

如何看待社会发展指标的综合评估

兼评“中国现代化研究报告”

谢立中(北京大学社会学系教授、哲学博士)

前一段时间,国内不少地方的领导人都纷纷表态,要使自己领导的那个地区实现“现代化”,国内最权威的科研机构之一——中国科学院的一个课题组也公布了“中国现代化研究报告”(载自科学出版社《2001中国可持续发展战略报告》一书,以下简称《报告》)。该《报告》根据自己提出的“现代化指标体系”,计算了全国及各地区现代化实现程度与中等发达国家水平的差距,以及全国及各地区实现现代化的时间表。对于这种计算结果,我们应该怎样看待?

与以往的一些社会发展综合评价指标体系一样,《报告》对中国“现代化”进程的评价和预测,也是采用近几十年发展起来的多指标综合评估技术。然而,正如众所周知的那样,尽管这种多指标综合评估技术在其他领域中的应用可能很有成效,但它在社会发展评估领域中的应用,应该说一直很不理想。从二十世纪六七十年代迄今,不少专家学者运用这种多指标综合评估技术开发、建立起来的各种社会发展综合评估指标体系,无一能够得到人们的公认和接受。盖其原因,就在于这种多指标综合评估技术在运用于社会发展综合评价和预测时,存在着许多至今难以解决的问题,从而在很大程度上影响了运用这种技术建立起来的那些综合评估指标体系的合理性与可靠性。

迄今为止尚未得出公认的指标体系

指标的选择就是一个难以合理解决的重要问题。每一个社会发展综合评估指标体系(包括《报告》中的“现代化”指标体系)都是由许多相关的指标元素所构成的,选择适当的指标成为建构这种综合评估指标体系的首要任务。但是,从众多相关的候选指标中选择哪些指标来作为指标体系的构成元素,就很难有一个唯一的、能为众人所接受的答案。目前人们一般采用由研究人员或课题组自定、通过专家调查确定以及运用有关统计技术(如主成分分析法等)确定这几种方法来完成指标选择的工作。前两种方法具有很大的主观随意性。由于指标的选择在很大程度上会受到研究人员或被调查专家个人知识、价值观念以及情感等多种主观因素的影响,由不同的研究人员(或课题组)来自行选择指标,或对不同的专家群体进行调查来选取指标,所得出的结果是会很不相同的。这也是为什么国内外尚无一种社会发展综合评估指标体系能够为人们所普遍接受认可的一个重要原因。由于不同的指标在变化的速度等方面情况不同,最终计算出来的综合评估指数也会大相径庭。

用统计分析技术来选择指标,主观随意性会有很大降低。以主成分分析法为例,它主要

是通过计算各种候选指标，在各个样本之间的相对差距来作为指标选择的依据，将各个样本之间具有相对较大差距的那些指标抽取出来。作为建构综合评估指标体系的元素，而将样本之间差距不大的那些指标加以排除。以这样方式来选取指标确实在很大程度上可以排除主观因素的影响（只要候选指标足够多），但它也有自己的弱点。一是在不同的样本集合中，同一项指标会具有不同的相对差距，因此，只要样本集合变化，用主成分分析方法挑选出来的指标内容就可能不同。可见，主成分分析方法也不能保证指标选择结果的唯一性。二是即使在一个确定的样本集合中，这种依主成分分析等统计分析方法选择出来的指标体系也只能够用来对样本进行排序，否则其有效性也是值得质疑的。

我们可以举一个很简单的例子来加以说明。设要对甲、乙两名学生的素质进行综合评估，需要建立一个综合性的评估指标体系，假定候选指标包括德、智、体三个方面，经有关素质质量表测定，甲在德、智、体三方面各自的得分是 5、5、5，乙在三方面相应的得分是 5、3、5。用主成分分析法分析，可以确定德、体两个候选指标在样本之间的差距为零，智力素质指标在样本间的差距最大。因此，可以只选择智力因素来作为对这两个学生综合素质进行评估的基本依据（也就是说这个综合评估指标体系只由一个指标元素构成就够了），用它来对这两个学生的综合素质进行评价，与同时用这三个指标来共同进行评价，得出的排序结果是一样的（这里暂时假设每个指标的权数是相等的）。但是如果不是用这个指标体系来对样本进行排序，而是要对样本的综合素质做连续测定与比较，那么只用智力这一个指标来对各个样本进行评价就会出现讹误。例如，假如以等权方式对上述三个方面的得分进行加总，那么学生甲的实际总得分应是 15 分，乙的总得分应是 13 分，若依次对两人进行排序，结果与用智力一项指标得出的结果一样。但若是要对他们两人之间的综合素质做连续数值上的比较，那么，以三项总分来看，甲的优秀程度是乙的 $1\frac{2}{13}$ 倍（ $15 / 13$ ），而若以智力一项得分来看，则甲的优秀程度是乙的 $1\frac{2}{3}$ 倍（ $5 / 3$ ），甲的优秀程度显然被夸大了。

上面提到的加权问题是建立社会发展综合指标体系中的另一个明显问题。通过上述方法选取出来的指标在经过无量纲化处理变成各种单个指数后，需要以各种方式加总起来，才能最后得出我们所需要的那个社会发展评价综合指数（就《报告》而言，就是用来评价各地“现代化”实现程度的那个最终的总分数）。但在对各个指数进行加总时，有一个对各个指数如何进行赋权的问题（不赋权即等于视它们为等权，如《报告》所做的那样）。以目前人们通常采用的那些确定指标权数的方法来看，同样存在着较大的问题。与进行指标选择这项工作一样，目前人们也主要是采用研究人员自定、专家调查和统计分析等方法来对指标权数加以确定。由于所谓赋权的过程，实际上就是对各个指标在整个指标体系中（也就是在社会发展的整体过程中）的相对重要性程度进行评价和确认的过程，因此如果采用前两种方法来对指标赋权，其结果同样会在很大程度上受到研究人员或专家们的知识背景、价值观念和情感等主观因素的影响。由不同的研究人员来对指标自行进行赋权，或通过对不同的专家集合进行调查来对指标进行赋权，所得出的结果也是会很不相同的。而只要指标的权数结构不同，运用同一个指标体系来对同一批样本进行测算和评价，其结果也就会很不相同。

采用有关的统计分析技术（如主成分分析法）来对指标的权数加以确定，其原理也如同指标选择过程一样，即是通过一定的统计分析方法，测算出各个指标在样本之间的相对差距，然后将各个指标在样本之间相对的差距大小来作为对指标进行赋权的依据。换句话说，就是相对差距越大的指标权数也就越高，反之则越低。这种方法虽然可以解决上述主观随意性问题，但同样有不合理的地方：首先是如前所述，同一个指标在不同的样本集合中会有不同的相对差距，因此样本集合的变化也可能导致指标权数的变化。此外，在一个确定的样本集合中，样本之间差距较大的那些指标并不一定真的就是人们公认为在实际社会生活中相对较重要的那些指标。因此以这种方式来确定指标权数，最终对社会发展过程进行测算的结果和人们的实际感受也就可能产生较大的偏离，不能真正反映人们感受到的社会发展的实际情况。

总而言之，由于存在着上述种种问题，使得社会发展综合评估成为一项难度很高的工作。拿我们这里所评论的《报告》来说，如果对《报告》所使用的“现代化”指标体系中的指标内容或各个指标的权数做一些变动（从目前的等权变成不等权，即赋予其中某些指标以相对更高的权数），那么无论是对中国及国内各地区现有“现代化”水平的测算结果还是对未来实现“现代化”年份的测算结果就都可能会与《报告》目前的测算结果很不一样。以你的指标内容及权数结构，计算出某地区将在2020年“实现现代化”；而以我的指标内容或权数结构，计算出来的结果则可能是提前或推后若干年。所以，对于此类的综合评估及测算结果，我们的态度只能参考而已。这恐怕也是国外很少有哪个国家或地区的领导人会认真地将这种综合评估的结果当作自己决策依据的一个重要原因。

真有一个“英格尔斯指标体系”吗

顺便谈谈所谓的“英格尔斯指标体系”。实际上除了中国以外，世界上根本不存在着一个用来对“现代化”进程进行综合评估的“英格尔斯指标体系”。据我所知，所谓“英格尔斯指标体系”是这样产生出来的：1983年美国著名社会学家英格尔斯教授应北京大学社会学系之邀到该系讲学，讲课中提到有一个叫拉西特（Russett）的学者曾经在70年代从60个特征对72个国家（包括发达与不发达国家）进行了比较研究，得出了一些有意思的结果。他主要对其中的十四个方面（人均国民生产总值、农业产值占总产值比重、城市人口比重等）做了简单介绍，并提出说可以以某个典型数字任意划一条线（譬如人均年收入3000美元），来区分发达国家与不发达国家，观察他们在其他各个指标上的区别，将它们看作是发达国家与不发达国家的另一些分界线。这里有两点要指出：第一，英格尔斯在这里介绍的只是一些分散的指标，并不是一个我们今天所理解的“综合评估指标体系”，无论是他还是拉西特似乎都不曾像我们中国学者今天所做的那样，用它来对各个国家的“现代化”进程进行综合评估；第二，英格尔斯只是在介绍拉西特的研究成果，而不是他自己的成果，因此，如果要叫一个名称，似乎也应该称之为“拉西特指标体系”更准确。这些讲课内容后来通过该系教师孙立平、韩明谟等人在文章或著作中加以引用而得以发表。中国社会科学院社会学研究所朱庆芳研究员看到后，觉得很有启发，遂将其进一步发展成为一个“综合评估指标体系”。这就是所谓“英格尔斯指标体系”的由来。它作为一个我们今天所看到的“综合评估指标体系”，实际上主要是中国学者的发明。当然，这些补论无关宏旨，只是写

到此处，顺便澄清一下而已，使人不再误以为国外也有什么“英格尔斯综合指标体系”。