指标
教育条件	教育资源	图书馆藏书量及生均藏书量
	校园硬件设施	生均校园建筑面积
		生均教室占有面积
	经费情况	总量教育经费,生均教育费用
	教师资源	教授比例
		长江学者比例
		院士比例
		正高级以上人数比例
		博导比例
科研条件	经费情况	总量科研经费,人均科研经费
	科研人员	科研机构人员比例

产出表

一级指标	二级指标	三级指标
教育质量	学科建设	国家重点学科
		博士、硕士学位授权一级学科
		硕士学位授权学科、专业
	教育基地	理科基地、文科基地、工科基地
		教育部重点实验室
		教育部工程研究中心
		教育部社科重点研究基地
	人才培养	本科生毕业人数
		博士点、硕士点、博士后流动站
	教师质量	授权专利、专著情况
		高校青年教师奖
		国家杰出青年科学基金
		跨世纪优秀人才
	学生质量	学生就业率
		全国英语六级通过率
		国际级奖、国家级奖、省部级奖、发明情况、拥有专利、专著情况
		全国百篇优秀博士学位论文入选情况
科研质量	管理队伍	管理部门总数,管理人员中获高级职称的人数比例
	科研基地	国家重点实验室,重点研究基地,国家工程研究中心,国家实验室,国家技术转移中心,国家级大学科技园和文化科技园数量,国家批准的文科研究中心,人文社科重点研究基地
	重大项目	获国家自然科学基金资助的项目,获国家社会科学基金的项目,973项目,863项目
	技术转让	技术转让合同数
		当年成交金额总数
	学术成果	SCI, S & N, EI, SCICITA, CSCD, CSCD-CITA, SSCI, A & HCI, CSSCI, 及新华文摘发表论文的质量

法如废除大学排名等,笔者是不赞成的,尽管目前国内大学评价体系还不够完善,但因其排行榜所产生的社会影响关系到一所学校的知名度和美誉度,并且具有增进社会公众对大学的了解、引导学生选择报考学校和社会投资流向、促进社会公众及高校关注高等教育质量问题、推动大学之间展开合作与竞争等方面的作用,因此高校必须重视大学评价及其

方法研究和利用它来为学校发展服务,我们不能因噎废食。但我们必须同时强调,从事大学排名研究工作必须要有合理的评价体系和正确的数据来源,这样的大学排名榜才有它的实际利用价值。虽然目前的大学排名指标体系还需进一步完善,但只要我们能以科学求实的态度去做好大学排名的任何一个环节的工作,我们一定能找到一个解决大学评价的好方法。

本文正是基于对我国大学评价的关注,介绍了大学评价方法中最重要的两大因素:指标体系和赋权方法的现状,并对其存在的问题进行了分析,在此基础上提出了一些初步的看法和改进的建议,希望通过这样的讨论对我国大学评价的发展起到一定的参考和推进作用。

参考文献

1 李志仁.我国应开展高等学校综合评估并选优排序[J].科研评价与指标,2003(1):20-23.

2 辛彦怀,胡小元.我国大学排序的历史与现状[J].高等教育研究,1998(2):74-77.

3 戴跃欣.大学评价体系及其排行榜透视[J].扬州大学学报,2003(1):12-16.

4 邱均平,宋恩梅.我国大学评价的现状 & 主要问题分析[J].图书馆论坛,2003(6):56-61.

5 武书连.“中国有多少排行榜”[EB/OL].http://www.edu.cn/20010917/3001903.htm,2007-10-11.

6 http://rank2001.netbig.com,2007-10-20.

7 武书连,吕嘉,郭石林.2004 中国大学评价[J].科学与科学技术管理,2004(1):61-63

8 http://cssd.acca21.org.cn/2003/news0220.html[EB/OL],2007-10-25.

9 王凌峰,郭丰.评说大学排行榜[N].科技日报,2003-02-12(7)

10 顾海兵.陷入误区且步入禁区的“中国大学排名”[J].中国高等教育,2003(7):22-24.

11 丁和根.传媒竞争力评价指标体系研究[EB/OL].http://www.xinwenjie.com.cn/tgch/20050318/20050318173852.htm,2007-10-25.

12 彭本红,孙绍荣,胡若.研讨厅中专家的权重确定及方案评价研究[J].上海理工大学学报,2004(2):160-163.

13 刘晓红.国外高等教育评估制度对我国高教评估的启迪[J].北京大学学报,2004(1):14-16.

14 陈玉昆.中国高等教育评价论[M].广州:广东高等教育出版社,1993.

15 Catherine A Palomba,Trudy W.Banta.Assessment Essential-planning,Implementing and Improving Assessment in Higher Education[J].Jossey-bass Inc.Publishers 1999(4).

16 兰云,蔡言厚.完善“中国大学排名榜”的若干建议[J].现代大学教育,2003(3):109-111.

作者简介

王 琴 女,中国科学院科技政策与管理科学研究所硕士。研究方向:评价与管理,公共政策。