

OECD/JRC 合成指标手册评论

——基于指标方法论的“机理进路”^①

邱东,郭同济

(江西财经大学统计与数据科学学院,江西南昌330013)

摘要:注重合成指标方法论的“机理进路”是避开测度陷阱、实现“数理”与“事理”统一的基础。文章从方法论比较的角度对 OECD/JRC 合成指标手册进行评论,从阐述合成指标的方法论的学理层面标准评价 OECD/JRC 手册的优势和不足并针对合成指标手册“十个构建步骤”提出意见和思考,在此基础上剖析了“手册悖境”的成因,指出打破悖境的重要指向正是注重合成指标的机理研究。中国学者应自信自强,通过发展具有中国自主知识体系的经济统计学科,在自主参与甚至积极引导国际统计规则的制定方面发挥应有作用。

关键词:多指标综合评价;OECD/JRC 合成指标手册;经济统计;方法论机理;手册悖境

中图分类号:F21 **文献标志码:**A **文章编号:**1000-2154(2025)02-0005-10

DOI:10.14134/j.cnki.cn33-1336/f.2025.02.001

收稿日期:2025-01-15

基金项目:国家社会科学基金重大项目“国际统计标准测度问题挖掘与中国参与的方法论基础研究”(18ZDA123);国家社会科学基金项目“加快构建中国特色哲学社会科学学科体系、学术体系、话语体系”研究专项“新时代中国特色统计学基本问题研究”(19VXK08)

作者简介:邱东,男,教授,博士生导师,经济学博士,教育部“长江学者”特聘教授,主要从事国民经济核算、经济测度、国际比较、中国统计改革与发展、多指标综合评价方法、可持续发展、宏观调节与区域发展等研究;郭同济,男,讲师,经济学博士,主要从事经济统计研究。

^①本文是《从方法提升到方法论:注重多指标综合评价的机理进路》^[1]的续篇,论文形成过程中与浙江工商大学苏为华教授有过深入的交流,交流主要内容以标明知识产权的方式纳入本文。苏为华教授和他的团队一直是国内研究合成指标方法论的主力,苏为华教授在2021年曾在科学出版社出版“国家哲学社会科学成果文库”《综合评价基本理论与前沿问题研究》^[2],该书应该以英文在海外出版,以显示中国学者在本领域突出的研究水平。很遗憾本文没能容纳该书的内容,也希望苏为华教授能撰写综述性论文,让国内同人对此领域的学术成果给予足够的重视。

Composite indicator 的中文命名,有的选用“复合”,有的选用“综合”,还有的选用“合成”。笔者选用“合成”。其理由在于,“复合”与“合成”都有英文 composite 的含义,但“复合”在中文里还有行为者(主体)之间关系修复之意,“合成”则没有这种歧义的可能。“综合”这个词使用范围比较广,作为专有名词的译名读者不大容易捕捉到英文 composite 的原意,而且在“价值指标”及其相关分析中,“综合”一词使用较多,composite indicator 正是与“价值指标”相对立的两大“可加性”方法之一,故而避开“综合”的译法,更有利于读者理解和把握指标方法的本意。

“机理进路”通常可译为 Mechanism approach,即刻意从方法的内在逻辑链接上深化和推进学科发展。但需要注意,这里的 Mechanism 不应该是“机械的机”,而是“有机的机”,社会经济现象属于“复杂有机现象”,如果理解上囿于“无机现象”的机械含意,则偏误颇多。笔者强调“经济统计意识”,此即一例。本文在第三部分从五个方面阐述了“方法机理”挖掘需要注意的内容。如果得到认同,则属于笔者独特的知识贡献。

Review of OECD/JRC's Handbook on Constructing Composite Indicators: Mechanism Approach Based on Indicators Methodology

QIU Dong, GUO Tongji

(School of Statistics and Data Science, Jiangxi University of Finance and Economics, Nanchang 330013, China)

Abstract: Attention to the “mechanism approach” of Composite Indicators Methodology serves as the foundation for avoiding measurement traps and achieving the unity of “mathematical reasoning” and “practical reasoning”. This paper provides a review of OECD/JRC's Handbook on Constructing Composite Indicators from a methodological comparison perspective, evaluating the strengths and limitations of the handbook by elaborating on the academic-theoretical standards underlying the Composite Indicators Methodology, and specific suggestions and reflections are proposed regarding the “Ten Steps” framework outlined in the handbook. Furthermore, it analyzes the root causes of the “Handbook Paradox” and emphasizes the mechanistic study of composite indicators constituting a crucial pathway to overcome this predicament. Chinese scholars should demonstrate intellectual confidence and self-reliance, developing an economic statistics discipline with Chinese autonomous knowledge systems, thereby playing an appropriate role in independently participating in and proactively shaping the formulation of international statistical standards.

Key words: multi-indicators composite evaluation; OECD/JRC Composite Indicators Handbook; economic statistics; methodological mechanism; handbook paradox

一、议题的提出

在社会经济统计中,人们往往需要得出可以比较大小或先后的“结果数据”,以便在不同事物之间做出选择和决策。而事物本身是多元的,对之的评价结果却需要是一元的,这便是笔者在《SFD 测度报告的经济统计学评论》^[3]中所概括的四大矛盾之一:评价的全面性与“结果数据”的单一指向性。也就是说,需要将所评价的多元事物做降维处理,将有关基础数据加乘汇总到一起,形成一个单一维度的“结果数据”。要解决“可加性问题”无非采用两种指标:一是价值指标,二是合成指标。

随着“经济测度的福利转向”,也由于价值指标本身的某些“测度陷阱”,合成指标在社会经济领域中逐渐变得非常流行,类似于“人类发展指数 HDI”这样的数据人们用起来非常顺手。不少学者还热心研发新的合成指标,俨然成为一个庞大的学科分支。然而,社会上流行着“偏数理轻事理”的倾向,似乎只要数学上可加,社会经济意义上便同样可加,人们不大注意某种合成指标的内在机理——其计算究竟为什么可行。即便作为国际标准的专门手册,方法“机理”的讲述也不充分、不透彻。而方法“机理”正是笔者在1986年开始就呼吁、倡导的研究重心,并由此形成了若干研究心得。

应苏为华教授邀请,笔者参加过他主办的两次专题国际研讨会。^①令人感到遗憾的是,第一次国际研讨会中当中国学者发言时,本来就不多的外国学者却全无踪影,这印证了笔者这些年来参加国际会议的一个感觉,欧美学者在与国人(当然也包括发展中国家的学者)进行学术交流时,大多只带着嘴巴,而没有带着耳朵和眼睛。他们压根儿没有从他者头脑里取得新思想——哪怕仅仅是学术信息的心理准备。

毋庸讳言,从总体上看,中国学者的经济统计研究水平低于发达国家的学者,乃至一些发展中国家的学者。仅21世纪以来,就有超过200部经济统计方法论研究的英文著作(含较长篇幅的“工作论文 working paper”)出版,的确需要我们认真学习,且“深度学习”。

然而学习并非迷信,总体落后也并不意味着在所有的经济统计学研究领域内,中国学者全都处于落后状态。例外是存在的:

在国际比较项目(International Comparison Program, ICP)机理研究上, ICP的“反常识数据结果”究竟应

^①第二次因为与教育部的会议冲突,笔者只提交了论文,即《从方法提升到方法论:注重多指标综合评价的机理进阶》^[1]。

该如何加以解释?在估算“基本类别 PPP”时,是否需要“纯价比假设”和“等价比假设”?这些质疑和思考难道不值得国际比较规则的制定者关注吗?

还有斯蒂格利茨教授主持的两份“经济测度报告”^[4-7],固然对当代经济统计学的学理进行了较为系统的剖析,体现了“经济测度的福利转向”,然而其中仍然隐含着不少“测度陷阱”,难道不需要进一步研究吗?“超越 GDP”已然实现了吗?

在多指标综合评价方法的研究上,中国学者的成果更是颇为可观,而且起步很早。从时间点上,笔者对合成指标方法论的系统研究从1986年开始,1987年向全国第一届中青年统计科学讨论会提交论文,《多指标综合评价方法的系统分析》1988年在《财经问题研究》(第9期)上发表^[8]。1990年笔者以合成指标为主题的经济学博士学位论文^[9]通过答辩,1991年由中国经济出版社出版^[10]。2000年苏为华教授同主题的经济学博士学位论文在厦门大学通过答辩^[11],2001年《多指标综合评价理论与方法研究》由中国物价出版社出版^[12]。比照一下,OECD/JRC 合成指标手册^[13]于2008年正式公布,^①比笔者的博士论文晚了至少17年(按出版时间算),比苏为华教授的博士学位论文也晚了7年。OECD/JRC 合成指标手册(2008版)并没有提及中国学者在此领域的研究成果,是这两篇博士学位论文等研究成果压根儿没有参考价值?^②还是基于“文明等级论”的一种学术态度?甚或出于“鲁滨逊视觉”(目无他人,唯余荒蛮)?苏为华教授认为,从理论或方法论上的研究水平看,上述两篇博士论文及其他众多中国学者的成果,在许多方面依然领先于OECD/JRC 合成指标手册(2008版本)。

2024年,笔者撰写了《从方法提升到方法论:注重多指标综合评价的机理进路》^[1],主要强调了两个方面:一是合成指标研究的系统性,再一是倡导合成指标方法论的“机理”研究,而非一味追求加大合成方法的技术难度。如果偏重方法机理的进路得到认可,那么笔者在1990年的研究不仅时间上领先,研究内容也更为丰富,至今仍然具有启示意义和推进作用,并不过时。基于这个判断,本文专门从方法论比较的角度对OECD/JRC 合成指标手册做一评论,作为前文的补充,主要包括以下内容:合成指标的方法论原理;注重方法“机理进路”;OECD/JRC 手册的结构;经济统计“手册悖境”。

二、合成指标的方法论原理

(一) 合成指标的功能

OECD/JRC 合成指标手册在“指标功能”上开宗明义,指出其关注与国家层面的产业竞争力、可持续发展、全球化和创新等方面,也即关注“国家显著事项”,这个表述界定了合成指标的层级和范围,可见其具有“宏观性”,是作为“国势学”乃至全球治理的定量指标工具,用以生产信息类公共品,是“社会基础结构”的重要组成部分。

需要注意的是,对经济统计学功能的理解存在着一个基本分歧,一派主张经济统计仅仅是对社会经济

^①苏为华教授介绍:据目前国外可考的文献,有关合成指标理论与方法的相对系统性成果,最早是2002年由欧盟的联合研究中心(JRC)Michaela Saisana 和 Stefano Tarantola 撰写的一份工作报告稿(State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development, EUR 20408 EN, European Commission-JRC: Italy)^[14],此后,该联合研究中心专家 Michela Nardo, Michaela Saisana, Andrea Saltelli 以及 Stefano Tarantola 等撰写了 Tools for Composite Indicators Building(EUR 21682 EN)工作报告^[15],随后 JRC 的 Michela Nardo, Michaela Saisana, Andrea Saltelli, Stefano Tarantola, Anders Hoffman, Enrico Giovannini 与 OECD 联合发布了 Handbook on Constructing Composite Indicators: Methodology and User Guide 等工作论文^[16]。而正式出版的则是2008年的合成指标手册^[13]。

^②广义地看合成指标(即不局限于经济统计学视域),苏为华教授介绍:中国国内自20世纪80年代以来就有一系列公开出版的关于合成指标理论方法研究及应用的论著,也有国内学者在国际期刊发表相关论文,甚至邀请过 JRC 专家来中国交流。

现实的“反映”，而另一派则主张经济统计应该是宏观经济管理的工具。两种功能主张使得统计指标的设计出现分歧，而对方法涵盖的范围也会有所差异。在合成指标领域，多目标决策方法在多大程度上纳入研究视野，就有不同见解。毕竟，就单一决策事项所采用的评价方法与旨在形成宏观信息系统的评价方法必定存在差异，需要操作者注意，在合成指标方法论研究中也需留心其影响。

（二）合成指标的定义和类型

OECD/JRC 手册是国际组织在本领域第一份国际标准，然而比较遗憾的是，该手册并没有给出“合成指标”的明确定义，方法原理讲述也还不够。

笔者于1990年将合成指标分成常规方法、^①多元统计分析和模糊综合评判三个大类^[9]。如果这个分类成立，OECD/JRC 手册竟然一直忽视了模糊评判，这是一个重大疏漏。笔者1990年博士学位论文就提出，“不确定现象”应该进一步区分为“模糊不确定性现象”和“随机不确定性现象”，这个基本区分意味着模糊综合评判方法存在和使用的法理性。

还有，OECD/JRC 手册明确提出，其内容不包括用于描述经济周期的“合成领先指标（Composite Leading Indicators, CLI）”，原因在于该类合成指标基于统计和计量经济学，与其他类型的合成指标差异相当大。然而，OECD/JRC 手册并没有展开此论点，我们需要明确，这类合成指标与其他合成指标究竟相差在哪里。而令人不解的是，在论述数据标准化方法时，^②OECD/JRC 手册还涉及领先指标，前后表述并不一致。

笔者认为，领先指标和其他合成指标就“合成”而言在本质上相同，只是其测度重心更偏于时间维度。苏为华教授对此做了深入剖析：领先合成指标 CLI 的计算，要求入选的指标必须是“先行于基准周期”的指标，通过计算时差相关系数或 K-L 信息量等方法构建“先行景气指标体系”，且先行时长或峰谷周期性比较接近，这正是“合成指标”理论所强调的“内部一致性”要求。^③虽然 CLI 的单项指标计算需要经过一个“相对复杂的因素剔除”（长期趋势、季节变动及不规则变动的剔除），但基于各指标的循环变动序列计算“增长率”，并对它们进行加权集成（通常是算术平均），这正是“常规方法”中最常用的“综合指数法”。而“相对复杂的因素剔除”正是合成指标计算中单项指标的同度量内在要求。从苏为华教授的剖析中可以看出：领先指标的计算与一般的多指标综合评价在操作上虽有不同，但从合成指标方法论角度看，并不足以将其排除在外。其实，领先指标系列在方法论上对其他类型的合成指标颇有启示，OECD 国家，特别是美国，在领先指标编制和使用过程中的经验和教训也值得认真总结并加以借鉴。

苏为华教授进一步指出：经济景气分析中，还有与 CLI 并存的同步合成指标（指数）、滞后合成指标（指数），以及与 CLI 平行的基于景气问卷调查的扩散指数（Diffusion Index, DI）都是合成指标方法论的具体应用，因为“无量纲化—加权—平均集成”都是不可或缺的步骤。从这个视角看，各类景气合成类指标的设计、计算与解读，特别需要接受合成指标方法论的指导。当前对诸如采购经理人指数（Purchasing Managers' Index, PMI）点数波动的过度关注，而轻视其合成计算的“简单粗暴”（只分“上升，下降，持平”三个方向，而忽视其程度），就是很值得商榷的。因此，苏为华教授在山东临沂商城景气指数、浙江舟山水产景气指数、永康五金景气指数等多个景气指数的设计中，就改成了5—7级模糊评价方法处理，而不是传统的3级处理。

（三）合成指标的“内部一致性”

OECD/JRC 手册第1.6节中指出，多元统计方法（主成分分析和因子分析）“可以根据相关性来对变量

^①所谓“常规方法”，即采用相对数和平均数进行多指标综合评价，非常典型的“政治算术”，邱东1990年博士学位论文从中概括出所有合成指标的“计算性质”：“统计相对数的加权平均”，多元统计分析用于多指标综合评价、模糊综合评判从计算的本质看都是其拓展。

^②见 OECD/JRC 手册第1.5节^[13]。

^③见本文第2.3节。

进行分组,然而,如果指标间没有相关性,则权重不能用上述方法来估计”。这段论述实际上指出了合成指标方法选择所面临的一个悖境,因为在变量选择时,为了减少各构成指标(子指标)的信息重叠,应该尽可能降低其间的相关性;而指标间的“低相关性”却又排斥多元统计方法的使用。除了加权方法的选择以外,在合成汇总方法的选择上也面临此种两难,值得深入思考。

OECD/JRC 合成指标手册的前言用表格列示了合成指标的六方面缺点。其中只有“指标和权重选择”和“忽略难以测度的维度”(其实这与前一条相同)这两条是方法本身的缺点;另有三条(解释不合理、被滥用、执行过程不透明)是方法使用不当造成的问题,并非方法的内在问题;“可能导致过于简单的政策结论”这条,则应该归结于内外两方面原因。当我们对方法优缺点做出评判时,恐怕更多的应该是注重其“内部一致性”。

OECD/JRC 合成指标手册文本中有的论述还不够严谨,作为国际标准,这恐怕是不应该出现的问题。例如,手册目录和第一部分第二节讲“选择变量”,而引言和表1中讲的则是“数据筛选”,二者虽然紧密关联,但还不能视为同一步骤。

苏为华教授指出,OECD/JRC 合成指标手册的专家工作小组 CC-COIN(Competence Centre on Composite Indicators and Scoreboards)最近几年组织了培训会,但在会上授课专家并没有完全根据 OECD/JRC 合成指标手册(2008版本)的步骤流程、内容方法来设计课程内容,从操作流程细节及其顺序到方法选择,都有不少变更。这表明,手册作为国际标准的稳定性是有一定欠缺的,用于指导与审核(Audit)合成指标编制实践,其权威性是可以被挑战的。

(四) 合成指标的系统观

笔者1990年在博士学位论文中提出,从系统角度看,“指标方法”无非“实物指标”“价值指标”“合成指标”和“指标体系”四种^[9]。合成方法属于“加总派”观点的两大分支之一,不过它完全否定了“价格”作为“同度量因素”和“权数”的稳健性,从而否定了价值指标计算的法理性,转而主张采取“去量纲”以求“同一性”从而达成“可加性”的合成方法,从而与价值指标方法并列。

经济测度中还有人认为,宏观指标根本无法达成其“可加性”,故而形成了“非加总派”的主张。然而,现实社会决策(对策)存在“决断信息”的需求,对一个问题,总是希望得到 yes or no 的单一结果数据,故而“加总派”观点仍然流行于世。

我们阐述合成指标方法,应该从系统的角度把握其历史地位,这样才能更好地使用该方法及其“结果数据”。

OECD/JRC 手册第一部分第一节论述合成指标构建的第1步,即“构建理论框架”。可见,指标合成处理这个定量过程也须从“定性分析”开始。有的人喜欢搞基于所谓前沿数理模型的“纯定量分析”,似乎文字表述越少,公式越复杂,就越高级。如果总是停留在抽象的数理空间,“纯粹”或许还可以,但如果搞应用并落实到现实社会中,则必须有一个与形成模型时抽象过程相反的逆过程——具象化,始终坚持抽象处理只能是将所谓定量分析囿于乌托邦。

需要做出区分的是,这里的“理论框架”是针对具体测度项目中特定测度对象而言的,并不是合成指标方法构建的一般理论,就是说,这部分的论述并没有从指标方法的系统角度来阐述合成指标的性质。实施具体测度项目,是其“必要性”和“可行性”二者相凑的过程,创新开发多指标综合评价方法,尽可能达成经济测度的相关性(或社会经济意义)。

OECD/JRC 手册第一部分第一节(1.1)中合成指标的发展论述以 GDP 测度为参照,并认为 GDP“理论和统计框架在过去的50年内逐步形成”,具备了“坚实的理论和经验基础”。在方法成熟度比较上需要注意的是,这个评价只能是就“价值指标”相对于“合成指标”而言,而且应该看到,正是因为价值指标作为一种“多维指标”尚存在“测度陷阱”,才有了开发“合成指标”的必要性。需要明确的是,合成指标方法的必要性并不只是因为测度领域的拓展,还出于已有价值指标方法需要类型改革和拓展的考虑。

(五) 合成指标处理的透明度要求

关于“透明度”,OECD/JRC 手册有这样一段论述:“在一些新兴地区,由于相关领域的经济研究尚处于发展阶段,综合指标(如竞争力、可持续发展、电子商务等)可能是非常主观的。”这里需要注意的是,将发展阶段与指标“主观性”直接挂钩,恐怕容易误导对经济测度原理的认知,似乎只要经济发达了,合成指标的“主观性”就消失了。其实,经济测度是由具有“主观性”的人来操作的,哪怕相关人员竭力追求测度的“客观性”,也无法完全消除“主观性”。手册引言中专门提及“合目的性原则(a fitness-for-purpose principle)”,等于公开认可了这一原则的“非中性”。

发达地区的测度操作或许更为“客观”,但也难免存在,甚至强化将发达国家的测度经验作为全球标准的风险,如隐含在“事实”背后的“见解”,以“客观”面目出现的“主观”认知,故而,为了在承认主观性的前提下尽可能追求测度结果的客观性,测度方法“透明度”应该是一般性的要求,正如 OECD/JRC 手册所言,“透明性必须是贯穿整个工作过程的指导原则”,故而,并不只是发展中地区才需要此方面的考量。

同时也应该注意到,合成规则和方法的透明度并不能保证编制具体合成指标时的操作透明度,规则和方法是否得到切实的贯彻,是需要严格审视的,其过程也需要公开。例如,“国际比较项目(ICP)”规则和方法是公开的,然而“购买力平价(Purchase Power Parity, PPP)”的计算过程能否公开还值得深究。还有,OECD/JRC 手册在引言中提出,合成指标应该“能够被分解成其子指标或具体值”,即合成的“逆操作”,这是一个颇高的要求,不知在社会实践过程中多少合成指标能做到这一步。

三、注重方法“机理进路”

单从数学工具的难易度和解读方式来看,OECD/JRC 合成指标手册似乎有些“落伍”和“过时”。出于提供使用指导和推广的目的,OECD/JRC 手册对所使用方法做了用户友好型的解读,这是其优点。需要注意区别的是,方法的通俗解说,偏重于操作层面讲述如何做(How questions),还不同于数理方法用于指标合成的机理剖析(为什么该方法可以用于该项多指标综合评价?上升到学理层面,Why questions)。而且合成技术又有所发展,毕竟十四年过去了,OECD/JRC 手册(2008版)的论述可能显得“陈旧”,恐怕难入偏好“技术路径(Technical Approach)”研究者的法眼。其实,方法论研究的成果是不会过时的。任何学科,学说史,特别是学说思想的梳理,总有其珍贵的学理价值。

然而,如果从社会科学应用的方法论提升这一视角看,偏重于“机理路径”,改变当下全球经济统计“手册(Handbook)”多、“脑书(Headbook)”少这种偏重于初级层次的格局,那么,该“手册”可以作为进一步撰写合成指标“脑书”的基础。所谓“脑书”,其深意正在偏重于指标的方法机理,偏重于指标方法的“内在一致性”和“外在一致性”。

苏为华教授认为:也正是从方法论的学理层面标准看,OECD/JRC 手册(2008版)还是很显不足的,无论是常规的“粗浅”方法,还是其 CC-COIN 家特别推崇的如 BoD(Benefit of the Doubt)方法^[17],以及新引入的一些多属性决策方法(Multi-Attribute Decision Making, MADM),远不及国内主要学者数十年前对“方法机理”的研究。

概而言之,“方法机理”至少需要关注以下几个方面:

第一,为什么要开展某个方面的测度?为什么要开发该指标?即指标方法的社会经济意义,或指标方法的时空相关性。在具体的合成指标编制过程中,这主要通过 OECD/JRC 手册的第一步骤来体现。

第二,在阐述指标方法“如何测度”问题(How questions)的基础上,进一步深入阐述“为什么可以那么做”的问题(Why questions)。苏为华教授认为:这是“数理”与“事理”的统一与相洽问题。数理必须从属于事理。学界非常普遍的“偏数理轻事理”倾向,同样存在于合成指标的构建活动中。这需要经济统计学界对指标方法的数理性作出如何合乎事理需求的深度剖析,这是响应“Why questions”的不可或缺的一步。笔者注意到,OECD/JRC 手册在第1.4节中明确指出,聚类分析的用处之一是作为“一种纯粹的、用于指标汇

总的统计方法”,然而,聚类分析究竟如何“去量纲”并“合成”,却语焉不详,这在方法机理上恐怕难以服人。笔者1990年博士学位论文^[9]对主成分分析(Principal Component Analysis)和因子分析(Factor Analysis)如何用于多指标综合评价,做了较为系统的说明,是对多元统计方法用途的一种鉴别。

第三,指标方法不可能尽善尽美,总是需要加以改进,而改进的动力之一便是其所隐含的“测度陷阱”,因此仅仅从正面研究测度缘由还不够,还需要从反面审视,指标方法是否存在且在哪里存在“测度陷阱”,即“为什么不能那么做”问题(Why not questions)。笔者1990年在博士学位论文中,对聚类分析无法独立地用作合成指标构建做了鉴别性说明,便属于这方面的研究。苏为华教授认为:“测度陷阱”的本质是指标方法论中“数理”与“事理”的“隐性脱节”。因此,(合成)指标方法的改进,必须遵循一些基本准则。2023年11月4日在第22次全国统计科学讨论会上,苏为华教授做了大会报告《综合评价技术的基本准则、方法逻辑与发展方向》^[18],对之在2021年“国家哲学社会科学成果文库”著作的基础上做了进一步阐述。

第四,任何指标方法都需要某些假设,提升测度品质不是不要假设,而是不能将计算(包含假设的)所得的“结果数据”绝对化,需要审视假设,特别是隐含的假设,对指标方法应用效果的影响。OECD/JRC 合成指标手册在构建方法格局中特别注意“后期处理”,设置了两个专门步骤,注重指标方法的“内在一致性”,这就提升了“指标方法论(Indicator Methodology)”研究的层次,是该手册的一个突出优势。在 OECD/JRC 手册第4.1节中,OECD/JRC 合成指标手册还设置了专栏,阐述“主成分分析”的假设,表现出手册编制者的睿智。需要进一步推进的是,如何将多元统计的一般方法论与其在合成指标编制的特殊应用相结合,即特殊方法论的拓展。

第五,指标方法的与时俱进。指标方法是为了动态地描述“国家显著事项”,描述对象自身会随着历史进程发生演变,描述指标当然需要与时俱进。从方法机理看,就是被测度对象是否发生了演变,发生了什么样的演变?演变对描述指标方法的机遇和挑战有哪些?如何准备测度条件以改革指标方法从而适应国势研判的要求?这体现了指标方法映射被评判现象时的“外在一致性”^①。

四、OECD/JRC 手册的结构

OECD/JRC 手册将合成指标的构建分为十个步骤:(1)理论框架;(2)数据筛选;(3)缺失数据插补;(4)多元分析;(5)标准化;(6)加权和汇总;(7)不确定性和灵敏度分析;(8)数据示例;(9)与其他指标的联系;(10)结果可视化。

对合成指标手册“十个构建步骤”的安排,笔者有这样几点意见和思考:^②

第一,“多元分析”作为一个单独步骤似有不妥,多元分析是数理统计的一般方法,并非合成指标方法的一个专有步骤,它在其他几个步骤中都需要使用,在合成指标的构建中发挥着重大作用,二者不宜混淆。作为 OECD/JRC 手册的第4步,主要是用来分析“数据结构”,其实多元分析方法在第5步和第6步中的作用也非常重要,手册将之放在第4步,某种意义上具有“方法准备”的意味。在 OECD/JRC 手册介绍合成指标和构建过程时,如果从内容衔接角度看,第4步应该直言为“数据结构分析”,与“选择变量”(第2步,含“数据筛选”)、“数据插补”一起作为“数据准备”板块的内容。^③而且,变量选择、数据筛选与数据结构分析、数据插补这几个步骤应该形成一个彼此互动的反复过程。

第二,“权数”在合成指标构建过程中,应该是一个“全过程概念”:其实从合成的第1步骤(即构建理论

^①指标方法的“外在一致性”还涉及与其他学科方法的关系。

^②笔者在《多指标综合评价:方法论反思》中对三种合成指标构造的步骤安排(分别由邱东、苏为华和 OECD 提供)做了比较,侧重于其共性^[19]。

^③第3步,数据结构分析或许应该安排在“插补”之前。

框架的定性分析)开始,到指数构建完成,权数的意味都存在着。选择哪个指标,这是“第一次加权”^①,未入选的指标其权数为零。缺失数据“补插”,意味着该项数据的权数不应该为零。标准化方法不同,其中隐含的权重也不同。汇总步骤时即便做“等权”处理,权数仍然存在,只不过数值都为1而已。如果接受权数作为“全过程概念”,那么将“加权”论述与“汇总”放在一起作为第6步,概念理解上就失之过窄。合成时有时需要专门加权处理,而有时合成方法本身就具有了加权处理的意味,并不需要再专门加权,故而在第6步中分开两部分论述,也有方法理解上的误导之嫌。

第三,“数据实例”(Back to the Details,第8步骤)也不应该是一个独立的“构建步骤”,似乎更适合并入(放在)“展示与推广”(第10步)中。OECD/JRC手册采取了“一例到底”的讲解方式,全书都用“技术成就指数(Technology Achievement Index, TAI)”为例,这恐怕是出于“教学目的”的考虑,一种撰写教科书的手法。其实,合成指标方法应该有不同类型,不可能存在适合于所有合成指标方法的典型事例。如果方法讲解可以做到“一例到底”,恰恰证明OECD/JRC手册的内容体系还存在方法缺欠。

第四,如果略作调整,第5步(数据的标准化)和第6步(加权和汇总)仍然是合成指标构建的“核心”板块,^②其后便是指标构建的“后期处理”板块,第7步(稳健性和灵敏度分析)主要关注合成指标的“内部一致性”,第9步(与其他指标的联系)则主要关注合成指标的“外部一致性”。需要注意的是,后续这些步骤还应该与前面的“数据结构”分析相衔接,不同的数据结构与不同的标准化方法、加权方法和汇总方法、稳健性分析、灵敏度分析相对应。这样,从“理论框架”到“数据准备”,到“核心操作”(非此不足以作为“合成指标”)到“后期处理”,构建方法各板块和步骤之间的逻辑顺序关系比较清晰。

第五,OECD/JRC手册分成两大部分,先讲构建合成指标,再讲指标构建的工具箱,编制步骤几乎讲了三遍(引言中还有提及)。笔者以为,OECD/JRC手册的行文未必需要做两大部分安排,或许将各步骤内容集中讲述,而第一部分第二章的“合成指标的质量框架”则放在最后,这样安排或许使方法的“机理进路”更为清晰,解读效果更佳。

苏为华教授指出:诚如其名,OECD/JRC手册是一份以操作流程为主线的工作指导文件,突出了合成指标构造的全程性,而忽视了合成指标理论方法的系统性,因此从严格意义上讲都不能够称为“理论与方法研究”的系统性著作。因为,异常值的识别处理与数据的插补、合成指标的深度分析、可视化与传播等,是任何一项完整统计研究活动(数据分析)的必要环节,并非合成指标构造的特有内容——至少手册编撰者没有立足合成指标分析的特殊性来探讨这些环节的内容。换言之,其中所述的“数理”及“事理”内容都是一般统计活动的共性部分,而非合成指标方法的个性部分。

五、经济统计“手册悖境”

一般而言,经济统计手册作为国际标准,推广应用过程中往往需要树立其威信,这样就需要为尊者讳,手册编撰者往往不愿意或不能(或囿于专业偏好)充分揭示指标方法之不足。然而,即便成为国际标准也不能固化,而应该与时俱进,当统计对象本身发生演变时,标准就需要修订,而修订的基础便是指标方法的机理探讨。经济统计手册主要阐述方法如何使用,有的手册也会就方法机理做一些基础性或概要性的阐述,不过往往不那么充分。是否深入探讨方法存在的问题,这是考验“手册”“可持续质量”的关键,也是手册编撰者面临的一种两难选择:究竟是偏好于短期推广,还是长期信誉?

应该看到,SNA从1993年版修订开始,就将修订过程中需要深入研究的国民核算方法论课题展示出来,不同的方法论见解显示出国民核算领域的认知水平和矛盾、核算问题的焦点和体系优化的方向,这些构成了SNA方法论研究的“金矿”(研究者正应该如同恩格斯先生所言:“从已有的思想材料出发”),从而

^①笔者在《多指标综合评价:方法论反思》中首次提出这个概念^[19]。

^②笔者在《多指标综合评价:方法论反思》中提出了合成方法中预处理、核心处理和后期处理三个阶段的划分方法^[19]。

成为下一轮核算体系修订的专业文化基础。

反观 OECD/JRC 合成指标手册,标准文本中缺少像 SNA 那样对未决问题坦白的表述。合成指标手册“引言”中也曾指出:“作为一个‘活’的产品,只要新的方法被提出,本手册就会不断更新发布新的版本。”然而从2008年至今,16年过去了,还没见到合成指标手册修订的新版本。似乎这些年并没有新的合成方法被提出,恐怕与本领域的研究态势不符,其实反倒表明了该手册的成长性比较弱。提升其成长性,一个重要指向便是注重合成指标的机理研究。

与此形成鲜明对照的是,国际社会涌现了各种各样的合成指标,据苏为华教授介绍,早些年就已经多达4000多种。这些新的合成指标是本领域最为鲜活的研究材料,如果系统地比照各种各样合成指标的编制和应用情况,充分地分析其相同点、不同点、优点和缺点,^①就容易发现合成指标方法的发展态势、存在的“测度陷阱”,就为方法改进和标准修订找到了新题材。

其实,OECD 和 JRC 在2008年编制合成指标方法手册时,并没有注意到发展中国家学者对合成指标编制议题的研究,没有注意到后发国家学者在该专题研究上的知识贡献。将笔者早年的合成指标研究与17年后的 OECD/JRC 合成指标手册对比,就编制原理的剖析(遵循方法的“机理进路”)而言,恐怕具有明显的差异,何者为优,值得深入比较。2014年笔者在第四届中国统计学年会上做了特邀报告,题目是《多指标综合评价:方法论反思》^②,再次就合成指标的方法论机理做了探讨。还需要强调的是,苏为华教授本人和他的团队在本领域长期深耕,取得了系列研究成果,也在海外发表了若干(英文)论文,颇值得重视。

通常,发达国家的学者是国际统计规则的制定者,而发展中国家的学者是这些规则的学习者、推广者,但这并不意味着后者没有专业思考的能力,没有质疑国际标准的“资质”,应该打破那种发展中国家人士在专业学术发展上一无是处的刻板印象。其实在20世纪80年代的中国,因为“宏观经济效益统计”的兴起,不能再沿用现成的“产值(价值)指标”,导致中国学者在合成指标研究方面投入了大量的精力深入思考,也发表了较多论文。笔者和苏为华教授的合成指标方法论研究,便建立在这个研究热潮的基础之上。

这个事例引发我们进一步思考:OECD/JRC 手册编制的参考文献是否比较齐全?有哪些重大遗漏?在已经收集到的文献中,所涉及的方法论问题是否得到了充分的讨论?还有哪些争议问题遗留了下来?合成指标方法论研究的主要方向究竟有哪些?OECD/JRC 手册引言最后一节为“后续内容(What's next)”,其线索是否偏于“技术进路”?是否需要拓展?

OECD/JRC 手册比较重视从“社会选择”和“多标准决策分析”中吸取经验,这种拓展非常好。然而也需要注意,就定距尺度和定比尺度的信息比较而言,定序尺度的规则和经验对其适用性究竟如何。我们知道,绝对值信息损失可能对不同时期排序变化产生影响,例如,A 在 t 期仅以微弱的优势排在 B 前面,到了 $t+1$ 期,A 与 B 易位的可能性就相对大一些。还有,排名位次在不同类型的博弈中其社会经济意义不同。例如在充分竞争的博弈中,排名靠前或许还有意义。但在零和博弈中,排名第一就是要全部占有,就是要保持绝对优势,而排名稍后者与排名非常落后者在利益享有上相差并不大,但在成本付出上却可能差异相当大,遭遇的恶意排斥也非常多。再有,大学排名顶多具有高校层级划分的粗略意义:某所大学在所有高校中大概处于哪个层级(档次),至于哪所大学排在哪所大学的前面或后面,具体名次上其实具有不确定性,未必令人信服。从这些事例中可以看出,合成指标手册只有比较系统地总结其构建经验,从中探索方法论机理,才能形成多指标综合评价的良性发展。

苏为华教授认为:中国经济统计界,在包括合成指标等在内的多个研究领域,都有自己独到的学术见解与理论成果,且有不少立足于中国实践的优秀应用性经济统计成果,其水平丝毫不逊于甚至胜于国外同类研究成果。苏为华教授呼吁:我们应该自信自强,通过发展具有中国自主知识体系的经济统计学科,在自主参与甚至积极引导国际统计规则的制定方面,发挥应有的作用。

^①即笔者1900年经济学博士学位论文所提出的“四点观”^[9]。

^②收入笔者《经济统计学科论》,中国财政经济出版社,2013年版^[20]。

笔者2021年4月21日在《光明日报》理论版发表了《新时代中国特色统计学问题研究与国际统计标准的中国参与》^[21],希望能为打破此种“手册悖境”提供思路。

参考文献:

- [1] 邱东. 从方法提升到方法论: 注重多指标综合评价的机理进阶[R]. 贵阳: 2024年第18届全国经济统计学年会大会报告.
- [2] 苏为华. 综合评价基本理论与前沿问题研究[M]. 北京: 科学出版社, 2021.
- [3] 邱东. SFD 测度报告的经济统计学评论[R]. 北京: 2021年清华大学经济社会统计问题前沿问题研究系列论坛报告.
- [4] 华汉阳, 朱启贵. 长三角现代化产业体系建设水平的测度研究[J]. 统计与信息论坛, 2024(12): 15-28.
- [5] 孙华平, 郭子晴, 孙梅. 偏向性技术进步提升了中国城市碳生产率吗? ——基于碳中和与稳增长的实现路径研究[J]. 济南大学学报(社会科学版), 2025(1): 106-123, 192.
- [6] STIGLITZ J E, FITOUSSI J P, DURAND M. Beyond GDP: measuring what counts for economic and social performance [M]. Paris: OECD Publishing, 2018.
- [7] STIGLITZ J E, FITOUSSI J P, DURAND M. For good measure: advancing research on well-being metrics beyond GDP [M]. Paris: OECD Publishing, 2018.
- [8] 邱东. 多指标综合评价方法的系统分析[J]. 财经问题研究, 1988(9): 51-57.
- [9] 邱东. 多指标综合评价方法的系统分析[D]. 大连: 东北财经大学统计学院, 1990.
- [10] 邱东. 多指标综合评价方法的系统分析[M]. 北京: 中国统计出版社, 1991.
- [11] 苏为华. 多指标综合评价理论与方法问题研究[D]. 厦门: 厦门大学经济学院, 2000.
- [12] 苏为华. 多指标综合评价理论与方法研究[M]. 北京: 中国物价出版社, 2001.
- [13] OECD, EUROPEAN UNION, EC-JRC. Handbook on constructing composite indicators: methodology and user guide [M]. Paris: OECD Publishing, 2008.
- [14] SAISANA M, TARANTOLA S. State-of-the-art report on current methodologies and practices for composite indicator development [R]. Ispra: European Commission, Joint Research Centre, 2002.
- [15] NARDO M, SAISANA M, SALTELLI A, et al. Tools for composite indicators building [J]. European Commission, Ispra, 2005, 15(1): 19-20.
- [16] NARDO M, SAISANA M, SALTELLI A, et al. Handbook on constructing composite indicators: methodology and user guide [M]. Paris: OECD Publishing, 2005.
- [17] 陈思博. 基于 BoD 多指标综合评价方法的改进研究[D]. 杭州: 浙江工商大学统计与数学学院, 2022.
- [18] 苏为华. 综合评价技术的基本准则、方法逻辑与发展方向[4][R]. 成都: 2023年第22次全国统计科学讨论会报告.
- [19] 邱东. 多指标综合评价: 方法论反思[J]. 经济统计学(季刊), 2013(1): 39-47.
- [20] 邱东. 经济统计学学科论[M]. 北京: 中国财政经济出版社, 2013.
- [21] 邱东. 新时代中国特色统计学问题研究与国际统计标准的中国参与[N]. 光明日报, 2021-04-21 (11).



(责任编辑 郭宝才 王 权)