

我国政府科技资源投入产出的典型相关分析

沈 赤¹ 章 丹² 李生校³

(1. 绍兴文理学院经济研究院, 浙江绍兴 312000; 2. 浙江工商大学工商管理学院, 浙江杭州 310018;
3. 绍兴文理学院经管学院, 浙江绍兴 312000)

摘 要: 研究采用典型相关分析方法对 1998-2009 年我国政府科技资源投入产出相关性进行了分析。研究结果表明, 科技经费支出额与我国政府科技资源投入状况相关性最强, 拥有发明专利数与我国政府科技资源产出状况相关性最强。与其他投入产出分变量间相关性相比, 科技经费支出额与拥有发明专利数之间的相关性最强。

关键词: 典型相关分析; 我国政府; 科技资源投入产出

中图分类号: F062.4

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2012.03.005

The Canonical Correlation Analysis of Input and Output of the Government's Science and Technology Resource

Shen Chi¹, Zhang Dan², Li Shengxiao³

(1. School of Economic Research Institute, Shaoxing University, Shaoxing 312000;
2. School of Business and Administration, Zhejiang Gongshang University, Hangzhou 310018;
3. School of Economic and Business, Shaoxing University, Shaoxing 312000)

Abstract: Canonical correlation analysis method is used to analyze the relationship between input and output of the government science and technology resources in 1998-2009 in this paper. The results show that: the impact of expenditures for S&T on the status of government investment in science and technology resources is greatest and the impact of the number of invention patents on the output of scientific and technological resources of state government is greatest. Compared with other pairs of variables, the expenditures for S&T and the number of invention patents are more relevant.

Keywords: canonical correlation analysis, government, input and output of the science and technology resource

1 引言

近年来, 我国政府期望通过不断增加科技资源投入提高科技能力, 从而增加国家自主创新能力。但是如何把握科技资源投入的方向已经成为我国政府所面临的一个重大现实问题^[1]。

现有针对科技资源投入产出问题的研究主要集中于识别投入过剩和产出不足的指标, 同时还包括对应的数量分析^[2-5]。但是缺乏对投入指标与产出

指标之间相关关系的整体分析。具体而言, 包括分析科技资源投入和产出指标中的关键指标, 以及能够对产出指标产生关键影响的投入指标。

针对以上问题, 本研究在确定我国政府科技资源投入和产出评价指标的基础上, 采用典型相关分析方法对 1998-2009 年我国政府科技资源投入产出数据进行分析^[6]。所谓典型相关分析, 是利用综合变量之间的相关关系来反映两组指标之间的整体相关性的多元统计分析方法。其基本原理是: 为了从

第一作者简介: 沈赤 (1964-), 男, 绍兴文理学院经济研究院院长, 副研究员, 研究方向: 科技管理。

基金项目: 国家软科学研究计划“政府优化科技资源配置的评价指标体系构建研究”(2010GXQ5D334)。

收稿日期: 2011 年 10 月 14 日。

总体上把握两组指标之间的相关关系,分别在两组变量中提取有代表性的两个综合变量 U_1 和 V_1 (分别为两个变量组中各变量的线性组合),利用这两个综合变量之间的相关关系来反映两组指标之间的整体相关性。这将丰富和发展科技资源投入产出理论,同时对于解释我国目前的科技资源投入政策具有重要意义。

2 投入产出评价指标体系

选择评价指标时要遵守以下几项基本原则:系统性原则、有效性原则、可操作性原则。系统性原则是指所建立的评价指标体系能够涵盖反映的科技资源投入产出的基本特征和整体状态;有效性原则是指所建立的评价指标体系符合进行评价对象的结构与状况,能合理地反映科技资源投入产出的结构特征;可操作性原则是指评价指标体系的设计在遵循系统性原则的基础上,保证数据资料的可获得和可量化,并且评价指标尽可能简化,不宜过多。

本文参考2001年联合国教科文组织统计部门分析报告中的投入与产出指标,依据选择评价指标所要遵循的可操作性原则,结合现有理论研究成果分析,筛选确定如下我国政府科技资源投入产出指标。

2.1 科技投入指标

(1)科技经费支出占GDP的比重(X_1)。通过查阅中国科技统计年鉴可得到1995–2009年的数据,通过查阅中国统计年鉴可得到1992、1993和1994年这3年的数据。

(2)万人口科技活动人员(X_2)。通过查阅中国科技统计年鉴,采用计算方法:科技活动人员/全国人口(万人),可得到1992–2008年的数据,但是2010年不再统计科技活动人员,因此,2009年的数据不可获取。但通过指数平滑法可以预测出2009年的科技活动人员的数据,然后再经过计算公式得到2009年的数据。

(3)科技活动人员(X_3)。通过查阅2009年中国科技统计年鉴可得到1992–2008年的数据,但是2010年不再统计科技活动人员,因此,2009年的数据不可获取,但是可以通过指数平滑法预测出2009年的数据。

(4)技术开发机构数。通过查阅中国科技统计年鉴和中国统计年鉴,发现与之最为对应的指标是中国统计年鉴中的指标:科技机构数,但是只能获

取1992–2000年的数据,同时,该指标能够被科技活动人员所取代,因此,该指标可以被删除。

(5)科技经费支出额(X_4)。通过查阅中国科技统计年鉴和中国统计年鉴,发现与之最为对应的指标是R&D经费内部支出,可获得1992–2009年的数据。

(6)高层次科技人才数(X_5)。通过查阅2009年中国科技统计年鉴,发现与之最为对应的指标是:科学家和工程师。可获得1992–2008年的数据,但是,2010年科技统计年鉴不再统计该指标,因此,2009年数据不可获取,但是可以通过指数平滑法预测出2009年的数据。

2.2 科技产出指标

(1)拥有发明专利数(Y_1)。通过查阅中国科技统计年鉴,可获得1993–2009年的数据。

(2)新产品出口销售收入(Y_2)。通过查阅中国科技统计年鉴,可获得1996–2009年的数据。

(3)高技术产业增加值(Y_3)。通过查阅2007、2008年高技术产业统计年鉴可得到1995–2008年的数据,2009年和2010年的高技术产业统计年鉴不再统计该指标数据,但是2008年的数据可以通过官方公布数据推出高技术产业增加值占全部工业增加值的10.2%。对于2009年的数据可以通过国家发改委网站公布数据获取。

(4)国内中文期刊科技论文数(Y_4)。通过查阅中国科技统计年鉴可得到1998–2008年的数据,2010年中国科技统计年鉴统计数据为2008年数据,因此,2009年数据不可获取,但是可以通过指数平滑法预测出2009年的数据。

综合以上数据的可获取性,只能选择1998–2009年的数据作为原始数据输入分析我国政府科技资源投入产出相关性。因此,本文共选取2个方面9项指标建立我国政府科技资源投入产出评价指标体系(图1)。

3 相关系数矩阵

我国政府科技资源投入指标间的相关系数如表1所示,产出指标间的相关系数如表2所示,投入产出指标间的相关系数如表3所示。从表1可以看出我国政府科技资源投入指标间的相关系数最小为0.9289,最大为0.9997,因此,投入指标间存在较大的相关性。从表2可以看出产出指标间的相关系数最小为0.9248,最大为0.9915,因此,产

出指标间存在较大的相关性。从表 3 可以看出投入产出指标间的相关系数最小为 0.9163，最大为 0.9884，因此，投入产出指标间存在较大的相关性。表 1、表 2 和表 3 表明，采用典型相关分析方法分析我国政府科技资源投入产出各指标间的关系是合适的。

4 典型相关系数、检验及其典型相关模型

典型相关系数的输出结果见表 4。可以看出第一典型相关系数的显著性概率为 0.000，第二、三和四典型相关系数的显著性概率分别为 0.196、0.383

和 0.515，因此，在 0.05 的显著性水平下，四对典型变量中只有第一对典型相关是显著的。因此，投入和产出相关性的研究可以转化为研究第一对典型变量相关变量之间的关系。我们采用第一对典型变量 U_1 、 V_1 来建立典型相关模型。其模型如下：

$$U_1 = -0.368X_1 + 2.207X_2 - 2.284X_3 - 0.816X_4 + 0.26X_5$$

$$V_1 = -0.607Y_1 + 0.165Y_2 - 0.242Y_3 - 0.322Y_4$$

由于典型权重只是保证了这一对典型变量间的相关系数在所有线性组合中是最大的，并且我国政府科技资源投入和产出各个变量之间相关性较大，可能存在共线性等问题，因此，利用典型权重来解

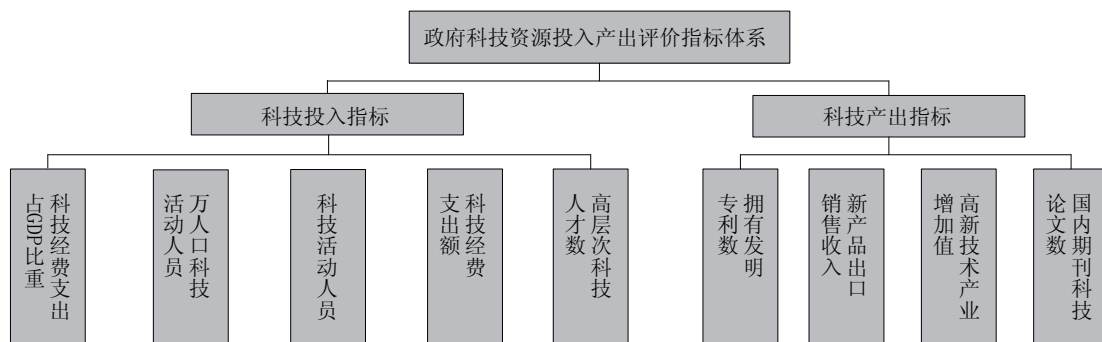


图 1 投入产出评价指标体系

表 1 投入指标间的相关系数

	X_1	X_2	X_3	X_4	X_5
X_1	1.0000	0.9289	0.9377	0.9403	0.9675
X_2	0.9289	1.0000	0.9997	0.9930	0.9717
X_3	0.9377	0.9997	1.0000	0.9943	0.9757
X_4	0.9403	0.9930	0.9943	1.0000	0.9615
X_5	0.9675	0.9717	0.9757	0.9615	1.0000

表 2 产出指标间的相关系数

	Y_1	Y_2	Y_3	Y_4
Y_1	1.0000	0.9248	0.9633	0.9483
Y_2	0.9248	1.0000	0.9793	0.9615
Y_3	0.9633	0.9793	1.0000	0.9915
Y_4	0.9483	0.9615	0.9915	1.0000

表 3 投入和产出指标间的相关系数

产出指标 \ 投入指标	Y_1	Y_2	Y_3	Y_4
X_1	0.9468	0.9163	0.9641	0.9821
X_2	0.9680	0.9641	0.9800	0.9560
X_3	0.9714	0.9655	0.9835	0.9622
X_4	0.9884	0.9537	0.9809	0.9605
X_5	0.9395	0.9602	0.9763	0.9762

释各个变量的相对重要程度是不妥的。现有研究表明，采用典型载荷和交叉载荷能够较好地反映各个变量的相对重要程度。因此，必须进一步进行典型结构分析。

5 典型结构

典型结构分析依据原始变量与典型变量之间的相关系数，反映原始变量与典型变量（包括自身典型变量和对方典型变量）之间的相关程度。当典型载荷的绝对值越大，相关性越大，对典型变量解释时，其重要性也越高^[7]。

由表 5 和表 6 可以看出，我国政府科技资源投入原始变量与第一对典型变量（ U_1 和 V_1 ）的相关程度都较高，而与其他典型变量的相关程度都较低。由表 7 和表 8 可以看出，我国政府科技资源产出原始变量与第一对典型变量（ U_1 和 V_1 ）的相关程度都较高，而与其他典型变量的相关程度都较低。进一步，可以分析 U_1 和 V_1 的含义： U_1 与 5 个投入变量之间的相关系数最小为 0.963，最大为 0.990（表 5），

所以 U_1 完全能够反映我国政府科技资源投入能力，同时与科技经费支出额（ X_4 ）之间的相关性最强， U_1 与 4 个产出变量之间的相关系数最小为 0.942，最大为 0.992（表 8），所以 U_1 与我国政府科技资源产出能力相关性较大，同时与拥有发明专利数（ Y_1 ）之间的相关性最强； V_1 与 4 个产出变量之间的相关系数最小为 0.943，最大为 0.993（表 7），所以 V_1 完全能够反映我国政府科技资源产出能力，同时与拥有发明专利数（ Y_1 ）之间的相关性最强， V_1 与 5 个投入变量之间的相关系数最小为 0.962，最大为 0.989（表 6），所以 V_1 与我国政府科技资源投入能力相关性较大，同时与科技经费支出额（ X_4 ）之间的相关性最强。

由以上分析可知，与其他因素相比，我国政府科技资源投入与科技经费支出之间存在最大的相关性，我国政府科技资源产出与拥有发明专利之间存在最大的相关性。因此，科技经费支出额是影响我国政府科技资源产出的最重要因素。

表 4 典型相关系数及其相关检验

	典型相关系数	Wilks的统计量	卡方统计量	自由度	显著性概率
1	0.999	0.000	53.034	20.000	0.000
2	0.892	0.071	15.894	12.000	0.196
3	0.754	0.346	6.371	6.000	0.383
4	0.445	0.802	1.325	2.000	0.515

表 5 投入原始变量与自身典型变量之间的典型载荷

自身典型变量 原始变量	U_1	U_2	U_3	U_4
X_1	-0.974	0.209	-0.044	0.037
X_2	-0.974	-0.141	-0.171	-0.017
X_3	-0.979	-0.118	-0.162	-0.018
X_4	-0.990	-0.122	-0.062	-0.035
X_5	-0.963	0.074	-0.226	-0.076

表 6 原始变量与对方典型变量之间的交叉载荷

对方典型变量 原始变量	V_1	V_2	V_3	V_4
X_1	-0.973	0.187	-0.033	0.017
X_2	-0.973	-0.126	-0.129	-0.007
X_3	-0.978	-0.105	-0.122	-0.008
X_4	-0.989	-0.109	-0.046	-0.016
X_5	-0.962	0.066	-0.170	-0.034

表 7 产出原始变量与自身典型变量之间的典型载荷

自身典型变量 原始变量	V_1	V_2	V_3	V_4
Y_1	-0.993	-0.084	0.069	-0.054
Y_2	-0.943	-0.049	-0.308	-0.119
Y_3	-0.984	0.002	-0.176	0.021
Y_4	-0.979	0.131	-0.157	0.024

表 8 产出原始变量与对方典型变量之间的交叉载荷

对方典型变量 原始变量	U_1	U_2	U_3	U_4
Y_1	-0.992	-0.075	0.052	-0.024
Y_2	-0.942	-0.044	-0.232	-0.053
Y_3	-0.983	0.002	-0.133	0.009
Y_4	-0.978	0.117	-0.118	0.011

表 9 原始变量的变异可被典型变量所解释的比例

	被自身典型变量 解释的百分比	被对方典型变量 解释的百分比		被自身典型变量 解释的百分比	被对方典型变量 解释的百分比
U_1	0.953	0.951	V_1	0.950	0.948
U_2	0.020	0.016	V_2	0.007	0.005
U_3	0.022	0.013	V_3	0.039	0.022
U_4	0.002	0.000	V_4	0.005	0.001

6 典型冗余分析

通过以上分析,典型相关系数描述了典型变量之间的相关程度,而典型载荷和交叉载荷描述了典型变量与每个原始变量之间的相关关系,但有时需要将每组原始变量作为一个整体,考察典型变量与变量组之间的相关程度,从而分析这些典型变量对两组变量的解释能力,以正确评价典型相关的意义,此时,需要进行冗余分析^[7]。

冗余分析的结果说明了典型变量对各变量组方差解释的比例(表9)。我国政府科技资源投入原始变量的变异可被自身的典型变量(U_1)所解释的比例为95.3%,表明典型变量 U_1 完全能够代表我国政府科技资源投入原始变量,我国政府科技资源投入原始变量的变异可被对方典型变量(V_1)所解释的比例为95.1%,表明典型变量 V_1 能够对我国政府科技资源投入产生重要影响;我国政府科技资源产出原始变量的变异可被自身的典型变量(V_1)所解释的比例为95%,表明典型变量 V_1 完全能够代表我国政府科技资源产出原始变量,我国政府科技资源产出原始变量的变异可被对方典型变量(U_1)所解

释的比例为94.8%,表明典型变量 U_1 能够对我国政府科技资源产出产生重要影响。

7 结论

本文所用的数据历时12年(1998—2009),属于小样本,少数据,同时涉及2组指标变量,对于这类数据进行回归分析结果很容易产生偏差。因此,将投入产出指标变量的实际调查数据输入计算机,按统计软件包PASW18.0中附带的命令程序进行典型相关分析,以判断影响我国政府科技资源产出的主要投入变量。研究得到如下主要结论。

(1)分析我国政府科技资源投入和产出评价指标及之间相关性,科技经费支出占GDP比重、万人口科技活动人员、科技活动人员、科技经费支出额和高层次科技人才数等指标能解释政府科技资源投入状况变异的95.3%,拥有发明专利数、新产品出口销售收入、高技术产业增加值和国内中文期刊科技论文数等指标能解释政府科技资源产出状况变异的95.1%,投入组指标能解释产出组指标变异的94.8%。

(下转第31页)

- [3] Schumann P A Jr, Ransley D L, Prestwood D C L. Measuring R&D Performance[J]. Research Technology Management, 1995(5/6):45-54.
- [4] Erik arnold, Katalin Balázs. Methods in the Evaluation of Publicly Funded Basic Research: A Review For OECD[EB/OL].[1998-03-01]. http://www.technopolis-group.com/resourcesdownloads/reports/022_eval_bas.pdf.
- [5] Brown M G, Sveson R A. Measuring R&D Productivity[J]. Research Technology Management, 1988(7/8):11-15.
- [6] Martin B R. The Use of Multiple Indicators in the Assessment of Basic Research[J]. Scientometrics, 1996,36(3):343-362.
- [7] 吴俊卿, 郑慕琦, 张志兴. 绩效评价的理论与方法: 在科研机构的实践[M]. 北京: 科学技术文献出版社, 1992.
- [8] 李正风. 基础研究绩效评估的若干问题[J]. 科学学研究, 2002(1):67-71.
- [9] 周寄中. 创新的基础和源泉: 基础研究的投入、评估和协调[M]. 北京: 科学出版社, 2008.
- [10] 戚湧, 李千目. 科学研究绩效评价的理论与方法[M]. 北京: 科学出版社, 2009.
- [11] 《全国科技进步综合评价》课题组. 全国科技进步统计监测及综合评价研究[J]. 统计研究, 1998(2):26-35.
- [12] 中华人民共和国科学技术部. 中国科学技术指标—2006[M]. 北京: 科技文献出版社, 2007.
- [13] 袁卫, 赵路, 钟卫, 等. 中国 R&D 理论、方法及应用研究[M]. 北京: 中国人民大学出版社, 2009.
- [14] 梁立明, 梁立华, 孔繁士, 等. 高等院校科研绩效评估的自报指标与源生指标[J]. 科研管理, 1998(6):44-48.
- [15] 王立平. 我国高校 R&D 知识溢出的实证研究——以高技术产业为例[J]. 中国软科学, 2005(12):54-59.
- [16] 晋彤驹. “投入产出法”的由来和发展[J]. 经济问题, 1981(2):40-43.
- [17] 阿尔弗雷德·马歇尔. 经济学原理[M]. 周月刚, 雷晓燕, 译. 北京: 中国城市出版社, 2010.
- [18] 亚瑟·赛斯尔·庇古. 福利经济学[M]. 朱泱, 张胜纪, 吴良健, 译. 北京: 商务印书馆, 2006.
- [19] 罗纳德·科斯. 社会成本问题[M]//盛洪. 现代制度经济学(上卷). 北京: 北京大学出版社, 2003:3-37.
- [20] 姜鑫, 罗佳. 高校 R&D 资源清查的问题及对策[J]. 中国高校科技与产业化, 2011(5):16-18.

(上接第 25 页)

(2) 对我国政府科技资源投入和产出指标中的关键指标进行识别, 并对其间的相关性进行分析。典型相关分析结果表明, 政府科技资源投入状况与科技经费支出之间相关性最大, 政府科技资源产出状况与拥有发明专利数之间相关性最大, 同时由于本文所选择的第一对典型变量 (U_1 和 V_1) 之间典型相关系数最大, 说明与其他典型变量相比, U_1 所代表的科技经费支出额和 V_1 所代表的拥有发明专利数之间存在最大相关性。

(3) 从研究可以看出, 我国政府科技资源投入组指标能够解释产出组指标的大部分变异。我国政府科技资源投入指标中的关键指标是科技经费支出额, 而产出指标中的关键指标是拥有发明专利数。与其他投入产出分变量间相关性相比, 科技经费支出额与拥有发明专利数之间相关性最强。因此, 为了保持我国政府科技资源高投入产出效率, 重点在于合理控制科技经费支出额, 并持续监控拥有发明专利数。本研究对我国政府科技资源投入产出相关性进行了研究, 得到了一些有意义的结论, 但是对于我国不同省份的科技资源投入和产出各指标之间

的定量关系的分析还有待进一步深入研究。

参考文献

- [1] 吴和成, 华海岭, 杨勇松. 制造业 R&D 效率测度及对策研究——基于中国 17 个制造行业的数据[J]. 科研管理, 2010,31(5):45-53.
- [2] Chandler G, Hanks S. An Examination of the Substitutability of Founder's Human and Financial Capital in Emerging Business Ventures[J]. Journal of Business Venturing, 1998(13):353-369.
- [3] 迟国泰, 隋聪, 齐菲. 基于超效率 DEA 的科学技术评价模型及其实证[J]. 科研管理, 2010,31(2):94-103.
- [4] 石风光, 周明. 中国地区技术效率的测算及随机收敛性检验——基于超效率 DEA 的方法[J]. 研究与发展管理, 2011,23(1):23-30.
- [5] 管燕, 吴和成, 黄舜. 基于改进 DEA 的江苏省科技资源配置效率研究[J]. 科研管理, 2011,32(2):145-150.
- [6] 孙玮, 王九云, 成力为. 技术来源与高新技术产业创新生产率——基于典型相关分析的中国数据实证研究[J]. 科学学研究, 2010,28(7):1088-1093.
- [7] 汪冬华. 多元统计分析与 SPSS 应用[M]. 上海: 华东理工大学出版社, 2010.