

## 清廉指数的腐败测评方法与局限性

过勇 宋伟

**内容提要:** 清廉指数是一种基于主观问卷调查得出的腐败测评结果,反映的是人们对一个国家和地区腐败程度的印象和感知。清廉指数的发布引起了国际社会对于反腐败议题的关注,推动了世界各国的反腐败行动。然而,清廉指数也存在着很多局限性,它作为一种主观调查评价方法,通过调查人们对于腐败的印象和感知来评价一个国家和地区的腐败程度,这种方法本身就具有不准确性和非客观性。尽管透明国际对清廉指数的数据来源和计算方法进行了不断的完善和优化,在2012年还对计算方法进行了改革,但是清廉指数在科学性和解释力方面仍然存在着很多不足和缺陷。此外,清廉指数对中国的测评存在着不公正对待,其调查样本和测评结果都不能够充分反映出中国腐败和反腐败情况的客观事实。

**关键词:** 清廉指数 腐败测评 局限性 公共治理

**中图分类号:** D73 **文献标识码:** A **文章编号:** 1003-3947(2013)05-0151-10

腐败是对政府合法性的严峻挑战,也是对全球经济发展的严重破坏。因此,一个国家的腐败状况往往是影响经济发展速度和外国直接投资水平的关键因素,受到各国政府、企业以及商业咨询机构的高度关注。在我们探讨腐败现象的过程中,一个基本的命题就是腐败测评。如何定量评价一个国家和地区在不同时期的腐败差异和变化情况?这不仅是探讨一些理论问题的基础,也是制定相关公共政策的重要立足点。然而,由于在腐败研究中存在隐秘性信息不对称和敏感性信息不对称的问题(Guo 2008),我们往往难以对腐败的真实发生发展情况做出准确的评价,这是腐败问题研究的最大难点。

本文简要讨论了腐败定量评价的重要理论意义和现实价值,对国际上现有的腐败测评方法和成果进行了介绍,并以影响力最大的清廉指数为例,对其方法和局限性进行了详细分析。我们的研究发现,尽管清廉指数对于推动各国的反腐败行动起到了一定的积极作用,但是作为一种腐败测评方法却存在着很多缺陷,评价结果具有较强的局限性,特别是对中国评价的不公正对待应引起我们的关注。

---

**作者简介:** 过勇,清华大学公共管理学院副研究员、廉政与治理研究中心副主任。宋伟,清华大学公共管理学院博士后、廉政与治理研究中心助理研究员。

**基金项目:** 教育部人文社会科学研究规划基金项目“我国廉政测评体系研究”(项目编号:13YJAZH029);清华大学自主科研计划“机构、法律和战略综合视阈的反腐倡廉体系比较与实证研究”(项目编号:20121087989)。

限于篇幅,部分参考文献略,有兴趣的读者请与编辑部联系。——编者注

## 一、腐败测评方法与文献综述

目前,社会上存在着很多认识腐败的误区,比如“腐败是经济增长的润滑剂”、“腐败越反越多”及“民主是治愈腐败的良药”等。这些似是而非的观点被广为传播,但是往往缺乏学术研究的支撑。

近年来,一些国际组织和商业咨询机构在测评各国腐败程度方面进行了大量的调查研究工作。胡鞍钢和过勇介绍了四个机构开发的五种常用指标,包括透明国际的清廉指数和行贿指数;世界银行的腐败控制指标;瑞士国际管理发展学院发布的《世界竞争力年鉴》中相关指标;世界经济论坛发布的《全球竞争力报告》中相关指标。此外,哥伦比亚大学的《国家能力调查》、经济学家智库发布的《国家风险服务和国家预测》、商人国际集团发布的《灰色地带动态评级》、香港政治和经济风险咨询机构发布的《亚洲情报通讯》等都有较为广泛的影响力,被作为评价各国腐败状况的重要依据(胡鞍钢、过勇 2001)。

在所有这些腐败测评指数中,清廉指数影响力最大,使用最为广泛。过勇对清廉指数的评价方法和结果、学术意义以及局限性做了讨论(过勇 2005)。宋旭光对清廉指数进行了统计学批判,认为清廉指数无论在统计主体的时间、空间属性上,还是统计技术方法本身都存在着很大的问题(宋旭光 2005)。兰斯多夫提出,尽管感知不能和现实相混淆,但是从已有经验来看,清廉指数和现实是一致的,清廉指数调查中获得感知是腐败实际水平的反映(Lambsdorff 2006)。伊恩认为,透明国际的清廉指数是目前可查的对腐败进行国别比较的最有价值的数据库(Senior Ian 2006)。过勇提出,国内外学术界关于腐败问题的实证研究中主要采用三类方法,除了基于这些调查进行评价的主观调查法之外,还有案件统计法和案件统计分析方法。主观调查法又可以具体分为两类,包括问卷调查法和综合评价法,后者的主要代表有清廉指数和世界银行的治理指数(过勇 2008)。斯蒂芬和保罗认为,清廉指数在概念和方法上都存在缺陷,而它的影响力可能会放大这种缺陷,造成清廉指数的滥用(Staffan Andersson & Paul M. Heywood 2009)。徐静对国内外腐败指数进行了比较研究,认为清廉指数在定义和方法论上都存在着一定的偏差性(徐静 2012)。

由于清廉指数的广泛影响力,本文首先对该指数发展和评价结果做一个简要介绍。从2009年开始,为了应对广泛的批评,透明国际成立了一个由14名专家组成的改革小组,本文作者之一过勇作为中国大陆地区唯一的代表参与整个过程。因此,本文也将对改革后的清廉指数方法进行介绍和讨论。

## 二、清廉指数的发展和评价结果

清廉指数是透明国际发布的评价各国和地区腐败状况的重要指数。透明国际(Transparency International)成立于1993年,总部位于德国柏林,是一个专门致力于反腐败的国际性非政府组织,截至2011年底已经在全球114个国家和地区建立了分会或会员组织。透明国际在成立之初,为了推动全球反腐败行动和计划,决定从1995年开始于每年秋季发布清廉指数(Corruption Perceptions Index)。由于德国教授约翰纳·伯爵·兰斯多夫首先提出了清廉指数的算法,因此他被称为是“清廉指数之父”。清廉指数的发布不

仅引起了人们对于腐败评价的极大关注,也提高了透明国际的知名度。透明国际认为:“清廉指数是透明国际最知名的工具,它已经得到了广泛认同,同时也把透明国际和腐败议题提上了国际性议程”(Transparency International 2005)。

清廉指数也可被译为“腐败感知指数”或“腐败印象指数”,顾名思义它是一种基于主观问卷调查得出的腐败测评结果。清廉指数被认为是“民意中的民意”,是一个反映人们感知腐败程度的主观测评体系。清廉指数所使用的测评数据并不是透明国际开展调查所获得的第一手数据,而是综合了多份有关调查报告经过数据处理而得到的。透明国际参考的报告大多都以一般民众、商人、官员和专家学者等为主,目的是尽可能地反映腐败测评的民意。由于透明国际参考的各个报告覆盖的国家并不相同,最初一般一个国家至少需要被三个报告调查过才能列入当年的清廉指数测评。近年来,随着清廉指数影响力的增加,被列入到清廉指数测评的国家和地区也逐渐增多,从1995年的41个增加到2011年的183个,2012年为176个。

清廉指数以评分和排名的形式展示测评结果,1995年至2011年采取的是十分制,2012年开始采取百分制。这两种计分方式都是得分越高,表示该国家和地区腐败程度越低。我们以十分制为例,可以将被评国家和地区的腐败程度分为四种类型:比较廉洁、轻微腐败、比较腐败、非常腐败,如表1所示。位于清廉指数排行榜前列的往往总是北欧国家,亚洲的新加坡和中国香港也属于这一档。他们分别代表着国家廉政体系的“北欧模式”和“亚洲模式”。轻微腐败的多是制度建设相对完善的发达国家和地区,也有一部分发展中国家(地区)和转型国家(地区)在前期的改革和发展中先行一步,进入了这一行列,如波兰、韩国和中国台湾等。比较腐败的多是那些正处在发展或转型之中的国家和地区,而非常腐败的几乎都是严重欠发达或转型陷入陷阱的国家和地区。中国(大陆地区)的得分如图1所示。

表1:国家和地区清廉指数得分的划分

类型	得分区间	主要代表国家和地区
比较廉洁	8.0~10.0	丹麦、瑞典、芬兰、挪威、新西兰、加拿大、新加坡、中国香港
轻微腐败	5.0~8.0	美国、英国、日本、法国、葡萄牙、捷克、波兰、韩国、中国台湾
比较腐败	2.5~5.0	阿根廷、泰国、墨西哥、巴西、中国、印度、赞比亚、埃及、越南、蒙古
非常腐败	0~2.5	俄罗斯、孟加拉国、肯尼亚、伊拉克、缅甸、印尼、阿富汗、朝鲜

清廉指数作为一个测评腐败的工具,尽管存在着很多缺陷,但是其对推动全球反腐败斗争的重要意义是非常显著的。透明国际在2006年的报告中曾指出:清廉指数得分低的国家应当作出重大改革,清廉指数为商人和投资者提供了重要信息(Transparency International 2006)。事实上,清廉指数的确引起了政府、商人和媒体,乃至普通民众的关注,政治家甚至以提高清廉指数为执政目标(Staffan & Paul 2005),商人以清廉指数得分情况来决定是否对一个国家和地区进行投资。正是因为清廉指数如此受到关注,其影响力不断提高,使人们逐渐忽视了清廉指数所存在的固有缺陷,而盲目的依据、引用、参考清廉指数对腐败作出判断。本文将通过清廉指数测评方法的研究,分析其存在的缺陷。

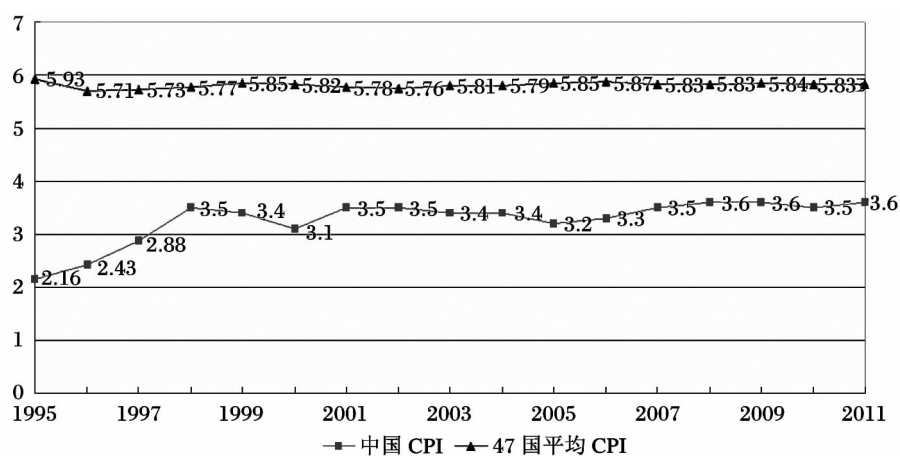


图 1: 清廉指数对中国的评价(1995 ~ 2011 年)

### 三、改革前清廉指数测评腐败的方法

清廉指数评分结果的得出主要有两个重要阶段:一是测评数据的选择和使用;二是对测评数据的技术处理。因此,如何选择测评数据、采用什么方法处理数据就成为决定清廉指数科学性和有效性的关键问题。

#### (一) 改革前清廉指数的数据来源

清廉指数的测评数据主要选自一些独立机构的调查报告,透明国际建立了一个选取相关报告数据的标准,主要有两个方面的要求:一是所选报告的腐败指标必须对国家间进行了量化评价;二是所选报告的腐败指标必须是与腐败高度相关的,测评结果反映的必须是腐败的水平。透明国际依据这样的标准进行初始数据的筛选,并根据一个国家至少在三个报告中被测评才能列入清廉指数的原则确定评价范围。

以 2010 年清廉指数所选取的 13 份报告为例,基本都满足上文所提到的透明国际选取报告的标准。13 份报告都有关于腐败的量化指标,并且与腐败高度相关。从调查对象来看,7 份报告中涉及专家,6 份报告中涉及商人,另有 1 份报告的调查对象是专家和公众。7 份报告没有提供调查样本数量和规模,13 份报告中被评国家和地区数量从 16 个到 201 个不等。透明国际对所选报告进行了数据相关性检验,结果表明这些来自不同机构的报告中的腐败指标具有相关性,透明国际以此证明报告对于腐败测评的有效性。

表 2: 2010 年清廉指数测评数据的来源

序号	报告名称	提供机构	评价指标	调查对象	报告年度	样本数量	被评国家
1	国家绩效评估等级	亚洲开发银行 ADB	透明度、问责制和公共部门的腐败	银行内外的国家工作人员和专家	2009	未提供	28
2	国家政策和制度评估	非洲开发银行 AFDB	透明度、问责制和公共部门的腐败	银行内外的国家工作人员和专家	2009	未提供	53

续表 2

序号	报告名称	提供机构	评价指标	调查对象	报告年度	样本数量	被评国家
3	贝塔斯曼转型指数	贝塔斯曼基金会 BTI	政府反腐败能力	网络记者和组织内外的专家	2009	未提供	128
4	国家政策和制度评估	世界银行 WB	透明度、问责制和公共部门的腐败	银行内外的国家工作人员和专家	2009	未提供	77
5	国家风险服务和国家预测	经济学家智库 EIU	公共部门腐败	专家人员	2010	未提供	135
6	转型国家	自由之家 FH	反腐败行动	专家和公众	2010	未提供	29
7	国家风险评级	环球透视 GI	腐败风险	专家	2010	未提供	201
8	世界竞争力报告	瑞士洛桑国际管理学院 IMD	国家效率中腐败状况指标	各国公司的中高层管理人员	2009	3960	57
9	世界竞争力报告	瑞士洛桑国际管理学院 IMD	国家效率中腐败状况指标	各国公司的中高层管理人员	2010	3960	58
10	亚洲情报通讯	政治和经济风险咨询公司 PERC	公共部门腐败感知	外籍企业高管	2009	1750	16
11	亚洲情报通讯	政治和经济风险咨询公司 PERC	公共部门腐败感知	外籍企业高管	2010	2174	16
12	全球竞争力报告	世界经济论坛 WEF	非法额外支出与贿赂	各国企业高管	2009	> 12000	133
13	全球竞争力报告	世界经济论坛 WEF	非法额外支出与贿赂	各国企业高管	2010	> 13000	139

## (二) 改革前清廉指数的计算方法

在确定清廉指数的数据来源报告之后,就需要对数据进行处理。由于每份所选报告的评价体系不同,如果要应用这些数据,就必须首先进行标准化处理,然后在进行清廉指数的数据修正。

清廉指数主要采取百分比匹配法进行标准化处理,是利用新的报告数据和前一年的清廉指数进行匹配测定,因此只有在新的报告数据和前一年清廉指数都被列入的国家才能进入新一年的清廉指数测评。具体做法是将前一年清廉指数中的最高得分被赋予在新的报告中排名第一的国家的标准化值,并以此类推。例如,假设某机构在 2013 年对五个国家的腐败状况进行了排名:美国 7.5 分、新西兰 5.4 分、越南 3.2 分、马来西亚 2.7 分、印度 2.4 分。在 2012 年清廉指数中,这些国家的分数分别是 8.9 分、9.7 分、3.5 分、6.2 分、2.8 分。那么,百分比匹配法就分别赋值美国 9.7 分、新西兰 8.9 分、越南 6.2 分、马来西亚 3.5 分、印度 2.8 分。百分比匹配法的关键因素在于所选报告中国家的排名,而不是得分,因此经过这种方法处理的数据可能会失去一些其他方面的信息而简单地与前一年清廉指数进行匹配。当然,这种方法的突出优点是在处理的过程中能够将得分限定在 0~10 分之间,对于结果展示具有一定的方便性。

在数据取得标准化值之后,可以将同一国家所参考的多份报告的标准化值进行简单平均值计算,但是这个简单平均值的标准差往往会小于前一年这个国家清廉指数得分的标准差。例如2011年清廉指数排名第一的是新西兰,如果在2012年所选取的报告中新西兰并不是都排在第一,那么2012年新西兰的清廉指数得分就会下降,反之如果2011年排名最后的国家在2012年所参考的报告中并不都是排在最后,那么得分就会上升。关键的问题是,如果参考报告具有不均一性,那么这种趋势就会持续下去,从而产生分数间多样性持续变小的趋势。因此,为了避免这种情况的发生,还需要对上面取得的简单平均值进行 $\beta$ 转换处理,同时使得处理结果保持在0~10分之间。 $\beta$ 转换的公式是:

$$10 * \int_0^1 (X/10)^{\alpha-1} (1-X/10)^{\beta-1} dX$$

如图2所示,经过 $\beta$ 转换后的数据,位于4.0~10.0之间的数据有了微小上升,而位于0~4.0之间的数据有了微小下降,这样就避免了标准差逐年变小的趋势。

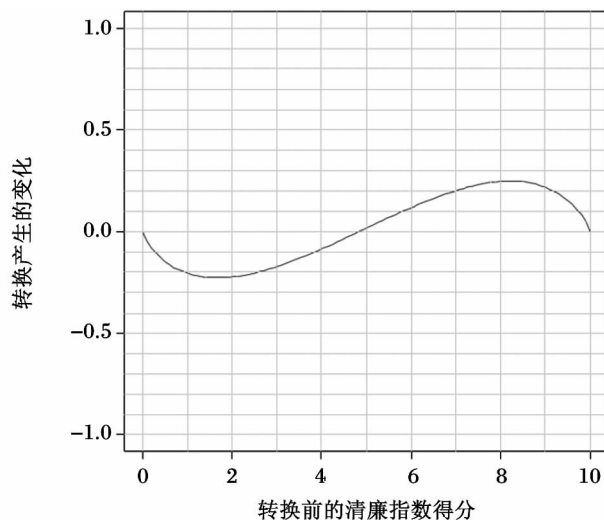


图2: $\beta$ 转换处理示意图

在得到清廉指数的标准值情况下,透明国际还提供了数据结果的置信区间。2001年之前,清廉指数置信区间是根据一个国家得分标准差及其在95%置信区间内的参数评估来提供的,这种方法的问题是置信区间很可能超出了0~10.0之间的范围,例如2001年芬兰清廉指数得分为9.9,其95%的置信区间为9.4~10.4,这不符合清廉指数的评分值域要求。因此,透明国际在2001年之后开始应用仿真方法学(bootstrap)的非参数方法进行置信区间的计算,主要思路是用随机替代方式对某个国家的报告数据进行重新抽样,计算平均值的变化情况。例如,一个国家的清廉指数参考了5份报告,其数值分别是3.0、5.0、3.9、4.4、4.2,随机替代可以变成如下数值5.0、5.0、3.0、4.4、4.4,这样平均值就由原来的4.1变为4.4,按照这种大量随机重新抽样后所计算的平均值分布,就可以反映清廉指数结果的准确性,该方法有效解决了置信区间超出0~10.0范围的问题。但是这种方

法的科学性决定于参考报告的数量,如果参考报告很少,那么置信区间就会向下偏移。此外,在清廉指数数据处理过程中,一直采用平均方法对参考报告的数据赋予权重。

#### 四、改革后清廉指数测评腐败的方法

为了应对世界范围内针对清廉指数的批评,透明国际于2009年成立了清廉指数工作组,开始对清廉指数的方法进行审核和改进,并提出了评估和选择新方法的两条标准:一是准确性,试图找出可以最大化程度减少误差并且在技术上能够更加准确地测评腐败的方法;二是可比性,使用百分比匹配法得到的评分结果是不具有可比性的,应当找到一个新方法使清廉指数随着时间的推移具有可比性;三是易于展示性,新方法应当能够让更多的人易于理解,并在交流展示过程中方便大家使用。

在经过深入论证和研究之后,透明国际决定从2012年开始运用简单平均法计算清廉指数,并将十分制变为百分制,以此来提示人们从2012年开始清廉指数方法已经与以往不同,并提出2012年的评分不能与之前的结果进行比较。这次清廉指数方法的变化主要体现在报告数据来源选取的变化和简单平均法的应用。

##### (一) 数据来源选取标准的提高

透明国际在原有基础上,对所参考的报告数据来源提出了更加细致的标准。一是报告数据必须来源于专业机构,并且清晰地记录了数据的调查方法,调查对象符合严格要求,例如对专家调查时,必须保证其具有真正的专家资格。二是报告数据反映的必须是公共部门的腐败问题,对腐败有着相似的界定。三是报告数据要有定量差距,如经济学家智库的统计报告采取0~4分整数制,数据之间差距太小,信息度就会下降(但这个报告仍然被透明国际所接受)。四是报告数据对于各个国家的评分必须是在统一标准下进行的,而不能对个别国家制定特别标准。五是报告数据应当是连续的,为此透明国际也加强了与各个报告机构的联系。

##### (二) 简单平均法的数据处理过程

透明国际认为,简单平均法是在现阶段清廉指数的最佳改革方法,该方法易于理解和传播,并且使得清廉指数具有了连续可比性,数据处理过程不再与前一年的得分结果相联系。运用简单平均法对数据来源进行处理,主要包括三个步骤:

第一,通过估算缺省值,生成每个报告数据来源的全球数据集。由于每个参考报告所测评的国家都不一致,因此需要估算没有包含在该数据集中的国家的分数,从而获得每个报告数据来源的全球完整数据集。但是只有当报告数据集中有全球性可比性的平均分和标准差时,才可以估算缺省值。第二,对数据进行标准化处理,使其得分在0~100分之间。例如经济学家智库的数据的得分范围是0~4分,0代表最清廉,4代表最廉洁,因此需要使用平均分和标准差把数据集中的得分转化为0~100分的得分范围,置换为0分代表最腐败,100分代表最清廉。这种标准化处理方法计算简单,并且保留了数据来源中的重要信息。第三,计算数据来源的平均分和方差。把每个国家可用的经过标准化处理的数据进行简单的算术平均,就是该国的清廉指数得分。但是算数平均计算中不包括第一步估算的缺省值,只计算原始数据中实际存在的国家的标准化后的平均分。然后计算每

个数据来源中每个国家的得分与最终的清廉指数得分之间的方差,以此衡量一个国家数据来源的数据分布情况。

通过以上分析可以看出,2012年清廉指数在数据选择和处理方面进行了很大的改革,一方面对于数据选择的标准更加严格和全面,在一定程度上提高了数据来源的可靠性;另一方面数据处理方法更加简单清晰,减少了在计算过程中对数据来源重要信息的剔除,增强了当年清廉指数结果的独立性,使清廉指数随着时间推移具有了一定的可比性。

### 五、清廉指数评价的局限性

清廉指数的发布和发展对于腐败测评研究产生了重要影响,对于引起世界范围内对腐败和反腐败议题的高度关注发挥了重要作用,这是清廉指数的核心价值所在。然而,我们也必须看到清廉指数所存在的局限性。清廉指数作为一种主观调查评价方法,通过调查人们对于腐败的感知和印象来评价一个国家和地区的腐败程度,这种方法本身就具有不准确性和非客观性。尽管透明国际对清廉指数的数据来源和计算方法进行了不断的完善和优化,在2012年还对计算方法进行了改革,但是清廉指数在科学性和解释力方面仍然存在着很多不足和缺陷。因此,我们需要正确看待清廉指数的评分结果,特别是对中国的评价。

#### (一) 清廉指数数据来源的局限性

通过上文对于清廉指数数据来源的举例和分析,可以看出清廉指数在选取数据来源的过程中存在着一些比较突出的问题。第一,所选报告的调查对象范围多为专家和商业人士。尽管改革后透明国际关注到了对于专家标准的界定问题,但是却仍然没有具体说明应当如何选择专家,而商业人士为什么能够成为评价腐败程度的主体,这个问题在有关报告中并没有给出具有说服力的解释,我们不能认为那些具有跨国经营管理经验的商人就更了解他国的腐败状况。第二,所选报告大多没有提供调查样本数量和规模,这就降低了有关报告中腐败评价结果的可信度。例如,在2010年清廉指数所选的所有报告中样本数量最多的为13000份左右,而亚洲情报通讯报告的调查样本仅为1750人,这样的调查规模却要对全世界各个国家的腐败状况进行测评,这显然是缺乏科学性的。第三,所选报告被测评的国家数量差距很大,这就隐含了国家评价标准的不一致性。如塞浦路斯在2010年的清廉指数得分根据4个报告计算而成,而韩国却根据9个报告计算而成。第四,所选报告的滞后性。如2010年选取报告中有7份报告是2009年公布的,事实上透明国际在2012年清廉指数改革之前,所选数据中有很多都是过去2~3年发布的报告,很难及时反映出腐败的变化情况。正如杰里米·波普所言“清廉指数最大的作用只是说明过去,现在的情况则更加复杂”(Williamson 2004)。

#### (二) 清廉指数计算方法及结果展示的局限性

透明国际从1995年发布清廉指数开始,就始终关注清廉指数的方法完善和优化,以此不断提高清廉指数的合理性和影响力。评价一个方法是否适用并不是由方法的简单性或复杂性决定的,而是由方法是否能够提高解决问题的科学性来决定的。无论是百分比匹配法还是简单平均法,清廉指数都存在着一些固有的局限性,这主要体现在以下几个方



面: 第一, 清廉指数的本质是对二手数据的加工应用, 尽管透明国际不断提高和细化数据来源选取的标准, 利用综合评价的方式减少单个调查报告由于其调查地区、调查人群等造成的偏差, 并通过改进方法的途径提高评价的可信度, 但是这些努力仍然无法改变计算方法使用二手数据可能存在缺陷的本质问题。第二, 透明国际因为意识到百分比匹配法所存在的巨大缺陷, 而改用简单平均法。尽管简单平均法提高了清廉指数的连续可比性, 但是仍然没有解决数据来源的权重问题。兰斯多夫对清廉指数采取相同权重的解释是“使用简单的方法对所有资料都赋予相同的权重, 可以满足可靠性和专业性的标准”(Lambsdorff 2006)。但实际上, 清廉指数所参考的报告在时间、样本、准确性等方面的差异非常大, 赋予相同的权重显然是不合理的。权重问题比可比性问题更容易被人们所忽视, 从而使人们盲目认可方法的改进。第三, 透明国际依据清廉指数的得分情况对被测评的国家和地区进行排名, 目的是更清晰地进行比较, 也正是因为这种排名形式, 更加引起了人们的关注。但事实上, 这种展示测评结果的方式并不科学。我们以 10 分制为例, 清廉指数将得分的小数点精确到后一位, 2011 年印度得分为 3.1 分, 排名 95 位, 越南得分 2.9 分, 排名 112 位, 仅 0.2 分之差排名却相差 17 位, 透明国际并没有能够对这样的结果进行有效的解释。此外, 由于每年测评国家和地区数量的不同, 排名并不具有连续可比性。第四, 完全准确地测评腐败发生的真实程度几乎是无法实现的, 清廉指数存在的一些缺陷需要一个更长期的改进过程。但是, 我们认为, 透明国际有责任让公众更多地了解清廉指数所存在的问题和局限, 而不是简单笼统地给出排名和得分。在透明国际网站上可以找到一些清廉指数的方法和数据来源介绍, 但是与排名和评分的影响力相比, 了解清廉指数形成过程的公众并不多。

### (三) 客观认识清廉指数对中国评价的局限性

自 1995 年清廉指数发布开始, 中国一直都是其测评的对象。从 1995 ~ 2011 年的评分结果来看, 中国得分和排名情况都呈现出了逐渐上升的态势, 这与中国政府近年来对反腐败斗争作出的努力不无关系。但是通过清廉指数对中国得分所参考的调查报告的研究, 我们发现清廉指数对中国有着较为严重的不公正对待。

以 2012 年清廉指数得分为例, 中国得分为 39 分, 是中国在贝塔斯曼转型指数(28 分)、瑞士洛桑国际管理学院(34 分)、世界各国风险指南(31 分)、世界经济论坛(47 分)、世界正义工程(45 分)、经济学家智库(38 分)、全球观察(32 分)、香港政治经济风险咨询机构(39 分)和透明国际行贿指数(55 分)9 个数据来源中的得分的平均分计算得出。这 9 份报告中调查对象定位为中国人的情况很少, 以透明国际行贿指数(Bribe Payers Index)为例, 该报告中明确提到“在全世界 30 个国家调查 3016 位商界管理人士, 在每个国家至少询问 100 位受访者。其中中国除外, 在中国只访问了 82 位人士”(Transparency International 2011)。而在其他报告中都没有明确指出其调查对象包括中国人。其他国家的调查对象是否真正了解中国? 是否对中国做出了公正评价? 他们对中国的主观感知和客观事实之间存在的差距有多大? 这些问题都成为中国清廉指数得分高低的影响因素, 这显然是有失公允的。

我们需要正确看待清廉指数对中国的评价。一方面, 清廉指数为我们提供了一个评

价我国腐败状况的国际视角,特别是商人的态度,这对于外商直接投资等现实问题有直接影响;另一方面,用科学严谨的态度认识清廉指数存在的缺陷和对中国的不平等对待,不能盲目将清廉指数作为反腐败研究和实践的完全科学依据。

事实上,透明国际充分意识到了清廉指数的局限性,为了修正其存在的缺陷,他们不仅对方法进行完善,而且也采取了其他措施。例如,透明国际于2003年开始发布全球腐败趋势测评(Global Corruption Barometer),又称“腐败晴雨表”,它反映的是人们对于过去一段时间内在腐败和反腐败方面的感知与经历,该测评在2010年和2011年报告中对全球近90个国家和地区的7万户家庭进行了具有代表性的民意调查。

#### 参考文献:

- 胡鞍钢、过勇 2001 “国际组织对各国腐败状况的评价体系概述”,《政治学研究》2001 4: 89—91。  
徐静 2012 “国内外腐败指数及其对比研究”,《中国行政管理》2012 5: 110—114。  
Lambsdorff J. Graf 2006. *Corruption and Reform of Institutional Economics: Theory, Evidence and Policy*. Cambridge: Cambridge University Press.  
Transparency International 2011. “Bribe Payers Index 2011.” Available at: <http://bpi.transparency.org/bpi2011>.

## Methodology and Limitations of the Corruption Perception Index

Guo Yong & Song Wei

(School of Public Policy and Management, Tsinghua University, Beijing)

**Abstract:** The corruption perception index (CPI) is based on the perception of corruption in each country or territory, and draws on relevant questions from a number of different data sources that capture the views of businesses and experts. The release of the CPI has attracted focus on anti-corruption and promoted anti-corruption efforts all over the world. However, the CPI has its limitations. For instance, it is inaccurate and does not capture the perceived level of corruption in a country or territory. Although Transparency International has adjusted and refined the methodology for calculating the CPI, updating the methodology in 2012, there are still some shortcomings and deficiencies in its scientific interpretation. In addition, China's score in the CPI does not fully reflect its corruption and anti-corruption status.

**Key words:** Corruption Perception Index (CPI); Measurement of Corruption; Limitation; Public Governance

(责任编辑: 郜继红)