

基于数据包络分析的地方高校科研绩效评价

沈立宏 赵怡

【摘要】 本文运用 DEA 的方法,对我国 31 个省份(直辖市)地方高校科研投入产出绩效进行了评价和排名,结果表明,全国地方高校科研绩效水平存在较大差异,科研绩效提升在规模上的控制相比于纯技术控制存在更大的改进空间。地方高校科研绩效与区域内的文化、经济、高等教育发展水平无必然联系,地方政府的支持力度,高校对本地区高水平科研人才的吸引力,科研政策科学性、有效性才是影响地方高校科研绩效水平的关键因素。

【关键词】 数据包络分析 地方高校 科研 绩效评价

【收稿日期】 2015 年 5 月

【作者简介】 沈立宏,宁波大学科学技术学院助理研究员;赵怡,宁波城市职业技术学院学生处处长助理、助理研究员。

一、引言

地方高校是指隶属于各地方政府,主要依靠地方财政拨款经费的普通高校。截止 2012 年底,在全国 2442 所普通高校中,由教育部及其他部委主管的为 113 所,其余的 2329 所地方高校占全国普通高校数的 95.37%^[1],成为我国高等教育事业的中坚力量。高校是实施创新驱动发展战略的一个重要科研载体,作为我国高等教育事业中的主体,地方高校以科技创新提升地区竞争力,服务区域经济,为地方经济发展提供智力支撑。

特别是近年来,地方政府对高校科研投入不断增加,地方高校的科研拨入经费已由 2007 年的 186.60 亿增加到了 2012 年的 430.98 亿元,短短 5 年间增长了 2.31 倍,年均增长率达到了 18.2%。面对这巨额的资金投入,如何提高资金利用率,增强科技成果的产出效能,是摆在政府和高校管理者面前的一个重要课题。这就需要在一定周期内,采用科学的评价方法对地方高校科研活动的投入产出绩效做定量和定性分析,得到科学、严谨、客观的绩效评价结果,进而实现对科研资源的合理配置。

国内外有关高校科研绩效的评价研究主要有两个视角:一个是单维的输出视角,只考量高校科研的产出,包括科研成果的数量和质量,并以此作

为大学科研能力排名的重要指标,如 Sombatsompop 等根据 web of science 数据库,采集了科研论文发表量、引用量、单篇文章引用率等指标数据,对泰国 24 所国立大学绩效进行了评估,并做了排名^[2]。另一个是双维视角,即投入/产出的效率视角。投入包括科研活动所需的人、财、物,产出则包括科学研究产生的各类科技成果。如王晓红等将人力资源配置、科技经费与项目作为科研活动的投入量,将科技成果、科技交流、科研奖励与收益作为科研活动的产出量,运用多指标综合评价法,选取了 57 所高校作为评价单元,进行科研绩效评价^[3]。

在具体的研究内容上,对高校科研绩效评价的研究主要集中于以下几个方面。第一,效率评价指标的选择和数据来源,不同的考评指标和数据源会产生不同的评价结果^[4]。如 Kao Chiang 等在对台湾 168 所大学管理类专业科研绩效进行评价时,选择了论文出版量、文献引用数、台湾国科会(Nation Science Council, NSC)奖励获得数作为考评的指标,发现评价结果同仅以论文数量和引用量作为指标的考评存在差异^[5]。在国内,戚湧等人将人力资源、科研平台、科研项目、经费投入作为投入指标,将论文与专著数、成果鉴定与获奖、知识产权与成果效益作为科研产出指标,并

本文系浙江省软科学计划项目“浙江省地方高校科研投入产出绩效研究——基于 DEA 的方法”(2014C35037)。

做了两者间效率的衡量^[6]。第二,对评价单元选取的研究,可分为几个层次,第一类是以某一专业或学科作为科研绩效的评价单元;第二类是以学校作为评价单元,来做学校的整体科研绩效;第三类是将某一区域内的高校作为评价单元。第三,对评价方法的研究,可分为定性研究和定量研究两大类。定性研究包括同行评议、360度绩效评价等;定量研究包括文献计量法、主成分分析法、层次分析法,如:史兴伟运用计量评价理论和方法开展高校科研成果量化评价^[7],郭银清运用层次分析法确定各指标的权重,借鉴成本收益法的思路,对高校科研效益综合评价模型进行了探讨^[8],还有诸如数据包络分析法、神经网络法、灰色决策评价法、模糊综合评价法等新兴的管理学评价方法都在科研绩效评价领域中有所应用。方法的创新为科研绩效评估提供了新的思路、新的工具。

总的来看,过往的科研绩效研究无论在指标选择,还是在方法运用上都有了较为详尽的探讨,但在研究对象上往往过多集中于某一专业领域或者单所高校,而对地区内高校整体科研绩效的研究相对缺乏,在为数不多的讨论区域高校科研绩效时,也往往将区域内部委高校和地方高校作为一个整体研究单元,而忽视了两之间巨大的差异性。本文将聚焦于地区内的地方高校,选取适当的评价指标,运用数据包络分析法,对各地区内地方高校科研投入产出绩效做出评价,分析原因,以期帮助决策者和管理者制定和实施更为科学有效的科研政策。

二、方法与数据

1. 研究方法。

数据包络分析法(data envelopment analysis, DEA)是美国运筹学家 Charnes 等人于 1978 年提出的一种效率评价方法。它无需考虑由输入到输出的中间环节和中间数据,只需使用最初和最终的各项数据,对同类决策单元(decision making unit, DMU)进行相对有效的评价。作为一种非参数评估方法,可以不预先估计参数或权重,这样在避免主观因素、简化算法方面有着不可比拟的优越性^[9]。DEA 方法一经问世,就受到了各个领域的关注。在数十年间,DEA 方法已经成为现代决策、评价领域中一项非常重要的研究工具。

2. 数据。

本研究主要采用《2013 年高等学校科技统计

资料汇编》的 2012 年有关数据中各地区地方高校科研相关数据。

三、地方高校科研绩效评价实证分析

1. DEA 模型与决策单元选择。

数据包络分析运用的关键是确定模型,经典的有 CCR 模型和 BCC 模型。CCR 模型假设规模收益不变(Constant Return to Scale, CRS),得出的技术效率包含了规模效益的成分,是一种综合技术效率模型。假设有 n 个 DMU,每个 DMU 有 m 种投入,记为 $x_i (i=1, 2, \dots, m)$,投入权重表示为 $v_i (i=1, 2, \dots, m)$;有 q 种产出,记为 $y_r (r=1, 2, \dots, q)$,投入权重表示为 $u_i (i=1, 2, \dots, q)$,要测量 DMU 记为 DMU_k ,其线性规划模型经等价转换和对偶处理,得到的对偶模型可表示为:

$$\left\{ \begin{array}{l} \min \theta \\ \text{s. t. } \sum_{i=1}^n \lambda_j x_{ij} + s^+ = \theta x_{ik} \\ \sum_{i=1}^n \lambda_j y_{ij} - s^- = \theta y_{ik} \\ \lambda_j \geq 0, j = 1, 2, \dots, n \\ s^+ \geq 0, s^- \geq 0 \end{array} \right.$$

其中, λ 表示 DMU 的线性组合系数, s^- 为引入的松弛变量, s^+ 为引入的剩余变量, θ 代表决策单元的效率值。用 θ^* 来表示模型的最优解, θ^* 的范围为 $(0, 1]$, $\theta^* = 1$, 表明评价单元 DMU_k 处于技术有效状态,此时 $s^+ = 0, s^- = 0$ 。当 $\theta^* < 1$ 则说明评价单元 DMU_k 处于技术无效状态。

纯技术效率的 BCC 模型是在 CCR 模型的基础上增加了约束条件 $\sum_{j=1}^n \lambda_j = 1 (\lambda \geq 0)$ 而构成,它基于规模收益可变 (Variable Return to Scale, VRS),能将规模效率排查在外,得出的技术效率排除了规模的影响,因此又可称为“纯技术效率”,其对偶模型可表示为:

$$\left\{ \begin{array}{l} \min \theta \\ \text{s. t. } \sum_{i=1}^n \lambda_j x_{ij} + s^+ = \theta x_{ik} \\ \sum_{i=1}^n \lambda_j y_{ij} - s^- = \theta y_{ik} \\ \sum_{j=1}^n \lambda_j = 1 \\ s^+ \geq 0, s^- \geq 0, \lambda_j \geq 0, \\ j = 1, 2, \dots, n \end{array} \right.$$

而评价单元规模效率的值就是 CCR 模型效率值

与 BCC 模型效率值两者的比值^[10]。同时可通过 CCR 模型,来判断 BCC 模型中各类决策单元所处的规模收益状态:

若在 CCR 模型中的所有最优解中, $\sum \lambda^* < 1$,则说明该决策单元处于规模递增状态,当投入量按同一比例发生变化时,产出量会以较大比例发生变化。当决策单元处于此阶段时,决策者应考虑增加投入量,扩大投入规模。若在 CCR 模型中的所有最优解中, $\sum \lambda^* > 1$,则说明该决策单元处于规模递减状态,当投入量按同一比例发生变化,产出量会以较小比例发生变化。当决策单元处于此阶段时,决策单元应考虑降低投入量,减小投入规模。若在 CCR 模型中的所有最优解中, $\sum \lambda^* = 1$,则说明该决策单元处于规模收益不变状态,当投入量按同一比例发生变化时,产出量的比例将不发生变化,处于最优状态。

因而决策者可以依据决策单元当前所处的规模收益状态做出相应的决策调整和规模控制,实现资源的优化、高效配置。

本研究将采用以上两种模型,以各省地方高校作为评价单元,综合计算各评价单元的综合效率值(TE)、纯技术效率值(PTE)、规模效率值(SE)以及规模收益状态(RTS),探讨各地区地方高校科研绩效。

2. 高校科研投入产出绩效评价指标选择。

科研投入一般是指用于投入科研活动所需的人、财、物等资源,而科研产出则相对较为复杂,在学术界存在不同的观点。基于评价指标“系统性、可行性、适用性、有效性、可行性”的原则,根据《高等学校科技统计资料汇编》中的具体指标,结合其他学者的研究成果,将“科研人员合计数(人)、科

技投入支出经费(千元)”两项作为输入指标;将“发表学术论文数(篇)、出版科技著作数(千字)、知识产权授权数(项)、科技成果技术转让金额(万元)”4项指标作为输出量。各个输入、输出指标单位量纲不同,魏权龄对此做了数学证明:决策单元的效率指数与各项投入数据与产出数据的量纲选取无关^[11],因此本文将直接采用《高等学校科技统计资料汇编》中有关数据,而不再对数据进行标准化处理。同时将 31 个省作为科研绩效评价单元(DMU),符合 DEA 模型对 DMU 数量的经验法则,即一般而言 DMU 数量不应少于投入和产出指标数量的乘积,同时不少于投入和产出指标数量之和的 3 倍^[12]。

3. 实证分析。

对各省地方高校 2 项输入指标、4 项输出指标,应用 DEA 计算软件进行综合效率值(TE)、纯技术效率值(PTE)、规模效率值(SE)以及规模收益状态(RTS)测量,得到各省地方高校科研绩效评价结果表,具体见表 1。

从表 1 可以看到,在全国(大陆地区,下同)31 个省份中,有 10 个省份的地方高校科研绩效综合技术效率处于 DEA 有效状态,占总体的 32.26%,有 16 个省份的地方高校科研纯技术效率处于 DEA 有效状态,占比 48.48%,有 10 个省份处于最佳规模状态,占比 32.26%,可以发现纯技术有效个数要明显高于综合技术有效和最佳规模状态个数,表明我国地方高校科研绩效提升在规模上的控制相比于纯技术控制存在更大的改进空间。

从地区分布看,东部地区综合技术有效的地方高校有 4 个,占比 36.36%,中部地区有 3 个,

表 1 各省份地方高校科研绩效评价结果

省份	综合技术效率	纯技术效率	规模效率	规模收益	省份	综合技术效率	纯技术效率	规模效率	规模收益	省份	综合技术效率	纯技术效率	规模效率	规模收益
北京市	0.76131	1	0.76131	递减	安徽省	0.774353	0.804132	0.962967	递减	甘肃省	1	1	1	不变
福建省	0.637414	0.643099	0.991159	递增	广西	0.878519	0.946431	0.928244	递减	贵州省	0.766622	0.849632	0.902299	递减
广东省	0.749897	0.751902	0.997333	递增	河南省	1	1	1	不变	宁夏	1	1	1	不变
海南省	1	1	1	不变	黑龙江省	1	1	1	不变	青海省	0.513982	0.56781	0.905201	递增
河北省	0.752669	0.776217	0.969663	递减	湖北省	1	1	1	不变	陕西省	1	1	1	不变
江苏省	1	1	1	不变	湖南省	0.816273	1	0.816273	递减	四川省	0.850561	0.854911	0.994912	递增
辽宁省	0.790067	1	0.790067	递减	吉林省	0.904072	1	0.904072	递减	西藏	0.681096	1	0.681096	递增
山东省	0.85475	0.890765	0.959568	递减	江西省	0.835609	0.841685	0.992782	递增	新疆	0.844017	1	0.844017	递减
上海市	1	1	1	不变	内蒙古	0.893199	0.922188	0.968565	递减	云南省	0.803966	0.864848	0.929604	递减
天津市	0.589244	0.60371	0.976037	递增	山西省	0.695313	0.705315	0.985819	递增	重庆市	0.907225	0.914214	0.992356	递增
浙江省	1	1	1	不变										
东部平均	0.83049	0.8787	0.94956		中部平均	0.879734	0.921975	0.955872		西部平均	0.836747	0.905142	0.924949	
全国平均	0.84839	0.90119	0.94366											

占比 30%，西部地区有 3 个，占比 30%，各地区较为接近，东部地区略占优势。但从地区地方高校整体来看，东部地区综合技术效率平均值为 0.830486，SD = 0.151648，低于全国平均值 0.8484，纯技术效率平均值为 0.878699，SD = 0.156887，也低于全国平均值 0.901189，中部地区综合技术效率平均值为 0.921975，SD = 0.104516，纯技术效率平均值为 0.921975，SD = 0.104516，西部地区综合技术效率平均值为 0.836747，SD = 0.156191，纯技术效率平均值为 0.905142，SD = 0.135849。表明中部地区地方高校科研绩效综合效率和纯技术效率水平都较高，西部地区两个指标都排名第二，而作为整体经济水平都较高的东部地区，其地方高校科研绩效无论是综合效率，还是纯技术效率反而都处于最后，并且标准差都是最大，说明在东部地区地方高校的科研绩效两极分化较为严重。

进一步分析规模收益状态，处于最优规模状态的东部地区有 4 个，占比 36.36%；中部地区有 3 个，占比 30%；西部地区有 3 个，占比 30%；结果同综合技术有效分布相同。处于规模递增状态的东部有 3 个，占比 27.27%；中部有 2 个，占比 20%；而西部则有 4 个，占比达到了 40%，而处于规模递减状态的东部有 4 个，占比 36.36%；中部有 5 个，占比达 50%；西部有 3 个，占比为 30%。这表明西部地方高校科研规模不够，应加大科研投入，以体现规模效应，而中部地区地方高校则需要控制科研投入，提高精细化管理，防止科研资源浪费，合理分配和利用资源，东部省份地方高校正好处于两者之间，需要因省而异，合理取舍。

为了更好区分各省之间地方高校科研投入产出绩效，特别是对 10 个综合效率值有效的省份进行更精确的区分和排序，本研究将引入两个虚拟最优、最劣决策单元，最优虚拟决策单元计为 DMU₃₂，其输入值选取原 31 个决策单元输入量的最小值，输出值选取原 31 个决策单元的输出量最大值，即

$$x_{i,32} = \min\{x_{i,1}, x_{i,2} \dots x_{i,31}\} \quad (i = 1, 2)$$

$$y_{r,32} = \max\{y_{r,1}, y_{r,2} \dots y_{r,31}\} \quad (r = 1, 2, 3, 4)$$

最劣虚拟决策单元计为 DMU₃₃，其输入值选取原 31 个决策单元输入量的最大值，输出值选取原 31 个决策单元的输出量最小值，即

$$x_{i,33} = \max\{x_{i,1}, x_{i,2} \dots x_{i,31}\} \quad (i = 1, 2)$$

$$y_{r,33} = \min\{y_{r,1}, y_{r,2} \dots y_{r,31}\} \quad (r = 1, 2, 3, 4)$$

构建成 33 个决策单元，建立以最优虚拟决策单元为效率值最大且最劣虚拟决策单元效率值最小的目标决策模型。引入最优、最劣两个虚拟变量势必会对原有 31 个决策单元排序和计算效率值产生干扰，为了既能对原 10 个最优决策单元进行排序，又能减少两虚拟变量对其余 21 个决策单元的干扰，本研究采用虚拟决策单元排序最优决策单元加现实变量排序非有效决策单元的组合排序法，即先在由虚拟决策单元参与的模型中，对 10 个最优决策单元进行排序，确定前 10，再加上原 31 个决策单元所确定的剩余 21 个非有效决策单元的排序，最终确定各省份地方高校科研绩效排名表，具体见表 2。

表 2 各省份地方高校科研绩效评价排名

东部地区		中部地区		西部地区	
省份	排名	省份	排名	省份	排名
北京市	24	安徽省	22	甘肃省	2
福建省	29	广西	14	贵州省	23
广东省	26	河南省	8	宁夏	6
海南省	1	黑龙江省	10	青海省	31
河北省	25	湖北省	7	陕西省	5
江苏省	9	湖南省	19	四川省	16
辽宁省	21	吉林省	12	西藏	28
山东省	15	江西省	18	新疆	17
上海市	4	内蒙古	13	云南省	20
天津市	30	山西省	27	重庆市	11
浙江省	3				

从表 2 可以看出地方高校科研投入产出绩效并非单纯受本地区高等教育和经济发展水平影响，经济文化发展水平较高的北京、天津分别只有 24 和 30，处于排名的末端，这显然与其在高等教育领域的地位是不符的，但仔细分析统计数据不难发现，在北京、天津等高等教育发达地区往往部委高校，无论在师资、科研力量、科院人员配置，还是科研资源的获取，都具有地方高校无法比拟的优势，这必然会对本地区地方高校的科学研究造成挤压，导致优秀科研人员因个人发展平台、科研资源等原因流向部委高校，科研产量低下，长此以往造成北京、天津等地方高校科研绩效排名落后的局面。而与之相反的一些经济、高等教育发展一般或者落后的省份，如海南、甘肃等省份，区域经济内少有部委高校的挤压，地方政府对这些高校有较大投入，国家也有专项的支持中西部科研计划，加之行之有效的科研引导、扶持管理政策，使得这些省份地方高校科研绩效排名处于前列。

四、结论

运用 DEA 的方法,采集 2013 年高等学校科技统计资料汇编中的相关数据,对我国各地区地方高校科研投入产出绩效进行了评价,可以发现海南、甘肃、浙江、上海、陕西、宁夏、湖北、河南、江苏、黑龙江等 10 个省份处于 DEA 有效状态,科研绩效相对较好。从整体上看,技术绩效要好于规模控制绩效;而分地区看,中部地区地方高校绩效综合效率和纯技术效率水平都较高,东部地区两级分化严重,拉低了东部地方高校的整体科研绩效,西部地区地方高校科研绩效介于两者之间。从各地区地方高校的规模收益状态来看,西部地方高校应加大科研投入,体现规模效应;而中部地区地方高校则需提高科研管理效能,控制科研投入,优化科研资源配置;东部省份地方高校各省份差异较大,需要因省而异,合理取舍。同时对各省份高校科研整体绩效排名可以发现,地方高校科研绩效与区域内的文化、经济、高等教育发展水平无必然联系,在一些部委高校云集的高等教育发达地区,如天津、北京等地,地方高校科研绩效水平反而较低。地方政府的支持力度,高校对本地区高水平科研人才的吸引力,科研政策科学性、有效性才是影响地方高校科研绩效水平的关键因素。

参 考 文 献

[1] 中华人民共和国教育部网站, <http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s7382/201305/152554.html>.

[2] Sombatsompop, Narongrit. Research performance evaluations of Thailand national research universities during 2007-2009. *Information Development*, 2010, 26 (4): 303-313.

[3] 王晓红等:《一种基于 DEA 和多指标综合评价的大学本科科研绩效评价方法》,《中国软科学》2004 年第 8 期。

[4] Garcia-aracil A., et al. Analysis of the evaluation process of the research performance: An empirical case. *Scientometrics*, 2006, 67(2): 213-230.

[5] Kao, Chiang., Pao, HweiLan. An evaluation of research performance in management of 168 Taiwan universities. *Scientometrics*, 2009, 78 (2): 261-277.

[6] 戚湧等:《一种基于 DEA 的高校科研绩效评价方法》,《科学与科学技术管理》2008 年第 12 期。

[7] 史兴伟:《利用文献计量学指标对高校科研量化评价研究》,《情报科学》2005 年第 11 期。

[8] 郭银清:《基于 AHP 的高校科研效益评价体系的构建》,《财政研究》2014 年第 2 期。

[9] 段永瑞、霍佳震:《基于数据包络分析的高校科研绩效评价》,《上海交通大学学报》2007 年第 7 期。

[10] Banker, R. D., Charnes, A., Cooper, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 1984, 30(9): 1078-1092.

[11] 魏权龄:《评价相对有效性的数据包络分析模型》,中国人民大学出版社 2012 年版。

[12] Cooper, W. W., et al. Sensitivity and Stability Analysis in DEA: Some Recent Developments. *Journal of Productivity Analysis*, 2001, 15(3): 217-246.

An Evaluation on the Performance of Scientific Researches in Local Colleges and Universities Based on Data Envelopment Analysis

Shen Lihong, Zhao Yi

With the help of data envelopment analysis, the study evaluates and ranks the input and output performance of scientific researches of local colleges and universities of 31 provinces or municipalities in China. The result shows that there are differences in scientific research performance among those local college and universities, and that there is still plenty of room for improvement in the control of the scale of scientific research performance compared to the control of only technology. There is no necessary correlation between the scientific research performance and the development of culture, economy, and higher education in the region. However, the support from the local government, together with the attraction to scientific research talents, and the reasonable policy are the key factors influencing the scientific research performance in local colleges and universities.