

中国工业化指数的计算与分析

崔向阳

摘要: 本文运用主成分分析计算出了我国各地区的工业化指数,并运用聚类分析、图形分析、相关分析法分析了我国工业化的特点,最后得出研究结论,必须处理好工业化和市场化、城市化、信息化的关系,发展第三产业,协调地区经济发展。

关键词: 工业化指数 主成分分析 聚类分析 图形分析 相关分析

中共十六大提出了新型工业化理论以后,工业化问题再度成为一个新的研究重点和热点。我国的工业化进程如果从“一五”算起,已经历了50年,我国的工业化水平到底是什么程度,这是大力推进新型工业化的基本问题。从国内学术界的研究现状看,学者们对我国工业化阶段的分析,主要还是运用国际流行的工业化理论,从各个方面作出判断,很少有人去尝试计算一个综合性的工业化指数。本文拟运用主成分法计算我国的工业化指数,并运用这一指数对我国工业化的有关问题进行分析。

一、工业化的概念

研究表明,工业增长、生产率增长与国内生产总值(GDP)增长之间存在密切的关系,这种关系也被称为卡尔多增长规律(Kaldor,1966)。一项对我国28个地区的研究(Hansen and Zhang,1996),证明了卡尔多增长规律对我国的适用性,研究使用了1965-1991年的时间序列数据,得出的回归结果为: $g_{GDP} = 1.79 + 0.56g_m$, $R^2 = 0.67$,式中, g_{GDP} 表示GDP增长, g_m 表示制造业产出,方程表明,中国28个地区之间产出增长率的67%的差别可以由各地区制造业产出增长的差异来解释。近年来,我国经济的高速增长也是与第二产业特别是制造业的高速增长联系在一起的,1994-2000年我国制造业对GDP增长的贡献分别为:39.4%、42.2%、43.4%、44.2%、23.3%、31.3%、45.9%^①,因此,工业尤其是制造业被称为“增长的发动机”。

但关于工业化的含义,在学术界有不同的理解。一般认为,工业化表现为工业产值比重和就业人口比重不断上升,同时农业产值比重和就业人口比重不断下降的过程。例如,《新帕尔格雷夫经济学大辞典》对工业化的定义是:“工业化是一种过程。首先,一般来说,国民收入(或地区收入)中制造业活动和第三产业所占比例提高了;其次,在制造业和第三产业就业的劳动人口的比例一般也有增加的趋势。在两种比率增加的同时,除了暂时的中断以外,整个人口的人均收入也增加了。”^②这种定义比较简明,也比较好理解,绝大多数学者持这种观点,例如,著名发展经济学家刘易斯、钱纳里、库兹涅茨等人均持相同或类似的观点。

我国著名经济学家张培刚认为,所谓工业化就是指“一系列基要(strategical)生产函数连续发生变化的过程”^③。这种变化最先发生于某个生产单位的生产函数,然后形成一种社会生产函数而遍及于整个社会。很显然,这是一种比较宽泛的定义,不但指出了工业部门自身的变化,还指出了工业化引起的工厂制度、市场结构等经济组织的变化。改革以前,我国工业化的特点是工业化与市场化、城市化相脱节,工

业化与人均收入、就业结构相脱节,造成了一种片面的、畸形的工业化。随着我国市场化进程的推进,这种状况有很大的改变,但还需要一个很长的过程,才能完全得到本来意义上的工业化带来的收益。很显然,造成这种状况与狭义的理解工业化有很大的关系,因此,我们认为,强调宽泛的工业化定义,比较符合我国的实际情况,对我国更有理论意义和现实意义。

二、指标的选取

从国内学术界的研究现状看,对我国工业化阶段的分析,大多是依据标准工业化理论来进行的,如胡健生的《对江苏工业化程度和所处阶段的判断》[载《江苏经济探讨》,1997(1)]、郭克莎的《中国工业化的进程、问题与出路》[载《中国社会科学》,2000(3)]。这些标准工业化理论主要是钱纳里的人均收入6阶段理论、库兹涅茨的产值结构和就业结构5阶段理论、霍夫曼工业结构4阶段理论等。

1. 钱纳里的人均收入6阶段理论

钱纳里等人从结构转变过程的角度将各国的人均GDP水平划分为6个变动时期(钱纳里,1989),如表1所示:

表1 钱纳里的人均收入6阶段(人均GDP)

(单位:美元)

	1964年	1970年	1982年	1994年
第1阶段	100~200	140~2803	364~728	946~1893
第2阶段	200~400	280~560	728~1456	1893~3786
第3阶段	400~800	560~1120	1456~2912	3786~7571
第4阶段	800~1500	1120~2100	2912~5460	7571~14196
第5阶段	1500~2400	2100~3360	5460~8736	14196~22714
第6阶段	2400~3600	3360~5040	8736~13104	22714~34070

钱纳里认为第2-5阶段是工业化时期,第1阶段和第6阶段是起始时期和发达时期。人均收入的不断提高是工业化的自然后果,综合反映了工业化的经济效益,因此,钱纳里模型已成为评价工业化阶段的标准理论。

2. 库兹涅茨的产值结构和就业结构5阶段理论

库兹涅茨根据对57个国家的原始资料处理结果,整理出1958年按人均GDP为基准的产业结构变化趋势,同时,他还根据1958年国内生产总值进一步考察了59个国家1960年劳动力在三次产业中所占份额,得出工业化5阶段理论,见表2、表3。

表2 1958年产值结构5阶段(%)

	70美元	150美元	300美元	500美元	1000美元
第一产业	48.4	36.8	26.4	18.7	11.7
第二产业	20.6	26.3	33.0	40.9	48.4
第三产业	31.0	36.9	40.6	40.4	39.9

表 3 1958 年就业结构 5 阶段 (%)

	70 美元	150 美元	300 美元	500 美元	1 000 美元
第一产业	80.5	63.3	46.1	31.4	17.0
第二产业	9.6	17.0	26.8	36.0	45.6
第三产业	9.9	19.7	27.1	32.6	37.4

工业化在产业结构变动中表现得最为充分,从三次产业产值结构的变动看,在工业化起点,第一产业的比重较高,第二产业的比重较低;随着工业化的推进,第一产业的比重持续下降,第二产业的比重迅速上升,而第三产业的比重只是缓慢提高。工业的扩展引起产业结构的迅速转变,这种转变又带动了就业结构的变动,因此,产业结构和就业结构变动的一般趋势,反映了工业化演进的阶段性。

3. 霍夫曼工业结构 4 阶段理论

霍夫曼分析了工业化过程中工业内部的结构变动状况,通过计算消费品工业与资本品工业的比重(霍夫曼比例),提出了工业化 4 阶段理论,见表 4。

霍夫曼的理论揭示了工业化过程中工业部门内部结构演变的一般趋势。因为,资本品工业与消费品工业相比,具有资本、技术密集的特点,同时资本品工业比例增大意味着工业加工程度深化,中间产品和最终产品比例扩大。所以,

霍夫曼比例越低,说明资本品工业规模越发展,相应地消费品工业比重越小,工业结构乃至整个产业结构高度越高。因此,霍夫曼比例反映了工业化的进程。

表 4 霍夫曼的工业结构 4 阶段

工业化阶段	霍夫曼比例
第 1 阶段	5
第 2 阶段	2.5
第 3 阶段	1
第 4 阶段	< 1

根据以上对工业化概念和标准工业化理论的研究,再考虑我国的实际情况和数据的可得性,我们认为可以选取以下 4 类指标作为计算工业化指数的依据:

(1) 工业结构指标:重轻工业比例;(2) 产业结构指标:非农产业产值比重;(3) 收入水平指标:人均 GDP;(4) 就业结构指标:非农产业劳动力比重。

三、工业化指数的计算

我们首先依据《中国统计年鉴 2000》中我国各地区 1999 年的数据,计算了我们选取的 4 个指标,原始指标的计算方法是:

$$\text{重轻工业比例 } x_2 = \text{重工业总产值} / \text{轻工业总产值}^{\text{④}}$$

表 5 各地区工业化综合指数

地区 x_1	名次 no.	重轻工业 比例 x_2	非农产业 产值比重 x_3	人均 GDP x_4	非农产业劳 动力比 x_5	工业化指数 z
上海	1	0.6400	98.0000	30 805.00	86.2000	10.00
北京	2	0.3708	96.0000	19 846.00	88.1000	7.91
天津	3	0.5880	95.1000	15 976.00	80.4000	7.08
浙江	4	1.3002	88.2000	12 037.00	59.4000	5.61
广东	5	1.3278	88.0000	11 728.00	58.8000	5.55
江苏	6	0.8315	87.0000	10 665.00	57.5000	4.65
福建	7	1.1717	82.3000	10 797.00	51.5000	4.36
辽宁	8	0.2577	87.5000	10 086.00	62.4000	4.19
山东	9	0.8702	84.1000	8 673.000	47.1000	3.59
黑龙江	10	0.2917	87.0000	7 660.000	51.7000	3.25
湖北	11	0.6436	83.0000	6 514.000	51.1000	3.12
河北	12	0.5520	82.4000	6 932.000	51.4000	3.04
山西	13	0.1697	89.4000	4 727.000	54.0000	2.97
海南	14	1.8813	62.5000	6 383.000	39.0000	2.21
陕西	15	0.4830	82.0000	4 101.000	43.0000	2.04
重庆	16	0.5165	80.7000	4 826.000	41.5000	2.00
吉林	17	0.3068	74.6000	6 341.000	50.7000	1.98
江西	18	0.5475	76.3000	4 661.000	44.8000	1.83
新疆	19	0.3257	77.0000	6 470.000	42.5000	1.79
湖南	20	0.5618	76.6000	5 105.000	39.6000	1.66
安徽	21	0.7599	74.5000	4 707.000	39.4000	1.65
云南	22	1.1522	77.8000	4 452.000	26.2000	1.64
青海	23	0.1247	83.0000	4 662.000	39.1000	1.58
四川	24	0.7542	74.6000	4 452.000	38.7000	1.57
宁夏	25	0.1909	80.1000	4 473.000	41.4000	1.51
内蒙古	26	0.3829	72.9000	5 350.000	45.4000	1.49
河南	27	0.6120	75.5000	4 894.000	36.4000	1.43
甘肃	28	0.1997	79.5000	3 668.000	41.1000	1.33
广西	29	0.7360	71.6000	4 148.000	34.6000	1.04
贵州	30	0.5126	70.6000	2 475.000	27.5000	0.06
西藏	31	0.5919	67.6000	4 262.000	24.1000	0.00

非农产业产值比重 $x_3 = \text{GDP 中第二产业比重} + \text{GDP 中第三产业比重}^{\text{⑤}}$

人均 GDP x_4 直接取自统计年鉴^⑥

非农产业劳动力比重 $x_5 = \text{年底第二产业从业人员比重} + \text{年底第三产业从业人员比重}^{\text{⑦}}$

计算综合指数的关键问题是如何确定各指标的权重,我们采用了处理这一类问题比较成熟的主成分分析法来计算。主成分分析的基本步骤是:(1)对样本数据进行标准化变换;

(2)计算相关系数矩阵;(3)计算相关系数矩阵的特征根和特征向量;(4)计算特征根的方差贡献率和累计贡献率;(5)根据累计贡献率确定主成分;(6)根据主成分的特征与对应的特征向量计算主成分;(7)根据主成分的方差贡献率计算综合得分。需要说明的是,我们使用了 SPSS 11.0 for Windows 来处理有关计算问题,由于对样本数据进行了标准化变换,计算出的综合得分有 19 个地区出现了负数,为了消除负数,我们又对综合得分使用了规格化变换的技术方法,规格化变换

的公式是：

规格化数据 = (原始数据 - 原始数据的最小值) / (原始数据的最大值 - 原始数据的最小值)，规格化变换后，规格化数据的取值范围在 0~1 之间，其中，最大数值为 1，最小数值为 0。为了符合习惯，我们将规格化变换后的数据乘以 10，工业化指数的最大值为 10，最小数值为 0。当然，也可以直接采用综合得分作为工业化指数，规格化变换与各地区的最后排名无关。

有关指标和最后计算出的各地区工业化综合指数如表 5 所示。^⑤

我们计算出的结果实际上是我国各地区工业化的相对指数，只说明各地区在全国工业化进程中的相对位置，并不说明各地区工业化的绝对水平，但可以用于截面比较和研究，如果采用这套方法计算出各年的工业化指数，也可以用于序时比较和研究。

四、中国工业化的地区分析

1. 聚类分析

为了更清楚地观察各地工业化的相对水平，我们使用了多元统计分析中的聚类分析方法，按照我们计算的工业化指数对各地进行了分类。聚类分析的基本思路是按照样本的距离和相似系数，把相近的归为一类，然后再逐步合并，直到所有样本合并为一个大类。通过聚类分析，可以恰当地对样本进行分类，避免了主观分类的随意性。

我们使用了 SPSS 11.0 for Windows 的 Hierarchical Cluster 程序来进行研究，根据电脑输出的树状图可以看出^⑥，可以按照工业化指数把全国分为 5 大类，第 1 类包括上海、北京、天津 3 个直辖市，工业化指数在 7.0 以上，可以称为工业化发达地区；第 2 类包括浙江、广东、江苏、福建、辽宁 5 个省，工业化指数在 4.1~5.6 之间，可以称为工业化较发达地区；第 3 类包括山东、黑龙江、湖北、河北、山西 5 个省，工业化指数在 2.9~3.6 之间，可以称为工业化一般地区；第 4 类包括海南、山西、重庆、吉林、江西、新疆、湖南、安徽、云南、青海、四川、宁夏、内蒙古、河南、甘肃、广西 16 个地区，工业化指数在 1.0~2.2 之间，可以称为工业化不发达地区，第 5 类包括贵州、西藏 2 个地区，工业化指数在 0.06 以下，可以称为工业化最不发达地区。

我们计算出的全国工业化指数的平均数是 3，中位数是 2，在平均数以上的地区只有 12 个。各类地区的平均工业化指数分别是：发达地区 8，较发达地区 5，一般地区 3，不发达地区 1，最不发达地区 0.03。最发达的 3 个直辖市的平均工业化指数高达 8，而最不发达的 2 个地区的平均工业化指数只有 0.03，相差非常悬殊。从上面的分析可以看出，我国的工业化进程极不平衡，大城市工业化水平高，农村地区工业化水平低；沿海地区工业化水平高，中西部地区工业化水平低；工业化水平高的地区少，工业化水平低的地区多。

2. 图形分析

为了找出影响我国各地区工业化水平的关键因素，有必要对各类工业化地区的特点进行比较，我们使用了多维数据图式法中最简单的雷达图和折线图，来对影响工业化的工业结构、产业结构、人均收入、就业结构等 4 个方面进行了比较分析。

我们首先将计算出的指标进行标准化处理，接着对标准化数据进行了规格化处理以消去负数，然后计算了 5 类地区的工业结构、产业结构、人均收入、就业结构 4 项指标的均值，作为各类地区各项指标的得分，计算的结果如表 6。

根据表 6 的数据，我们使用 Microsoft Graphy 2000 画出的雷达图和折线图如图 1 和图 2。

表 6

	工业结构	产业结构	人均收入	就业结构
工业化发达地区	0.23242	0.95399	0.69658	0.95000
工业化较发达地区	0.48565	0.67887	0.30313	0.52844
工业化一般地区	0.21675	0.63887	0.15624	0.42124
工业化不发达地区	0.26828	0.38591	0.08624	0.25175
工业化最不发达地区	0.24343	0.18591	0.03154	0.02658

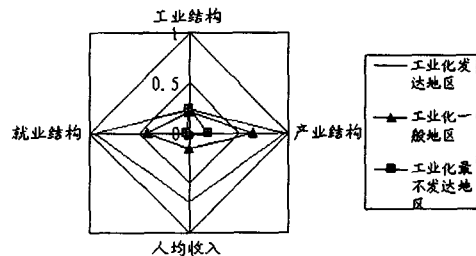


图 1

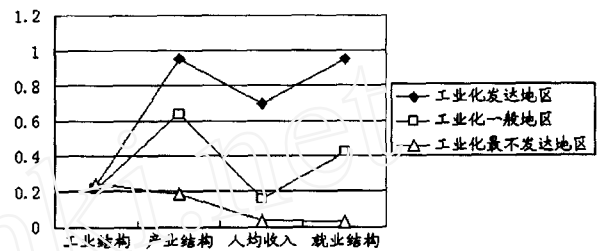


图 2

为了避免线条紊乱，我们只画了发达地区、一般地区、最不发达地区三条线。但已经清楚的看出，工业结构对工业化水平影响不大，产业结构有一定的影响，人均收入、就业结构 2 项指标对工业化影响最大。

从原始数据来看，工业结构指标即重轻工业比例 x_2 的全国的均值是 0.6，中位数是 0.5，发达地区的平均数是 0.5，并不突出，最不发达地区是 0.55，也不落后。从产业结构指标即非农产业产值比重 x_3 来看，全国的均值和中位数均是 80，发达地区的平均数是 96，一般地区是 85，最不发达地区是 69，不同地区的差距不是很大。从收入水平指标即人均 GDP x_4 来看，全国的均值是 7 802 元，中位数是 5 350 元，发达地区的平均数是 22 209 元，一般地区是 6 901 元，最不发达地区是 3 368 元，发达地区是一般地区的 3 倍、最不发达地区的 6 倍，不同地区的差距非常大。从就业结构指标即非农产业劳动力比重 x_5 来看，全国的均值是 48，中位数是 45，发达地区的平均数是 85，一般地区是 51，最不发达地区是 26，不同地区的差距比较明显。因此，从这 4 项指标来看，决定工业化水平高低的最关键因素是 2 项指标，第一是人均收入，第二是就业结构。

3. 相关分析

我国工业化的重要特点是在计划经济的条件下开始起步，在工业化迅速推进的过程中市场化没有得到相应的发展，造成了很多结构性偏差，例如，工业化与城市化脱节，工业化与第三产业的发展脱节，工业化与人均收入的提高脱节，所有制形式过于集中等许多问题。改革以来，我国的资源配置方式逐步转向了市场经济，但工业化与市场化的关系仍是人们关注的重要问题。

我们使用了国内某研究机构^⑦发布的中国市场化指数和我们计算的中国工业化指数，对 1999 年我国 30 个地区的工业化水平与市场化水平的关系进行了测算。该机构发布的 1999 年中国各地区的市场化指数如下：上海 6.59、北京 6.30、天津 6.58、浙江 8.24、广东 8.33、江苏 7.04、福建 7.28、辽宁 5.60、山东 6.22、黑龙江 3.97、湖北 5.53、河北 6.70、山西

4.57、海南 5.65、陕西 4.48、重庆 5.57、吉林 4.51、江西 5.12、新疆 2.90、湖南 5.99、安徽 5.40、云南 3.39、青海 2.00、四川 5.29、宁夏 2.69、内蒙古 3.45、河南 5.97、甘肃 4.02、广西 5.28、贵州 3.86(该套市场化指数缺西藏的数字)。

为了确定最佳模型,我们使用 SPSS 11.0 for Windows 分别进行了线形模型、对数模型、双曲线模型、复合曲线模型、S 曲线模型、二次曲线模型、三次曲线模型、幂函数模型、增长曲线模型、指数模型的适配,最后发现三次曲线模型的可决系数最大,我们决定采用三次模型,最后得到的计量模型如下:

$$zz = 11.2188 - 6.9733 * sch + 1.4601 * sch^2 - 0.0840 * sch^3 \quad R^2 = 0.45 \quad F = 7.09 \quad sigf = 0.001$$

上式中 zz 表示工业化指数, sch 表示市场化指数。可决系数达 0.45,说明各地区的工业化与市场化相关性比较高,45.5%的工业化水平差异可以用市场化来解释。而且,如果剔除上海、北京、天津 3 个变异较大的样本, R² 的值可以提高到 0.70。

通过对图 3 拟合曲线的观察,我们发现,工业化水平与市场化水平的关系呈倒 S 形,曲线有两个拐点,在两个拐点之间,曲线的斜率大于零,随着市场化指数的增大,工业化指数迅速爬升,在第一个拐点之前和第二个拐点之后,两者的相关关系不明显。进一步的研究发现,第二个拐点是京津沪引起的,如果剔除这 3 个发达地区,在图形中就看不到第二个拐点了。因此,在我国的现阶段,对绝大部分地区来说,市场化对工业化的推动作用不是变小了,或是一次性的,而是越来越大,推进市场化可以迅速推动工业化的发展。

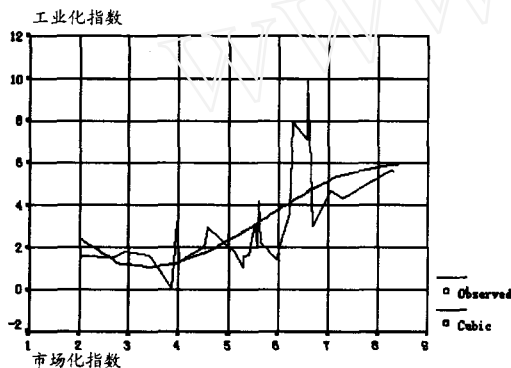


图 3

五、研究结论

1. 以市场化促进工业化,优化工业化的制度环境

在工业化过程中,要进一步发挥市场机制的资源配置优势。要加快发育要素市场,鼓励民间资本投资,保护私人产权,培养内源性增长机制。特别是要处理好虚拟经济和实体经济的关系,大力发展资本市场,拓宽融投资的渠道。我国现在已相对不缺资本,缺的是资本转化为投资的制度供给。要进一步放开,凡是向外资开放的产业,都应允许民间资本进入。政府的管制越少越好,要从审批经济转向登记经济。

2. 以城市化促进工业化,理顺工业化和城市化的关系

城市具有规模效应、积聚效应,是信息中心、资金中心、物流中心,城市的发展对工业化有极大的促进作用,我国工业化水平最高的地区正好是我国最大的 3 个城市,但这并不是巧合。各种要素不断向城市集中是工业化阶段的一个普遍规律,我国曾长期人为割裂工业化和城市化的关系,现在看来,城市发展不足已经成为中国工业化发展过程中的最大制约因素,必须要大力发展城市化,提高城市化水平,加快农村富余劳动力向非农业和城镇转移的速度,特别是要发挥中心城市的带动作用。近年来,江浙一带不同地区经济发展

的差异,完全可以用该地区与中国的经济中心上海的空间距离来解释,各地必须创造条件,主动接受中心城市经济辐射。

3. 以信息化带动工业化,推进产业结构优化升级

我国是一个人力资源丰富的国家,要充分发挥我国的人力资源优势,既要充分利用工业化和信息化来提升劳动生产率,实现产业结构的升级换代,又要着眼于扩大就业,发挥我国的比较优势。在处理信息化与工业化的关系上,应重点放在用信息化改造传统产业上,一方面,用信息化改造传统产业,可以提高传统产业的国际竞争力,吸纳更多的劳动力;另一方面,传统产业的不断升级,也可以为信息产业的发展提供坚实的基础和广大的市场。要注意把我国具有比较优势的人力资源和先进的生产要素相结合,使比较优势转化为竞争优势,不断提升我国的产业竞争力。

4. 发展第三产业,提高人民收入水平

制约我国工业化水平的一个重要方面是人均 GDP 太低,必须尽快提高人民收入,大力发展第三产业正是一个突破口。大力发展第三产业,可以增加消费,带动经济增长;可以提供更多的就业岗位,加快农村剩余劳动力的转移;可以积聚人气,促进城市化进程;可以引入更多竞争,加快市场化进程。随着工业化的发展,第一产业逐渐萎缩,第二产业基本稳定,第三产业的发展空间最大,大力发展第三产业是解决我国经济中诸多矛盾的重要手段,特别是要加快发展现代服务业,提高第三产业在国民经济中的比重。

5. 协调地区经济发展,全面建设小康社会

我国工业化的重要特点是地区发展不平衡,必须继续坚持西部大开发战略,加快中西部地区的发展。中西部地区是东部地区重要的市场来源、要素来源和产品来源,东部地区的持续发展离不开中西部的支持。中共十六大报告提出要全面建设小康社会,东部地区率先基本实现现代化,中西部地区要增强自我发展能力,加强东、中、西部经济交流和合作,努力实现地区经济的协调发展和共同富裕。

注释:

- ①根据《中国统计年鉴 2001》(北京,中国统计出版社,2001)有关数据整理计算。
- ②《新帕尔格雷夫经济学大辞典》,中文版,第 2 卷,861 页,北京,经济科学出版社,1996。
- ③张培刚:《农业与工业化》,武汉,华中工学院出版社,1984。
- ④⑤⑥⑦《中国统计年鉴》(2000),表 13-5.3-9.3-9.5-3,北京,中国统计出版社,2000。
- ⑧具体的计算过程备索。
- ⑨为了节省篇幅,省略了原图。
- ⑩国民经济研究所 2001 年 8 月发布了包括 5 个方面、19 个指标的中国各地区市场化进程相对指数。

参考文献:

1. Kaldor, N., Cause of the Slow Rate of Economic Growth of the United Kingdom Cambridge. Cambridge University Press, 1966.
2. Hanson, J.D. and Zhang, J., A Kaldorian Approach to Regional Economic Growth in China. Applied Economics, June, 1996.
3. 钱纳里、鲁宾逊、赛尔奎因:《工业化和经济增长的比较研究》,中文版,上海,上海三联书店,1989。
4. 库兹涅茨:《现代经济增长》,中文版,北京,北京经济学院出版社,1989。
5. 霍夫曼:《产业结构问题研究》,中文版,北京,中国人民大学出版社,1997。
6. 雷钦礼:《经济管理多元统计分析》,中文版,北京,中国统计出版社,2002。

(作者单位:南京大学商学院 南京 210093)

(责任编辑:N)