

基于因子分析的湖南省区域创新环境研究

黄玲雁

(湖南工业大学商学院,湖南 株洲 412007)

[摘要] 文章以湖南省14个地市作为研究对象,并且构建了湖南省区域创新环境评价指标体系,然后通过因子分析的方法获得了湖南省区域创新环境的综合指标,以此对湖南省14个地市的创新环境进行了评价与分析,最终揭示了湖南省区域创新环境的差异和特征。

[关键词] 区域创新环境; 因子分析

[中图分类号] F127.53 **[文献标识码]** A **[文章编号]** 1673-0755(2014)01-0059-05

引言

区域创新环境是可持续发展的动力,对产业和区域整体具有持续推动作用^[1]。湖南省地处我国中部地区,是国家实施中部崛起战略以及联结东、西部地区的重要桥梁和纽带。2010年,全省实现生产总值16037.96亿元,比上年增长18.6%^①。本文以湖南省14个城市为研究对象,使用因子分析方法,定量对湖南区域创新环境进行分析,以期对湖南社会的创新发展具有一定的参考价值。

一 区域创新环境研究概述

区域创新环境的概念最早是由欧洲创新环境研究小组(GREMI)于1985年提出来的。当前对于区域创新环境的认识还未获得一致的认同。不过总体来说,对于区域创新环境的研究包括以下几个方面。其一,区域创新环境的内涵及构建。Martina Fromhold-Eisebith以德国亚琛(Aachen)地区作为研究对象,研究了创新环境与社会资本的差异性和一致性,通过分析发现,差异性和一致性相互作用,一起为区域创新系统的发展和重构提供良好的环境^[2]。陈伟从区域创新系统理论的角度出发,着重归纳了区域创新系统的内涵、结构和差异化的分类,并且在此研究上提出了以城市群作为单位的区域创新演进方向^[3]。王灏晨通过对广西省的区域创新系统进行分析,设计出了广西省区域创新系统模型,并且运用新系统动力学模型对动态仿真结果进行分析,发现创新环境中的人才因素和资金因素是制约广西区域创新系统建设的瓶颈要素^[4]。其二,区域创新环境

因素和区域经济发展关系。盖文启指出区域创新环境的建设对区域经济的发展有着重大的影响,尤其是区域内日益创新的软环境扮演着十分关键的角色。他还结合了实际情况指出在区域创新环境方面我国高新技术产业区发展过程中所存在的不足^[5]。王树林(2002)认为区域创新环境对于区域经济竞争力的提升有巨大的作用,是带动区域经济发展的火车头。区域创新环境可以作用于区域内的创新资源,使其得到整合和优化,从而提高区域内的创新能力,实现经济的快速增长^[6]。其三,创新环境如何影响创新绩效。李习保根据区域创新系统理论,结合知识生产函数模型,采用随机前沿模型,利用我国30个省份1998—2005年的创新活动数据,实证分析了创新环境对于创新活动效率的影响,研究指出:科技研发经费支出和政府的财政支持对创新效率有着明显的促进作用^[7]。岳鹄通过分析省级创新产出1997-2006年的面板数据,实证检验区域创新环境对创新产出效率的影响,结果显示:中国30个省市创新能力的显著差异,不但是因为各地区RD投入存在差异,创新环境存在着的差异也发挥着巨大的作用,这进一步说明了区域创新环境对区域创新能力确实具有重要影响^[8]。张莹以《中国区域创新能力报告》的数据为基础构建了panel-data模型,着重选择了5个分量实证分析了重庆市的区域创新环境对于创新绩效产生怎样的影响,通过回归分析,揭示了重庆市在区域创新环境投入方面存在的一系列问题,并且提出了解决方法^[9]。许婷婷研究了江苏省区域创新环境的现状以及创新产出的变化趋势,

分析了区域创新环境的差异,并且给出了相应的建议,为今后相关的研究提供了借鉴^[10]。

二 湖南省区域创新环境评价

(一)评价指标体系的建立

参考《中国区域创新能力报告》一书,结合已有的指标体系,考虑到数据的可得性后(因为有些数据指标在《湖南省统计年鉴》中并没有被统计),本文构建了用于描述湖南省区域创新环境的指标,一共 11 个,分别是高新技术产业总产值(S1)、大中型工业企业当年科技活动经费支出总额(S2)、外商直接投资中实际利用投资(S3)、企业办科研机构(S4)、公路长度(S5)、图书馆个数(S6)、公路货运量(S7)、地方教育支出(S8)、金融机构人民币贷款(S9)、大中型企业科技活动人员(S10)、新产品产值(S11)。这些指标基本上也包括了基础设施环境、市场环境、科技金融环境等。

(二)数据来源与相关性的检验

本文以湖南省 14 个地市为数据样本,数据来源于 2010 年《湖南省统计年鉴》,具有较高的可信度。因子分析是以相关性为基础的,所以有些数据是不适用的,因此在做因子分析前有必要对样本数据进行一些必要的检验。本文使用 SPSS 软件对指标数

据进行相关分析,经过检验,各指标间的相关度可以满足因子分析要求。

(三)创新环境因子分析

Thurstone 在 1931 年首次提出了因子分析,其概念起源于 20 世纪初有关于智力测验的统计分析。因子分析是用少量的因子来说明多个变量之间的关系,被描述的变量通常是能够实际观测的随机变量,而那些因子则是难以观测的潜在变量^[11]。简单的来说因子分析就是把紧密联系的变量归到一起,而剩下的则属于不同类别,它们的相关性则比较低。通过降维把具有较高相关度的变量聚在一起,这便于解释变量的特征,从而使问题分析过程得到简化。因子分析以相关性为前提,从协方差矩阵或相关矩阵入手,认为大部分变异是由于某少数几个公共因子导致,而把余下的变异称为特殊因子。因子分析的基本步骤是首先建立因子模型;然后要对载荷矩阵进行估计,此时建立的因子模型还是一个初始模型;再采用方差最大化正交旋转方法对因子进行旋转;最后利用因子得分综合矩阵计算出因子综合得分。

具体操作步骤如下:

首先运用 SPSS 软件对数据进行处理,可以得到描述统计量,见表 1。

表 1 描述统计量

	均值	标准差	分析 N
公路货运量 万吨	9688.0714	5728.19047	14
大中型企业科技活动人员 人	5953.5000	7386.75220	14
高新技术产业总产值 万元	4.5979E6	5.81940E6	14
金融机构人民币贷款 亿	800.0464	1555.01357	14
大中型工业企业当年科技活动经费支出总额 万元	127668.2857	1.73179E5	14
外商直接投资实际利用投资 万美元	37031.5000	56394.78536	14
企业办科研机构 个	33.6429	37.84911	14
公路长度 公里	16285.5600	4776.01763	14
图书馆个数 个	8.8571	3.10972	14
新产品产值 万元	1.7324E6	2.34209E6	14
地方教育支出 万元	249618.7857	1.10571E5	14

其次进行 KMO 检验,得到的结果见表 2。

表 2 KMO 和 Bartlett 的检验

取样足够度的 Kaiser-Meyer-Olkin 度量	.771
近似卡方	219.127
Bartlett 的球形度检验	df 55
	Sig. .000

KMO 检验主要是用来研究变量之间的偏相关性,其值处于 0~1 之间,越接近 1,则表明变量之间的相关性越强,做因子分析的效果也就越好。经检

验,KMO 的统计值为 0.771,大于 0.7,因此适合进行因子分析。

再利用 SPSS 软件分别求得相关系数矩阵、特征值和累积贡献率,见表 3。根据特征值大于 1 的原则,可以判断选取 2 个主因子比较合适,其旋转后的累积贡献率达到 88.882%,已经大于 85%,能够解释指标体系中变量的大部分方差。通常选取累计方差贡献率大于 85%,就能够保留重要的信息。所以提取 2 个公共因子,分别命名为 L1、L2。

表 3 解释的总方差

成份	初始特征值			提取平方和载入			旋转平方和载入		
	合计	方差的 %	累积 %	合计	方差的 %	累积 %	合计	方差的 %	累积 %
1	7.696	69.961	69.961	7.696	69.961	69.961	7.229	65.718	65.718
2	2.081	18.921	88.882	2.081	18.921	88.882	2.548	23.164	88.882
3	.459	4.174	93.056						
4	.411	3.736	96.792						
5	.200	1.814	98.606						
6	.067	.606	99.212						
7	.052	.470	99.682						
8	.014	.129	99.811						
9	.011	.103	99.914						
10	.007	.060	99.974						
11	.003	.026	100.000						

由于因子模型存在着不唯一性,所以能够因子旋转。设 T 为一个正交矩阵,由于 $TT' = I$,所以因子模型 $\vec{x} - \vec{u} = A \vec{F} + \vec{\varepsilon}$ 等价于 $\vec{x} - \vec{u} = (AT)(T'\vec{F}) + \vec{\varepsilon}$, 而后的载荷矩阵为 $B = AT, \vec{G} = T'\vec{F}$ 是公共因子。如果模型不易于解释,那么作一个正交变换 T,把模型变为 $\vec{x} - \vec{u} = (AT)(T'\vec{F}) + \vec{\varepsilon} = B \vec{G} + \vec{\varepsilon}$,然后再在新模型中寻找因子的合理解释。现实中由于初始得到的因子不一定能够反映问题的实质特征,它们所代表的实际意义不一定容易解释,此时的初始载荷矩阵和评价指标的结构关系模糊不易辨别,因此就有必要采用 Varimax 方差最大旋转法,通过正交旋转,它能够使得各个因子上所具有较高载荷的变量数目达到最小,由此能够简化对因子的解释。通过旋转可以发现,旋转之后每个公因子上的载荷分配显得更加的清晰,因此这比没有旋转时更容易解释各因子的意义,见表 4。

因子载荷作为公共因子与变量的相关系数,其绝对值越大,则表示该因子与变量的关系更为紧密,因而更能代表这个变量。通过表 4 观测可以看出, L1 在金融结构人民币贷款、高新技术产业总产值、外商直接投资实际利用投资、大中型企业科技活动人员、大中型工业企业当年科技活动经费支出总额、新产品产值和企业办科研机构上具有较大的载荷,本文将将其命名为科技金融动力因子。L2 在图书馆个数、公路长度、公路货运量和地方教育支出上有较大的载荷,本文将将其命名为社会基础因子。

在所建立的因子模型里,根据因子分析得到的结果,总体里的原有变量已经被分解为公共因子和特殊因子的线性组合 $X_i = a_{i1}F_1 + a_{i2}F_2 + a_{im}F_m + \varepsilon_i, i = 1, 2 \cdots p$;据此可以用原有变量的线性组合 $F_j = b_{j1}X_1 + b_{j2}X_2 + b_{jp}X_p, j = 1, 2 \cdots m$ 来表示公共因子,称之为因子

得分系数,用它能够计算每个观测记录在各个公共因子上的得分,把公共因子不可测量转换为可以测量。由 SPSS 可以得到因子系数矩阵,见表 5:

表 4 旋转成分矩阵

		成份	
		L1	L2
高新技术产业总产值	万元	.989	.107
大中型科技活动人员	人	.980	.026
大中型工业企业当年科技活动经费支出总额	万元	.979	-.006
企业办科研机构	个	.957	.186
金融机构人民币贷款	亿	.933	.160
外商直接投资实际利用投资	万美元	.932	.226
新产品产值	万元	.922	.122
公路长度	公里	-.184	.884
图书馆个数	个	.136	.875
地方教育支出	万元	.672	.702
公路货运量	万吨	.569	.608

根据因子得分系数可以得到因子得分计算公式:

L1 得分 = $0.032 \times \text{公路货运量} + 0.153 \times \text{大中型企业科技活动人员} + \cdots + 0.039 \times \text{教育支出}$;

L2 得分 = $0.219 \times \text{公路货运量} - 0.083 \times \text{大中型企业科技活动人员} + \cdots + 0.252 \times \text{教育支出}$ 。

要获得各地区的综合实力,可以对 2 个因子的得分进行加权求和,权数为方差贡献率,表 3 已经计算得出了 L1 的为 65.72%,L2 的为 23.16%。在 SPSS 中可以直接得到第 1 列因子得分 FAC1_1 和第 2 列因子得分 FAC2_1。据此可以进一步计算出湖南省 14 个地市的创新环境综合评分,综合得分的计算公式如下:

综合得分 = $65.72\% \times \text{FAC1_1} + 23.16\% \times \text{FAC2_1}$ 。

表 5 成分得分系数矩阵

		成份	
		L1	L2
公路货运量	万吨	.032	.219
大中型企业科技活动人员	人	.153	-.083
高新技术产业总产值	万元	.147	-.047
金融机构人民币贷款	亿	.133	-.018
大中型工业企业当年科技活动经费支出总额	万元	.156	-.097
外商直接投资实际利用投资	万美元	.126	.012
企业办科研机构	个	.134	-.008
公路长度	公里	-.115	.417
图书馆个数	个	-.063	.382
新产品产值	万元	.135	-.034
教育支出	万元	.039	.252

计算得到的结果见表 6。

表 6 湖南省 14 个地市创新环境综合得分

城市	L1	L2	区域创新环境	
	得分	得分	综合得分	排名
长沙	3.16922	0.50152	219.89	1
株洲	0.72729	-0.80964	29.04	2
衡阳	-0.03965	0.98216	20.15	3
郴州	-0.34887	0.97469	-0.35	6
岳阳	-0.00024	0.49895	11.54	4
邵阳	-0.62265	1.01853	-17.33	8
湘潭	0.39000	-1.70426	-13.85	7
娄底	-0.27648	-0.07507	-19.91	10
益阳	-0.40796	-0.36819	-35.34	11
永州	-0.71840	1.18989	-19.41	9
常德	0.16844	0.54819	1.63	5
怀化	-0.60798	-0.15503	-43.55	12
张家界	-0.48709	-1.90141	-76.05	14
吉首	-0.61234	-0.70034	-56.46	13

三 湖南省区域创新环境差异

(一)湖南省区域创新环境差异特征

湖南省区域创新环境整体水平从总体上看,许多城市区域创新得分为负,这说明总体的创新环境水平偏低;全省的创新环境差异比较明显,投资环境较好的是湘北以及京广沿线地区。长沙以 219.89 的总分位居第一位,株洲和衡阳以 29.04 和 20.15 分列第二和第三位,岳阳和常德紧跟其后排在第四位和第五位。以上地区由于地理位置较好,交通发达,地区经济发展较为迅速,各项基础设施建设比较完备,金融和商业贸易以及科技发展的水平较高,因此创新环境较好。排名第七的是湘潭,它的科技金

融动力因子得分有 0.39,单一看 L1 这一个因子,它的排名是第四,但是湘潭的社会基础因子得分偏低,仅有 -1.7,致使综合排名下拉。郴州的科技金融动力因子的得分并不是很高,但是它的社会基础因子较高,使得它的综合排名超过了湘潭,排名第六,因此 L2 因子在一定程度上对创新环境的建立与完善具有推动作用。排名处于中下游的是邵阳、永州、娄底和益阳。湘西地区创新环境整体水平较差,怀化、吉首和张家界的排名不理想,无论是科技金融动力因子还是社会基础因子的得分都为负数,属于全省倒数几名,创新环境有待提高。怀化市作为湘西地区的主要经济中心和重要的铁路枢纽,有着相对较好的基础设施,其创新环境要高于其他两城市。

本文的因子分析得到的结果基本上反映了 2010 年湖南省区域创新环境的实际情况。唯一有点意外的是作为湖南省的“两型社会”之一的湘潭,排名到了第七位(将在下文的成因分析中说明)。

(二)湖南省区域创新环境差异成因分析

长沙作为湖南省的省会,是湖南省的政治、经济中心,因子评分排名第一,远远高于排名第二、第三的株洲、衡阳,这充分体现其经济实力和科技发展状况。长沙拥有四个国家级开发区,高新技术企业的快速发展带动了区域企业创新意识的提升;地理条件优越,为招商引资提供了很好的平台;这些都为创新活动注入了活力。

株洲地处湖南省东部,是非常重要的交通枢纽,被誉为“火车拉来的城市”。京广线、湘黔线、浙赣线、沪昆高速铁路和武广高速铁路都在这里交汇。株洲市享有“两型社会”建设综合配套改革试验区、中部崛起战略等 6 个国家级政策。由于有政策上的支持,基础设施的建设也比较完善,经济科技的发展动力充足,创新环境呈现出良好的态势。

衡阳市作为湖南省的一座老工业城市,近些年社会经济发展的速度放缓了,但是在湖南省依旧是重要的交通、商贸、文教发达城市。在社会基础因子上的得分比较高,说明在它公路长度、图书馆个数,公路货运量 and 教育支出等方面拥有一定的优势,这提高了其区域创新环境的综合水平。

岳阳和常德是湖南省的新兴发展城市,近些年社会经济增长相当明显。常德作为湘西北的桥头堡,在国家东西互动的战略中扮演着重要的角色,经济商业比较发达,德山经济开发区获批升级为国家级开发区。岳阳资源丰富,工业生产体系较为完备,区位条件优越,也拥有一个国家级经济技术开发区。两市政府对科学技术和教育方面的投入较多,支持

较大,说明政府在努力地营造较好的区域环境。

郴州作为广东的“后花园”,借助依靠广东的有利区位优势,经济发展也较为迅速。郴州经济开发区在 2003 年被批准为高新技术产业开发区,目前已经形成了机械、电子、新材料、食品、化工、建材、冶炼和医药等八大支柱产业,创新环境良好。

湘潭在湖南省的 14 个地市中面积是最小的,人口排在倒数第三。作为长株潭两型社会建设试验区之一,应该有着良好的创新环境,但是排名却居于第七位。这可能和选取的指标有关。由表 1 可以知道,公路货运量和公路长度这两个指标的平均值分别是 9688.0714 万吨、16285.56 公里,而湘潭这两个指标分别都只有 5665 万吨和 7700.26 公里,这确实确实影响了 12 因子的得分。不过这从一个侧面可以反映出在长株潭城市的发展中,长沙起着带头作用,湘潭相对落后,在社会基础设施建设方面还应该加强。

邵阳、永州、娄底和益阳的科技金融动力不充足,科研机构不多,科技活动人员较少,吸引的 FDI 不多,创新环境有待进一步提高。

相比之下,湘西地区的怀化、张家界和吉首因为区位优势相对较差,区域内基础设施条件不是很好,交通不便利,信息交流不顺畅,能源缺乏,经济基础薄弱,发展程度偏低,外商企业注资较少;高新技术企业也少;企业办的科研机构怀化只有 10 家,张家界仅仅只有 3 家,而位居第一的长沙有 151 家,这严重阻碍了科技金融动力因子的增长。

四 结语

本文利用 SPSS 软件对湖南省创新环境进行了因子分析,得出了各地市的创新环境的综合得分和排名,发现湖南省区域创新环境总体水平不是很高,创新环境的地区差异也是客观存在的。只有充分地认识到这些差异,才能制定有益于各地市协调发展的机制,实现湖南省的全面创新。

然而,本文的实证也存在着不足。首先选取的

数据是来自 2010 年的《湖南省统计年鉴》,这仅仅是对 2010 年的湖南省创新环境的评价,有一定的局限性。其次,本文指标体系的构建是基于指标的可行性和数据的可得性上建立的,还不能完全绝对地反映创新环境的优劣。第三,缺少城市群之间的比较,如果可以把城市按照地域进行划类,得到的结论可能更具有可比性。

注释:

① 湖南省地区生产总值(亿元)的数据来源于湖南省 2011 年统计年鉴。

[参考文献]

[1] 张文忠. 区域创新环境与企业发展研究[J]. 软科学, 2003, 17(6): 25-28.

[2] Martina Fromhold-Eisebith. Innovative milieu and social capital complementary or redundant concepts of collaboration- based regional development[J]. European Planning Studies, 2004, 12(6): 747-765.

[3] 陈伟. 区域创新体系的内涵-分类及城市群创新演进趋势[J]. 当代经济, 2012(9): 84-85.

[4] 王颢晨. 基于系统动力学的广西区域创新系统研究[J]. 科学学与科学技术管理, 2008(6): 66-71.

[5] 盖文启. 论区域经济发展与区域创新环境[J]. 学术研究, 2002(1): 60-63.

[6] 王树林. 区域创新环境与区域经济发展[J]. 商业研究, 2002(24): 59-60.

[7] 李习保. 中国区域创新能力变迁的实证分析: 基于创新系统的观点[J]. 管理世界, 2007(12): 18-30.

[8] 岳鹤. RD 投入、创新环境与区域创新能力关系研究: 1997-2006[J]. 当代经济科学, 2008, 30(6): 110-116.

[9] 张莹. 区域创新环境对创新绩效影响的实证研究[J]. 科技管理研究, 2009(2): 2-10.

[10] 许婷婷. 江苏省区域创新环境对创新产出影响分析[D]. 南京: 南京航空航天大学硕士学位论文, 2012: 2-6.

[11] 杜强. SPSS 统计分析从入门到精通[M]. 北京: 人民邮电出版社, 2009: 271-282.

Studies on the Regional Innovation Environment of Hunan Province Based on Factor Analysis

HUANG Ling-yan

(Hunan University of Technology, Zhuzhou 412007, China)

Abstract: This paper firstly builds a regional innovation environment evaluated system, and then applies factor analysis for getting major indexes. At last, the paper makes evaluation and analysis on regional innovation environment of Hunan province, to reveal the difference and characteristics in regional innovation environment of Hunan.

Key words: regional innovation environment; factor analysis