

法治评估及其中国应用^{*}

钱弘道 戈含锋 王朝霞 刘大伟

摘要：法治评估是基于评估法治状况和现实发展的需要而兴起的。世界银行全球治理指数中的法治指数、世界正义工程的法治指数实践带动了部分国家与地区在更广泛层面展开法治的量化评估。中国的法治评估实践，经历了立法司法领域的评估考察、法治评估的全面探索两个阶段，表现出“嵌入性”、政府主导、法治理解广义化、法治评估区域化等特征。法治评估应该在“试错”中逐步实现区域评估的规范化、体系化与制度化，并作为区域法治建设的新增长点推动地方政府法治目标的实现。

关键词：法治评估 法治指数 法治增长点 法治政府

作者钱弘道，浙江大学光华法学院教授（杭州 310008）；戈含锋，浙江大学光华法学院博士后（杭州 310008）；王朝霞，浙江大学光华法学院博士后（杭州 310008）；刘大伟，辽宁大学法学院副教授（沈阳 110136）。

在经济全球化的过程中，法治的理念和价值被广泛承认。从国际经济领域的投资合作，到人道援助的成效考虑，一国的法治状况都受到相当多的关注。^① 在这种

* 本文是新世纪优秀人才支持项目“中国法治评估研究”（NCET-06-0531）、211 工程项目“转型期法治的理论、制度与实证研究”（20300 * 123210301）、司法部委托项目“法治评估体系和法治指数研究”的阶段性成果。参加本课题或法治指数评审的成员有司法部研究室王公义教授、中共中央党校石泰峰教授、中国政法大学江平教授、国家统计局统计科学研究所吕庆喆研究员、复旦大学孙笑侠教授、中国社会科学院李林研究员、中国社会科学院李步云研究员、中国人民大学张志铭教授、清华大学林来梵教授、上海交通大学郑成良教授、浙江省人大法制委员会胡虎林主任委员、国家行政学院胡建淼教授、《中国纪检监察报》原总编孟祥峰、浙江大学夏立安教授、浙江大学钱弘道教授、香港大学戴耀庭教授。参加课题的还有余杭区司法局局长毛新利、原局长马其鏢、副局长马永年、原副局长郑红；陈建六、姜斌、王帅、钱无忧、范凯文、段海风、吴海燕、何博、梁燕妮、王梦宇、钱国玲、唐文薇、武威、张辉、于晓琴、滕志杰、金卓、邵佳、徐珉川、陈琳、栾婧、李文辉、刘君斌、刘俏蕾、郭楠楠、郁甲利、侯望、姚艺文、王晓茹、汪真燕、王丛、胡淑丽、项欢等同志。

① 例如，在为避免跨国经济交往风险而编制的国家风险评估指南中，法治水平是确定一国是否属于高风险国家的重要依据。

情况下，客观和准确地评估一国法治状况就成为一种现实需要。一些国际组织和国内机构开始这方面的尝试和努力，并取得了一定效果。

近年来，中国一些地方的人民代表大会、人民法院开始对相关工作进行量化和绩效考核；浙江、江苏、广东等地政府在推进区域法治化的过程中开始关注以法治评估为契机展开的制度创新。这些法治评估的探索和实践，是对全球法治研究的回应，更是对中国法治发展客观需要的回应。这一背景下，分析法治评估的域外实践，总结中国各地法治评估尝试的特征和经验，在中国语境下探索法治评估方法和技术的运用，进而预测法治评估在中国的发展，具有重要的现实意义。

一、域外法治评估的兴起条件、模式和反思

域外法治评估涉及国际社会与国家两个层面。不同层面、不同区域的法治评估，在评估主体、评估目的、评估方法上存在一定差异。但是，作为一种相似的实践，法治评估的普遍兴起必然存在共同的背景因素，分析这些因素是研究中国法治评估实践的一个前提。另外，由于各国经济政治结构状况、法治发展水平的差异，在评估内容方面，一国（地区）范围内的评估实践与别国相比，具有较大的不可比性。而国际层面的评估则由于以国别（地区）为评估对象，使用统一的指标进行对比，因而其评价指标具有更为广泛的意义和适应性。

（一）域外法治评估兴起的条件

域外法治评估的兴起和发展基于多方面因素，既体现了法治本身的发展，也反映了全球化条件下人类联系和合作日益强化，需要对共同关注的目标及其状况进行客观评价和认识的尝试与意图，同时，与社会科学领域研究方法的发展和推进也有直接的关系。法治评估本身就是理论上基于评估现实需要而展开的总结和运用。

在法治研究中，用确定构成法治的各个要素或指标对法治进行细化分析是相当通行的做法。^① 马克斯·韦伯认为，法治秩序的主要指标是科层制行政活动的持续性、客观性、权责分明、等级化监控、严格区别公与私的关系、禁止官职的买卖和

^① 季卫东对这种细化法治指标（要素）的研究目的进行了分析，认为建立一套法治指数的主要意义可以概括为：（1）对不同社会体制和文化进行比较；（2）为改造权力结构提供更清晰的蓝图；（3）使法制建设的具体举措和绩效的评价趋于统一化（参见季卫东：《以法治指数为鉴》，《秩序与混沌的临界》，北京：法律出版社，2008年，第55—56页）。季卫东此处所说的法治指数，应该是指法治指标，因为第一、二项中使用的示例，如马克斯·韦伯、富勒的分析都属于指标的论证，没有进行数字化的指数推算和确定；只有第三项使用了相对客观和量化的指数方式评估法治状况。

继承、公务员人格独立和专业化等等。^① 富勒提出了法律八个方面的“内在道德”，包括“法律的一般性”、“法律应该被颁布”、“原则上的不溯及既往”、“清晰”、“不自相矛盾”、“不要求不可能之事”、“连续性与稳定性”、“官方行动与公布规则的一致性”，这实际上就是对法治形式正义标准或指标的一个细化。^② 张文显认为，根据不同时期法治的原理和各国法治实践，可以把不同法治国家的共性化的标志分为形式标志和实质标志：形式标志指法治国家的外在表现方式及实现法治国家的技术条件，主要包括完备统一的法律体系、普遍有效的法律规则、严格的执法制度、公正的司法制度、专门化的法律职业等；实质标志是指依据法治的精神而形成的理性化制度的确立和运作，涉及法律与政治、公共权力与国家责任、权力与权利、权利与义务等方面的关系。^③ 这些对法治构成要素的分析，为现实中的法治评估提供了学理资源。

在社会治理和国际合作实践中，存在着对人口、环境、经济、政府与市场、全球联系等^④人类社会特定方面发展目标进行细化评价的需要。与理论层面的抽象和表述不同，对这些目标的执行情况进行考察、监督，需要更明确的衡量指标和方法。此时，借助统计数据 and 模型进行评估，由于其更客观、具体和量化而受到目标推进者的关注。例如，基于确立和考察人类共同发展的目的，2000年联合国千年峰会号召所有国家为实现一系列量化的、有时间限制的发展目标而努力，该目标后来成为千年发展目标，进而产生了一系列管理计划和评价体系。^⑤ 在国家层面，作为验证行政改革目标的一个重要工具，政府绩效评估在一些国家已经相当成熟，并成为公共管理学的一个重要课题。^⑥ 这些评估实践为法治评估提供了重要的参考。

法治作为当前共同关注的目标，各国和国际社会存在对其发展水平和推进状况进行评价的需要。在国际层面，通过一定的指标设置和指数推算对一国的法治发展水平进行赋值评估，已经成为分析、评价该国投资水平和制度状况的重要做法，如世界银行全球治理指数中的法治指数、世界正义工程的法治指数等。在国家层面，对于后发国家来说，建构型法治的推进更需要一个客观和量化的分析评价标准来衡量法治发展的水平，校准法治建设的局部目标定位。

法治评估受到关注，也是经济全球化的一个重要表现。在全球化早期，随着经

① 参见季卫东：《以法治指数为鉴》，《秩序与混沌的临界》，第55—56页。

② 参见富勒：《法律的道德性》，郑戈译，北京：商务印书馆，2005年，第55—97页。

③ 参见张文显主编：《法理学》，北京：法律出版社，1997年，第244—245页。

④ 世界银行《2006年世界发展指标》五个方面的考核统计数据，参见世界银行本书编写组：《2006年世界发展指标》，中国财政经济出版社译，北京：中国财政经济出版社，2006年，第vi、viii、270页。

⑤ 世界银行本书编写组：《2006年世界发展指标》，第1页。

⑥ 参见陈永国编著：《公共管理定量分析方法》，上海：上海交通大学出版社，2006年，第6页。

济贸易联系的加强,对投资环境、交易安全等进行评价的需要逐渐强化。在这些基于经济因素开展的评估实践中,法治的重要内容——政府和制度备受强调。^①与专门的法治评估相比,在基于经济交往需要而进行的评估中,经济因素与制度因素共同受到关注;而在法治评估中,制度因素则是主要或者唯一的考察对象。

社会科学研究方法和技术手段的进步,促进了法治评估的发展。在早期,对某一社会问题和目标进行评估的研究以描述和定性分析方法为主,如前述马克斯·韦伯、富勒等关于法治内涵指标或要素的分析。随着自然科学的发展,定量分析方法在社会学科领域开始广泛运用。计算机等技术手段的发展进一步使定量方法得到拓展。在这样的背景下,法治评估成为精确的科学研究过程,以指数形式进行的不同指标的厘定和分析构成法治评估的重要内容。^②

(二) 世界银行和世界正义工程的法治指数

从1996年开始,世界银行连续推出年度《全球治理指数报告》,该报告成为衡量被考量国家政府施政水平的一个重要依据。其中,对不同国家的法治状况进行评估并计算相应的法治指数,是世界银行全球治理指数(WGI)的重要内容。

全球治理指数将治理界定为“一个国家权力运行的传统和机制”,法治是治理的六项下位指标之一。^③在《全球治理指数报告》中,计算法治指数首先是通过设置多个法治下位变量(指标)而实现的。变量所具有的确性和可观测性使得对法治这一抽象概念进行测量成为可能。全球治理指数的法治指标设置容纳了世界经济论坛全球竞争力报告(GCS)、盖洛普世界调查(GWP)等对法治的多样化理解,并据此对法治指标进行厘定。^④指标设置后需要调查收集相关数据。全球治理指数依据的是来自30多个不同组织提供的,围绕治理指数设置的上百个个体变量的调查数据。^⑤全球治理指数还包括对公司和家庭的问卷调查,多种商业信息公司、非政府

① 如世界银行《2006年世界发展指标》中的投资环境指标包含政策不确定性、腐败、法院、犯罪、监管和税收管理、金融、电力、劳动力8项指标,前4项指标本身就是法治的重要内容(参见世界银行本书编写组:《2006年世界发展指标》,第vi、viii页)。

② 有关国内外法治指数的研究情况,参见戴耀廷:《香港的法治指数》,《环球法律评论》2007年第6期。

③ 其他五项治理下位指标为发言权与问责、政局稳定与无暴力、政府效能、监管质量和腐败控制(参见<http://info.worldbank.org/governance/wgi/index.asp>,2011年11月5日)。

④ 参见<http://info.worldbank.org/governance/wgi/pdf/rl.pdf>,2011年11月5日。

⑤ 如2010年数据涉及非洲指标(AFR)、非洲发展银行国家政策与机构评估(ADB)、亚洲发展银行国家政策与机构评估(ASD)、欧洲银行重建和发展过渡报告(EBR)、世界经济论坛全球竞争力报告(GCS)等30个组织机构(参见<http://info.worldbank.org/governance/wgi/pdf/rl.pdf>,2011年11月5日)。

组织、多边国际组织以及其他公共机构关于相关指标的主观评估，在主体选择上注意地区的均衡和普遍，以保证对法治认识的全面性和代表性。世界银行全球治理指数目前已成为衡量各国法治水平的重要参考。

世界正义工程（WJP）^①设计了一套衡量国家法治化程度的量化评估工具，目标是建立发展一个强有力、低成本且有效率的方法，可以在较大范围国家间灵敏追踪法治水平在较长时期内的变化。

世界正义工程法治指数的研究由维拉司法研究所和阿尔特斯全球联盟主持，二者共同开发了一套评估法治抽象概念的具体标准，并在印度昌迪加尔、尼日利亚拉各斯、智利圣地亚哥和美国纽约四个城市进行了试验。在法治指标的确定上，世界正义工程试图为世界各国设定一个普遍适用的法治框架体系，但又沿袭了传统的西方法治理念，其逻辑立足点在于对人性的理性怀疑与对权力滥用的担忧。世界正义工程确定的法治指标包括四个原则：负责的政府；经公布且稳定的保护基本权利的法律；可行、公正、有效的程序；可实现的司法正义。在此基础上，确定了16个一级指标、68个二级指标，构建了对法治进行充分有效评估的基本框架。^②该框架实际上是对西方法治政府体制的提炼，是西方宪政体制的缩写。

（三）域外法治评估活动的反思

以全球治理指数和世界正义工程法治指数为代表，主要兴起于国际社会及西方发达国家的此类评估实践，致力于提供一套具有普遍适用性的评估方法。世界银行和世界正义工程两大组织因其各自形成了一套体系化的评估测度方法而成为法治评估的标杆，并在全世界范围内形成了示范效应和影响力，引发了持续的“法治指数化”运动，成为各国或地区进行治理评估与法治评估探索的模本。

针对这些法治评估活动，一直存在争议。^③同时，关于法治指标的内容与法治指数的制定方法，也存在不同认识：前者体现了长期以来对法治的认识分歧，后者表现为不同技术方法之争。

在法治指标的选择上，我们既应当关注法治的共同性，也当考虑法治的差异性。国际上的法治评估活动，以法治的共同性为前提，制定了一套标准，以此评估所有国家的法治水平。在法治共同性的前提下，国际层面的法治评估具有了在各国间对比的可能，在不同评估方式间也可以进行优劣判断。国家层面的法治评估活动，则以法治

① 2007年，世界正义工程被创建为美国律师协会的一个下属机构。2009年，该组织成为独立的非盈利机构。

② 参见 <http://www.worldjusticeproject.org/rule-of-law-index/>，2011年11月5日。

③ 如本文第一作者2008年访问斯坦福大学，与有关学者讨论世界正义工程的法治指数，斯坦福大学有若干位学者就发表了不同意见，认为仅调查四个城市不能较准确反映一个国家的法治水平，如中国这么一个大国，不同地区存在着很大的差异。

的差异性为出发点，在考虑本国（地区）实际情况的基础上制定自己的评估标准。

在法治评估领域，基于对法治共同性和差异性的不同理解，一方面，一些国际组织（如联合国开发计划署、经合组织和世界银行等）和西方评估机构追求在不同国家间建立一套普遍适用的评估尺度，对各国的治理状况进行整体性测量和比较；^①另一方面，这种“普适性的标准”在面对各民族国家的历史文化、政治制度和经济发展水平时，在客观性、公正性和权威性方面一直饱受质疑。

在当前的法治评估实践中，单一国家或地区更多采取“因地制宜”的个性化制定方案。相对于国际层面的法治评估，一国范围内的法治评估具有更强的针对性和灵活性。中国形成了自己独特的法治发展路径，中国应该在法治评估方面享有发言权。因此，在国际视野下，结合中国法治现状创新法治评估模式非常必要。

二、中国的法治评估实践

中国的法治评估实践，既受国际上法治评估研究的影响，也是中国法治建设发展的逻辑延伸。中国的法治评估大体分为两个类型或者发展阶段。

（一）立法司法领域的专项评估

在中国，法治评估在受到关注之前，实践中已经产生对特定领域法律现象的评估考察。此类评估与法治评估具有大体相同的方法论背景，但并不完全等同。

首先是立法后评估。此类评估首先盛行于西方发达国家，主要是由政府或立法机关采取类似检验、测试、数据分析等技术性手段，对立法成本、立法效益、立法效果进行评估的一种活动。^②近年来，立法后评估在中国也受到关注。^③由于特定的社会背景差异，国内外立法后评估在关注内容和取向上有所不同。^④与法治评估

① 参见俞可平：《中国治理评估框架》，俞可平主编：《国家治理评估——中国与世界》，北京：中央编译出版社，2009年，第5页。

② 刘松山：《全国人大常委会开展立法后评估的几个问题》，《政治与法律》2008年第10期。

③ 2004年，国务院颁布《全面推进依法行政实施纲要》中规定“积极探索对政府立法项目尤其是经济立法项目的成本效益分析制度。政府立法不仅要考虑立法过程成本，还要研究其实施后的执法成本和社会成本”，此后，国务院启动了对《信访条例》、《艾滋病防治条例》、《蓄滞洪区运用补偿暂行办法》等行政法规的立法后评估。2011年，全国人大常委会将《科学技术进步法》、《农业机械化促进法》作为立法后评估的试点对象（参见《积极稳妥开展立法后评估工作——访全国人大常委会法制工作委员会》，《中国人大》2011年第14期）；地方人大开展的立法后评估，比较有代表性的是上海市（参见许安标：《立法后评估初探》，《中国人大》2007年第8期）。

④ 有学者指出，大多数西方国家立法后评估对象主要限于立法本身的成本和效益，但对立法体制、立法程序和立法后的实施情况较少评估。这主要是因为这些国家的立法制度

相比,立法后评估涉及范围较窄,且主要关注成本效益分析。良好的立法是法治的前提,立法评估可以视作法治评估的一个方面,但无论是评估方法还是评估范围,二者均有明显不同。

其次是司法评估。司法被视为法治发展的重要环节,司法评估本身可以为法治评估提供数据支持。中国开展了依托计算机信息技术进行的以量化为主的标准化考核。一些地方法院制定审判流程规程及相应的审判质量效率评估体系。^① 依托现代信息管理系统进行司法评估的主要意义,在于为法院提供了科学评估审判质量和效率的工具。

再次,其他领域对法治的评估也值得关注。例如,在联合国开发计划署和商务部委托的“中国治理评估框架”课题中,法治被列为该框架下12项善治目标之一,该框架下的法治又包括10项具体指标,^② 构成法治评估的基本内容。此类评估关注法律状况的整体,不同于前述相对局部的立法司法领域的专项评估,^③ 但其对法治的考察往往更为宏观,且多服务于法治以外的目标,因此同样不能代替专门的法治评估。

(二) 法治评估的全面探索

近年来,香港、北京、江苏、浙江、上海、深圳等中国发达地区开始了法治评估的全面探索。这里选取部分地区的法治评估情况,从参与力量、目标、法治的基本要素三方面进行分析,考察其共性和差异,探求中国法治评估的特色、优势和不足。

香港与内地具有近似的文化背景,其法治化程度相较内地又具有先发优势。杭州市余杭区推出的法治指数就是以香港为首要参照对象。

香港的法治评估由香港社会服务联会(HKCSS)发起,是“香港社会发展指数计划”(Social Development Index, SDI)中的一项。^④ 研究者运用质化和量化相结合的方法制订了2004、2005年度的指数。在对法治的理解上,研究者首先确定认识法治的两个进路:价值性进路和体制性进路。前者更关注法律是否确认某些基本人权和价值;

① 已相当成熟,法律制定后基本不存在得不到实施的问题(参见刘松山:《全国人大常委会开展立法后评估的几个问题》,《政治与法律》2008年第10期)。

① 如上海、江苏、浙江等省市的法院审判质效评估尝试(参见张国香、宁杰:《审判质量效率评估体系研讨会召开》,http://www.chinacourt.org/public/detail.php?id=189893,2011年11月3日;金维、沈法:《构建审判质量效率评估体系》,《江苏法制报》2005年2月1日,第6版;张光宏:《人民法院绩效考评体系的科学构建》,《浙江审判》2009年第4期)。

② 参见俞可平:《中国治理评估框架》,《经济社会体制比较》2008年第6期。

③ 这是其与法治评估的近似之处,但往往因为法治评估服务于“治理”这样的更宏大主题,法治评估的意义难以得到充分体现。

④ 参见戴耀廷:《香港的法治指数》,《环球法律评论》2007年第6期。

后者着重考察政府是否通过法律和在法律下行事。研究者认为,就法治的测量要求说,体制性进路范围较窄,但易于指数的确定。根据体制性进路,研究者将法治分解为七个方面的条件和表现:法律的基本要求;依法的政府;不许有任意权力;法律面前人人平等;公正地施行法律;司法公义人人可及;程序公义。^①香港法治指数在数据获取方式上主要采用专家打分,补充采用民众调查,这种方法既通过法律精英层反映了法律运转的现状,又反映了民众对法律的感受,体现了法治运作的整体效果。

余杭法治指数借鉴了香港模式,同时根据地区法治发展的现实进行了创新。余杭法治指数以余杭区法治建设的9个目标——党委依法执政、政府依法行政、司法公平正义、权利依法保障、市场规范有序、监督体系健全、民主政治完善、全民素质提升、社会平安和谐——为依据;^②在此基础上逐层分解,最终形成包括27项主要任务、77项评估内容的指标体系。其评估结构可以用“1(1个指数)4(4个层次)9(9项满意度调查)”来概括:1个法治指数,是以数字方式对年度余杭法治发展水平进行最直观简练的评价;4个层次为总指标、区级机关指标、乡镇指标、农村社区指标;9项群众满意度调查包括党风廉政建设、政府行政工作、司法工作、权利救济、社会法治意识程度、市场秩序规范性、监督工作、民主政治参与、安全感。^③这些指标设置既与内地的政治经济制度特点相联系,也不可避免会受到当前中国法治发展初期阶段的影响,前者决定了指标较为稳定的基本特征和核心构成,后者则说明部分局部要素的可调整性。

由于地区经济政治体制、法治发展水平等差异,香港与余杭两地法治指数在评估方法、对基本概念的认识等方面存在较多差异。

第一,香港和余杭采用了不同的法治指标。香港的七条指标主要围绕立法、执法和司法审判,对民主政治、市场经济等法治外围制度和理念关注较少。余杭的九条是根据内地地区法治的发展状况及中国特色社会主义法治的要求设定的,民主政治完善、市场规范有序等构成其法治指标的重要内容。这些体现了对于法治不同内涵认识的指标设置,与两地法治发展的不同阶段性有直接的联系。

第二,余杭借鉴香港模式,侧重量化分析技术,同时不忽略定性研究方法,体现了主观判断与客观数据的结合。在客观数据的获取上,余杭的数据信息主体多于香港,如余杭囊括党委依法执政、政府依法行政等方面的数据,香港对这方面则不太关注。其原因一如前述,是香港法治已经得到较充分发展,开始独立于政治、经

① 参见戴耀廷:《香港的法治指数》,《环球法律评论》2007年第6期。

② 参见余杭法治评估体系课题组:《法治量化评估的创新实践——余杭法治报告》,《中国法治发展报告 Na 6(2008)》,北京:社会科学文献出版社,2008年,第370—372页。

③ 参见余杭法治评估体系课题组:《法治量化评估的创新实践——余杭法治报告》,《中国法治发展报告 Na 6(2008)》,第366页。

济基础的必然体现。而余杭虽然在内地属于经济较为发达的地区，但对作为政治、经济制度全面发展表征的法治来说，其仍然从属于我国法治整体落后这一状况，故而有关注法治的基础性要件，如政治上的民主发展，甚至政府依法执政等等仍然是需要关注考核的内容。

第三，余杭的群众满意度调查相较香港更为广泛。每个年度，余杭法治指数要做1000—1500份的民意调查，问卷对象涉及街道居民、外来务工人员、农民等各个群体。^①群体的特征区别于香港。此外，能展开较大规模民意调查也与政府主导和推动有很大的关系。

第四，余杭法治评估的重点在于通过评估发现问题，提出建议，推进法治建设。香港模式在这点上相对淡化。^②反映了我国内地法治建设的积极推进和建构式特征。香港则主要是进行一个法治状况的考察和评析，推进的意图和急迫性并不明显。

第五，香港的评估机构是非政府机构，比较独立。^③余杭的评估机构的评审组是由来自大学、研究所的专家组成，^④但其独立性仍遭到一些质疑。^⑤为此，评审组专家专门在香港设立了中国法治研究院，后期余杭法治指数均由中国法治研究院完成。

除香港和余杭外，北京、^⑥深圳、^⑦昆明、^⑧四川^⑨等地也开始进行法治评估指标的探索。以北京和深圳的法治评估为例，二者存在如下差异。

-
- ① 钱弘道：《余杭法治指数的实验》，《中国司法》2008年第9期。
- ② 法治余杭课题组：《余杭法治指数报告》，《中国法治发展报告 Na 7（2009）》，北京：社会科学文献出版社，2009年，第385页。
- ③ 香港的社会服务联会属于独立于政府之外的民间机构，经费、办公场地、人力都有一定局限。香港法治指数的课题，是香港社会服务联会委托香港大学法律学院戴耀庭教授牵头完成的。
- ④ 余杭法治指数的设计制定，是余杭区委区政府委托浙江大学光华法学院牵头完成的（参见法治余杭课题组：《杭州余杭法治指数报告》，《中国法治发展报告 Na 6（2008）》，第386页）。
- ⑤ 质疑的问题主要是，政府是否会影响专家们的独立判断，“法治指数”是否会成为一种新形式的政绩工程等（参见《“法治指数”：无公信力则无生命力》，《新京报》2008年4月7日，第A02版；江德华：《谨防“法治指数”变为“自说自话”》，《21世纪经济报道》2008年4月8日，第3版）。
- ⑥ 参见杜晓：《官员是否依法行政 “法治GDP”说了算》，《法制日报》2010年3月23日，第4版。
- ⑦ 参见周成新、王成义主编：《深圳市法治政府建设指标体系（试行）解读》，深圳：深圳出版发行集团、海天出版社，2009年，第323页。
- ⑧ 参见昆明市委法治昆明建设领导小组办公室：《昆明市召开新闻通报会发布〈法治昆明综合评价指标体系〉》，[http://www.sft.yn.gov.cn/newsview.aspx?id=1333522 &DepartmentId=1](http://www.sft.yn.gov.cn/newsview.aspx?id=1333522&DepartmentId=1)，2011年11月3日。
- ⑨ 参见四川省司法厅课题组：《四川省“法治指数”构成与法治环境评估体系》，<http://www.scsf.gov.cn/sfyj/1287.jhtml>，2011年11月3日。

第一，在指标设计上存在“既有评价”与“目标设置”的区别。北京指标设置有助于对过去法治建设成效的考察，如在二级指标“民主政治建设”中，相应量化考察指标有“党组织提出的立法建议数量；人大代表选举中的违法案件数量；选民直接选举人大代表参选率”等，这些主要需要基于过去的数据和事实得出。深圳市则主要是设置建设法治政府的目标，如在“政府工作法治化”项目下规定了事项内容、完成形式、完成情况、责任部门、进度安排等，属于完全的目标定向，难以评估已发生的事项和情形。

第二，在指标设计上存在着“定性要求”和“定量评价”的区别。北京指标的设置考虑到数据统计收集工作的可行性，并进行了相应的模型设计，为定量研究奠定了一定基础。深圳指标则以定性要求居多。从国外法治评估实践以及研究方法的发展来看，客观精确的定量分析无疑更符合法治评估的要求。

北京和深圳的实践集中反映了当前国内法治评估的两种基本模式，其共同名称下的不同实践其实存在较大差异。但是，无论评估模式有多大差异，法治评估的意义和价值已经受到基本的承认和关注。总体看，法治评估已被认为是当前或转型期推动法治发展的一个抓手。在这种背景下，运用科学的方法测度法治指数、进行法治评估，十分必要。

三、法治指数的测度方法

相比于各国法治现实状况和对法治本身的认识理解，方法的选择和运用具有更大的可比性和可通约性。用指标化或指数化^①的方法衡量一国（地区）的法治水平，已然是一种趋势。其根本原因在于，指标化或指数化传递了一种可观测的、标准化的审视法治的方法，这套方法是包括指标、标准、数据、评估主体、数据分析、指数计算等在内的方法体系。

（一）指标的设计

1. 法治基本维度或一级指标的确定

确定指标的方法包括对已有文献、资料的分析，引入专家的理解、讨论等。经

^① 指标与指数是两个不同的概念。指标是为了评估的需要，对评估的内容或要素所作的进一步解释，如将依法行政、遏制腐败作为法治评估指标。指数是以量化的方式确定指标的权重及其等级，并以一定的公式计算出法治水平的分值作为评估结果，如香港2005年的法治指数为75分，余杭2007年的法治指数为71.6分。指标是评估实践中必不可少的一个要素，有评估必有评估指标；但是有评估未必有法治指数，在中国的一些评估实践中，仅对评估指标进行分析判断，在确定指标的权重时也可能运用量化方法，但最终并不以数值衡量法治水平。

验判断法、德尔菲专家评价法、穆迪优选法、相关分析法和聚类分析法等都是行之有效的办法。

在确定指标前，应确定评估的原则或者关键要素。世界银行用以测量全球治理水平的六大全球治理领域，是由世界银行学院在获取 200 多个国家关于治理的二手资料的基础上，根据全球治理的国际基准对治理的多个方面进行反复比较、排序，剔除不适宜的治理要素，最终筛选出来的。世界正义工程经过与 100 多个国家的 17 个专业领域的领导、专家、学者、普通工作人员的长期探讨，得出法治的四项基本原则，并经过广泛调研与试点，提炼出评估不同国家法治状况的指标。^①

也有一些研究机构用了不同的方法，如人权研究中心和 Bertelsmann 的转型指数引入了外部专家参与治理指标的筛选、确定。^② 联合国开发计划署和经合组织的 Metagora 项目在设置治理指标时均考虑不同国家治理的差异性以及针对不同国家进行评估的方法的差异性，采用了公民抽样调查法。香港法治的“七项条件是从不同有名的法治研究著作中归纳及有系统地罗列出来的”，是任何法治实践所不能或缺的七个方面。^③ 余杭法治指标的九个方面是余杭区委与法学理论及实践专家反复探讨，通过经验判断，并通过采用德尔菲专家评价法得出来的。^④

对比不同机构法治指标设置的不同方法可以发现，法治指标方法的选择不仅仅是个方法问题，在更大程度上是与对法治内涵的理解与认知相关的，用何种方法设置法治指标不单单取决于方法本身的优劣，更取决于评估目的、评估层次、评估对象等多方面内容。世界银行等国际层面的法治评估，其指标设置更多考虑的是具有普遍意义的法治理念，突出法治指标对所有评估国家普遍存在的法治核心问题的反映和揭示，强调法治指标的适用性。突出对所有被评估国家法治领域核心问题和关键要素的总体反映和揭示，弱化了法治指标在不同国家的适用差别。中国的法治评估实践，由于国内本身发展阶段所限，通常更强调作为法治发展基础的制度性要素，包括政治、经济、社会背景等综合性因素。由于不以排序、比较而是以改善区域法治水平为目的，法治指标的设置通常从各地区的实际情况出发，以揭示地方法治进程和法治状况的改善程度为目的，以评估地的法治建设目标为基准，因而具有强烈的国别和地区特色。

2. 一级指标的细化和分解

通过上述方法确定的一级指标不能直接用来测量，因为这些指标是原则的、抽

① 转引自赵昕编译：《可以量化的正义：衡量法治水平的十六项“法治指数”》（上），《人民法院报》2010年6月18日，第5版。

② 周红云：《国际治理评估指标体系研究述评》，《经济社会体制比较》2008年第6期。

③ 戴耀廷：《香港的法治指数》，《环球法律评论》2007年第6期。

④ 钱弘道：《余杭法治指数的实验》，《中国司法》2008年第9期。

象的法治精神的阐述，可测性差。因此，必须对其进行细化和分解，将其转化为明确的、可操作的次级评估指标。

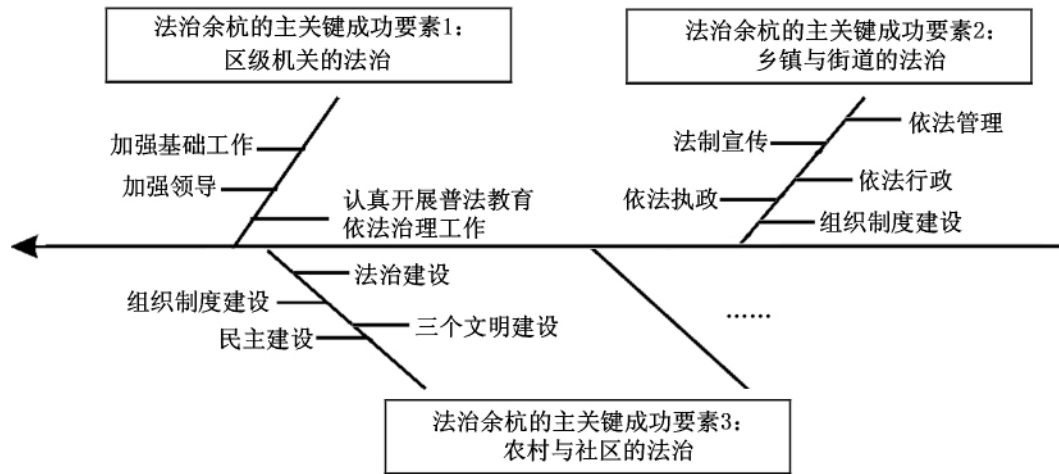


图1 余杭法治指标的要素分解法：鱼骨图的描述

细化和分解一级指标的方法既有定性的，也有定量的。定性的方法主要是指从法治指标的上级指标逐层分解到下级指标的推理方法，也被称为要素评估法或鱼骨图法。在鱼骨图中，鱼头代表法治，鱼骨架上的刺（主刺）代表法治的多重价值目标，这些目标进一步细分为法治评估要素，即鱼身上更细小的刺（小刺）。也就是说，在进行法治评估时，首先确定法律的价值目标，并藉此细化出量化的价值要素。通过对这些要素的评估，进而考察法治的价值实现程度。余杭法治评估在细化评估指标时，先由区级九个目标分解出二级指标，再层层分解到区级机关、乡镇和街道、农村和社区三个层次，最终形成余杭法治指标的“149结构”。^①图1以鱼骨图的形式，展现了余杭四个层次的法治指标的分解和细化过程。“149结构”中的“4”所蕴含的意思是将余杭的法治指标体系划分为“1+3”模式。在图中，鱼身象征“1”，是一个总体性指标；这个总体性指标进一步分解为“3”，即围绕鱼身形成的三根主刺。围绕每个主刺，又分化成若干小刺。以区级机关这一层级为例，加强领导、认真开展普法教育依法治理工作、加强基础作用等指标都是区各职能部门法治建设所必需的法治指标；法制宣传、依法执政、依法管理、依法行政等是乡镇与街道这一层面的法治分解出来的体现该层次法治特色的个性指标。

3. 权重的设置

法治可以从多方面来理解，所以法治评估的指标通常是由多个向度的指标围成的一个指标群。在世界银行“全球治理评估”中的六个治理指标中，同时包含了集成指标和单一指标，形成了一个指标体系。不同指标对于评估对象的贡献不能等量

^① 参见钱弘道：《余杭法治指数的实验》，《中国司法》2008年第9期。

齐观，必须通过权重区分不同指标的重要性。

确定指标权重的常用方法也很多。在世界银行2004年组织实施的“国家政策与制度评估”项目中，16项指标被不均衡地分成了4组。每组指标的权重被均等地设置为25%，在每组内，每项指标的权重也相等。只是到了更下一级的指标时，才出现了不同指标之间的差别。在这种情况下，二级指标间的最大差距为8.33%和5%（第三组和第四组25%的总权重被分配给了5个组内指标，每个指标均占5%），三级指标间的权重的最大差距为6.25%（8.33%的二级指标权重仅被分配给下属的二个三级指标，各自比重占75%和25%）和1.0%（5%的二级指标被均等地分配给了5个三级指标）。^①所以，不同指标间重要性的大小并不完全是由该指标在政策与制度治理中实际发挥作用的大小决定的，而是由指标的多少决定的。指标被分的越精细，数量越多，则其权重数越小；相反，指标被划分的数量越少，每个指标被赋予重要权重的几率就越大。这样纯粹基于操作的便捷性来设置权重的做法从方法论来看有一定的合理性，但是对法治本质的反映却在准确度上大打折扣。

香港和余杭在确定法治指标权重时均采用德尔菲法，^②通过匿名的方式将事先设计好的权重打分表以信函的方式传送给专家进行重要性评价。香港和余杭的最后权重确定非单独进行，而是同法治指标水平打分融合在一起。每位评审专家对每个指标进行权重评分的同时，也对该指标所代表的法治水平的高低进行打分，由此通过权重与法治水平得分的乘积反映该指标的最终得分。^③不同专家关于同一指标的权重评价的平均值将作为最终计算法治指数的函数中的一个变量。余杭分别选取了20位内部组和外部组评审成员，对表中所列九项余杭法治建设目标进行逐一权重设置。在设置时，采用的是1到10的绝对数值形式（不同于世界银行项目的百分比形式的权重）。在收集到每位专家关于每个指标的权重数后，余杭法治评估课题组对不同专家针对同一个指标的权重分数进行了处理。这种处理首先是计算调整后的平均分。调整后的平均分是指所计算的指标权重及指标得分在去掉一个最高分及最低分之后计算的18位评审专家打分的平均值。余杭与香港两地共同的一个特色是，两地

① 刘铎、程宇编译：《世界银行：国家政策与制度评估》，俞可平主编：《国家治理评估——中国与世界》，第159—161页。

② 德尔菲法是以不记名方式征询专家对某类问题的看法，并对专家意见进行统计整理，再将经过整理的调查结果反馈给各位专家，让他们重新考虑后再次提出自己的看法，并特别要求那些持极端看法的专家详细说明理由。经过几次反馈，大多数专家的意见趋向集中（参见何俊德：《项目评估——理论与方法》，武汉：华中理工大学出版社，2000年，第36页）。

③ 参见余杭法治评估体系课题组：《法治量化评估的创新实践——余杭法治报告》，《中国法治发展报告No.6（2008）》，第370—372页；戴耀廷：《香港的法治指数》，《环球法律评论》2007年第6期。

不但让评审成员以量化打分的方式参与确定指标权重，还让评审成员用文字描述每项指标权重大小的理由。^①这在一定程度上提高了权重设置的精准度，更为重要的是，这种权重设置的科学性是专家根据各法治指标在地区法治发展中实际发挥的作用进行估量的，相对纯粹的方法论增添了对于地方语境的现实考量。

（二）法治指标数据的收集

评估信息与数据的收集，可以直接运用法律统计数据、调查数据，也可以通过不同方法的组合进行数据获取。

法律统计数据大多来自被评估单位或国家统计局等提供的工作信息（如司法部门的工作报告、年度工作成果等）或统计数据。世界银行全球治理指数获取的数据主要是靠一些机构提供，自主调查比较少。但是值得研究的是，这些数据能否提供评估所需要的信息？这些信息是否能够充分反映法治现状？如果承认法律统计数据只是对法律现实的一部分记录，那么为了保证法治评估的真实性和可靠性，通过社会调查获取第一手数据就十分必要。

社会调查数据主要是通过测量目标群体对法治指标的主观感知获得的。已有的法治评估大多数是对目标群体进行随机抽样。世界正义工程法治指数所依据的数据主要有两个来源：（1）一般人口调查（GPP），由当地主要的调查公司进行，每个国家选取三个城市、共1000名受访者作为代表样本；（2）专家问卷（QRQ），主要对民商法、刑事司法、劳动法以及公众健康专家进行问卷调查。^②

香港的法治指数评估同时采用了法律数据和社会调查数据。^③余杭也同时运用了法律数据和社会调查数据。法律数据主要是一些能够反映余杭区法律实践的数据，如被控犯罪人数、挽回经济损失数、反贪局的调查案件数、行政复议案件数、上诉案件率、民商（刑事）法律援助案件数等统计性数据。^④社会调查数据包括民众法治满意度的主观性认知调查^⑤所获得的数据等。以组合方式获取法律数据的做法主要是为了弥补单一方法获取数据的片面性、信息不完整性与信息失真的状况。尤其当既有的法律统计、司法统计无法直接对应于设置的法治指标，以及存在统计失真的情况下，补充运用一手调查信息，是妥当的做法。将民众对法治满意度的主观评价纳入社会调查中不但增强了评估数据的科学性，而且这本身就彰显了法治评估过程的

-
- ① 参见余杭法治评估课题组：《法治的增长点——区域法治评估的实验》（未刊稿），杭州：2009年，第120页；戴耀廷：《香港的法治指数》，《环球法律评论》2007年第6期。
- ② 参见 <http://www.worldjusticeproject.org/rule-of-law-index/>，2011年11月5日。
- ③ 戴耀廷：《香港的法治指数》，《环球法律评论》2007年第6期。
- ④ 参见余杭法治评估体系课题组：《法治量化评估的创新实践——余杭法治报告》，《中国法治发展报告 Na 6（2008）》，第388页。
- ⑤ 参见钱弘道：《2008 余杭法治指数：数据、分析及建议》，《中国司法》2010年第3期。

民众参与与民主化，可以说，法治评估本身就体现了对民主法治化进程的推进。

（三）评估主体、数据分析及法治指数的计算

从主持主体看，法治评估可分为政府主持的评估与非政府组织主持的评估。此外，邀请哪些人参与评估，也是影响评估结果的重要环节。“内部人”和“外部人”的选择是评估中的重要步骤。将公共政策评估中的“利益相关者”理论^①引入法治评估是一种有效的方法，该方法将普通民众作为法治评估主体，在评估中植入了更多民意因素。余杭和香港在评估人员的组成上都同时包括“内部组”和“外部组”。余杭的外部组评估意见占82.5%。^②香港的外部组将独立于政府之外的律师、教授和民众作为法治评估的参与者，可信度较高。^③

香港和余杭法治指数计算运用的都是加权平均法，^④即先计算出每位评估主体对每个指标的评价分值，这一分值等于该指标权重得分与指标等级得分的乘积。然后，运用加权平均法计算出每位评估主体对所有评估指标评价的总和，即为该专家对评估区域法治总体水平的量化评价。基于每位专家的法治水平评价，在剔除无效评价（去掉最高分与最低分的方式）的基础上合计出最终的法治指数。

评估实践证明，对法治可以进行量化评估。每种方式都有其优势和不足。因为法治是一项具有价值判断的评估对象，对其客观性评估存在一定限度。因此，法治指数测定方法应最大限度地坚持科学化，这是法治评估发展的一个基本趋势。

（四）法治评估方法的科学性问题的科学性问题

现有的域内外法治评估实践在一定程度上展示了指标设置、数据收集及指数计算的方法与技术，同时也引起了对法治评估方法科学性的争论。评估学意义上的方法在多大程度上适用于对法治的评估？域外法治评估的方法与技术是否或者在多大程度上可以被中国借鉴？法治评估如何才能更科学？这些都值得深入研究。法治评估不同于经济评估，在讨论法治评估的科学性问题上，需要更广阔的视野、更开阔的思路。

法治能否量化？提出这个质疑是可以理解的，毕竟价值判断是法治评估不可或缺的内容。正因为法治存在主观价值判断，所以我们不能完全套用纯粹的自然科学

-
- ① 法治评估中的利益相关者是指那些能够直接影响立法、立法活动、立法评估活动或被立法、立法活动、立法评估活动所直接影响的人或团体，包括立法主体、执法者、司法者、社会公众、立法评估活动发起人、立法评估主体以及传播媒介。
 - ② 参见钱弘道：《余杭法治指数的实验》，《中国司法》2008年第9期。
 - ③ 参见戴耀廷：《香港的法治指数》，《环球法律评论》2007年第6期。
 - ④ 参见戴耀廷：《香港的法治指数》，《环球法律评论》2007年第6期；钱弘道：《余杭法治指数的实验》，《中国司法》2008年第9期。

观念方法。法治很难绝对量化，但也不是说绝对不能量化。事实上，许多传统上无法量化的社会科学的研究对象已出现一种“指数化”的趋势。^①但是，我们必须认识到，法治无论作为一项原则、价值目标抑或具体制度，其内涵都极为丰富，如何化约为数字化的标准的过程极其复杂。一方面，由于民主与法治所具有的国家垄断性特征，体系庞大且作为一个有机的整体发挥作用，使得对法治的评估涉及多层次、多维度的标准；另一方面，法治指标不全是由简单的统计数据或其他容易检测到的实验数据构成，一般会涉及到价值判断，并且受制于部门利益与政务公开程度，法治数据的采集并不容易，^②这些情况都会影响法治评估的客观性和准确性。

指标设计直接关系评估方法的科学性。在法治指标方面，各国是否应该适用全球通用的普适性指标体系，以及我国各地区是否应适用全国统一的法治指标都存在分歧。作为评估原理与方法论在法治领域延伸的跨学科的产物，法治评估除了体现其方法的跨学科特色外，其本身的“问题意识”表明，它更是一个关乎政治理性、实质正义的问题。在某种程度上，评估方法的科学性不是通过其自身方法的完善就能绝对实现的。评估有其天然的局限性，必须依靠与方法紧密相关的特定国家或地区政治体制、法治演进模式、民主法治的实现程度等法治资源，才能实现法治评估的科学化。

评估的一般原理与方法本身并不限定评估的对象与内容，具有“价值无涉”的特性。但是，对于法治的评估、制度的评估、政府绩效的评估等社会实践的评估都具有与国家愿景与时代使命相一致的价值导向，这一方面增加了评估方法运用的复杂性与难度，另一方面也提醒我们，评估方法只是评估科学性的形式，评估内容本身才是最值得关注与引起重视的实质所在。一味陷入方法论的陷阱中，可能导致本末倒置，滋生评估自发性、盲目化、形式化^③等弊端。

四、中国法治评估的特点、问题及应用预测

通过考察、比较、研究国内外的法治评估实践，可以总结出中国当前法治评估的特点与存在的问题，进而对中国法治评估的前景进行预测。

（一）中国法治评估的特点

第一，法治评估的“嵌入性”。“嵌入性”是指中国的法治评估从属于依法治国

① 民生指数、幸福指数、安全指数等等都是表现。

② 参见占红洋、李蕾：《初论构建中国的民主、法治指数》，《法律科学》2010年第2期。

③ 有学者在评价国内政府绩效评估实践时总结评估存在的主要问题：（1）评估的自发性、半自发性以及跟风效仿现象严重；（2）评估的运动式、突击式以及形式主义痕迹明显；（3）忽视开展政府绩效评估的基础性工作（参见高洪成：《“异体评估”视域下的政府绩效评估研究》，沈阳：东北大学出版社，2009年，第78—79页）。

方略，从属于法治现代化这一框架，属于中国经济体制和政治体制改革的重要环节，具有转型期的时代特点。这种法治评估内嵌于地方政府的法治建设工作之中，作为其一个环节而存在，这种既有别于域外实践、又体现了具有中国社会转型期的时代特点的做法，直接决定了中国的法治评估的“嵌入性”特点。法治评估活动在不同国家有不同的运作方式，不能抛开一个国家制度结构的特点，单独谈论法治评估方式的优劣。

第二，法治评估由政府主导。除香港外，中国当前在各地进行的法治评估体系主要由政府主导开展，或者由政府直接界定编制，如深圳市的做法；或者由政府发起，由有关学术机构承担并主要服务于政府的政策导向和要求，如北京市和四川省的做法。2004年国务院在《全面推进依法行政实施纲要》中提出：“全面推进依法行政，经过十年左右坚持不懈的努力，基本实现建设法治政府的目标”，政府主导模式是一定时期内法治实践需要的产物，是这一特征的集中反映。相比而言，国际上法治评估的发起、主持主体多为国际组织或者民间机构，其并不服务于某一特定的政府。但是，国际组织发起的评估并不能保证其在发达国家与发展中国家划分视角之下的完全中立。中国的法治评估实践有助于否定“西方中心主义”的价值偏见，这可能是中国实践的重要价值之一。

第三，法治理解广义化。中国的法治建设作为中国近代以来救亡图存的现代化整体进程的一个组成部分，现代化基础的全面建构必然对中国的法治评估产生影响，于是我们看到各地评估法治过程中对法治的广义界定，在一些国家和地区不属于法治的内容也被纳入评估。中国往往将民主等传统上不属于法治或者与法治并列的内容作为法治评估的重要变量。比如，北京市法治建设指标体系在法治的二级指标中纳入了“民主政治建设”项，具体包括选民直接选举人大代表参选率、政协委员提案采纳率、被国家机关邀请参加执法监督的政协委员比率、村委会（居委会）选举结果当场公布率等。这说明了中国正处于法治发展的初级阶段或者上升阶段，与法治相关的制度建设与实践不得不在发展中受到重视。

第四，法治评估区域化。当前，中国的法治评估基本上是在各地开展。^①某种程度上可以说，地方政府推进法治建设的过程中存在一种“县际竞争”^②的现实状

① 《世界银行治理报告》的法治指数、世界正义工程法治指数是基于同一计算方式和相对稳定的数据来源针对各国的评估，为国别之间的比对提供了较好的参考。而包括我国在内的以国家为单位进行的评估则由于主体视野的局限性或者主要关注于国内法治状况的评估，对国家之间的有效比照性不足。

② 县际竞争最早由张五常解释中国经济发展奇迹时提出，即以县政府为主体、以一个县的地域为边界的县与县之间相互竞争的发展模式。但郑京平和蔡定剑从不同方面提出了不同看法（参见张五常：《中国的经济制度》，北京：中信出版社，2009年，第24—28页；郑京平：《中国“县际竞争”发展模式的利弊分析》，《中国国情国力》2010年第9期；蔡定剑：

况。中国建设法治政府主导模式和县际竞争的现实为法治评估的推广提供了直接的动力。国务院法制办、司法部正在探讨出台在全国范围内进行法治评估的指导性意见，但最终还是落脚于各地的评估。

第五，法治评估方法以定性描述为主。国际上相对成熟的法治评估实践都根据一定的计算方法收集数据进行量化统计，并最后得出较客观准确的数字。中国内地的法治评估则更多停留在定性研究阶段，很多法治评估实践仍然属于传统计划经济思路下的拟定指标作业，尚未进入量化阶段。

（二）中国法治评估需关注的问题

第一，政府主导法治评估可能对中国法治评估的指标设置产生实用性导向。如北京市指标体系框架中的三项一级指标中，“政府工作职能”指标占据主导地位，指标设置直接服务于依法治市的需要。这样的指标设计担当着为政府提供信息和直接服务政府决策的任务，被赋予了太多工具性的期望。法治评估的作用主要应是评价现状、揭示问题，而非解决问题。

第二，由政府规划的指标体系可能受政府政策的制约，具有较强的实用导向性，受地方政府的行为机制等因素的影响。一方面，由政府主导介入的法治评估“由于能得到政府和各部门的协助”，^①易于开展，更具可持续性。另一方面，政府规划的指标体系也容易限制法治评估的作用，指标设置上普遍存在将法治与依法行政混同，将法治简单化为依法行政的倾向，尤其是对公众参与的内容回应不够。公众参与有利于评估过程的透明化和考核内容的问责化，并增强评估过程和结果的可验证性。对于具有实质内容的法治评估，强调公众参与，使得公众对于关系自身切身利益的事项具有一定的决策评价可能，也有利于在评估中强化政府的服务角色和职能。

第三，法治评估对法治的广义界定，反映了中国法治建设对提升其他配套制度的需要。法治的建设和推进需要依托政治、社会和经济的整体发展。在考虑法治评估意义和作用时，对政府主导型法治评估的特点扬长避短，找到一个合适的平衡点十分关键。

第四，评估区域化过程中指标设置、评估方法等的多样性，一方面反映了各地法治发展水平和对法治推进关注点的差异和侧重，另一方面又限制了各地比照借鉴的可能。这种不同标准之间的考核甚至评分，由于标准不同、缺乏参照系，在很大程度上只能局限于对本地区法治水平的历时性考察，而不能展开横向比对，其效用

^① 《理性看县际竞争》，<http://news.sina.com.cn/pl/2008-10-11/081316435178.shtml>，2011年11月3日。

^② 王称心、蒋立山等：《现代化法治城市评价——北京市法治建设状况综合评价指标体系研究》，北京：知识产权出版社，2008年，第79页。

受到极大的限制。^①

第五，深入法治评估方法本身，还需要考虑评估的科学性和有效性问题。

首先是指法治指标设计和评分比重的科学性问题。从当前中国不同地区的法治评估实践来看，有两个突出问题：其一是除香港和余杭外，其他地区总体上是将政府、党委、法院等机关的依法行为作为指标的基本内容，较少涉及相对人部分（诸如公众参与、民主决策等内容）。党政机关人员对法律法规的制定执行状况有更详细的了解，对于法治也有更专业化的认识，这种侧重固然有利于对法治的精确评估，^②但法治的内容远为广泛和复杂，尤其是，法治的基本理念最终是以人为本，以被管理对象和社会为本的。指标设置上存在将法治与依法行政混同，将法治简单化为依法行政的倾向，这种法治指数能否全面反映法治建设的内容是存疑的。其二，内地的法治评估实践中都引入了人数可观的专家研究设计法治评估体系，各地基本上都采取了“专家主导、政府支持”的模式。这种模式的结果是，法治评估中专家在评分比重中占据较大的份额。^③在相当一部分专家没有全程参与或者对评估地区的实际了解较少的情况下，占据如此重要的评分比重，可能会在一定程度上影响评估结果的科学性，^④并可能会引起民众感觉与评估结果之间的脱节。

其次是法治评估结果的有效性问题。从宏观角度讲，中国内地的法治评估都是在政府支持下进行的，而政府支持本身又在一定程度上是县际竞争的产物。如何保证法治评估不以一种新形式旧内容的方式出现，就成为影响其有效性的重要内容。从微观角度讲，法治评估（指数）的目的和价值在于为法治的进展提供一份可资参考的“数据资料”，评估结果显示的不足之处应该成为完善法治建设的重要内容。因此，实现相应的有效反馈与实践应用是影响法治评估有效性的重要方面。

① 这一问题已受到中央政府的注意，统一各地法治评估指标（目前主要是围绕法治政府建设方面）的尝试已经初步展开（参见《国务院办公厅关于推行法治政府建设指标体系的指导意见（讨论稿）》，<http://fzxx.gansudaily.com.cn/system/2009/12/10/011384875.shtml>，2011年11月3日）。

② 西方社会政府管理部门也强调“内部人员”对法律法规更为了解这一优势（参见 Evert Vedung, *Public Policy and Program Evaluation*, New Brunswick, NJ: Transaction Publishers, 1997, p.117）。

③ 例如，余杭法治指数中“专家组”评分占全部分值的30%（参见余杭法治评估体系课题组：《余杭法治指数报告》，《中国法治发展报告 No.7（2009）》，第366页）。

④ 从信息角度讲，专家提供的是理性信息，而民众提供的是价值性信息；在价值性信息领域，相比于民众，专家并没有任何优势。在法治评估领域，大量信息属于价值性信息而非理性信息（参见王锡锌：《公众参与和行政过程——一个理念和制度分析的框架》，北京：中国民主法制出版社，2007年，第195页）。

(三) 法治评估在中国应用的预测

中国法治评估运动既是对全球化的回应，也是中国特色市场经济发展、国内区域“县际竞争”的产物。在建设社会主义法治国家成为基本共识的前提下，法治评估所具有的工具理性特色与政府推进型的法治建构模式具有天然的亲和性，以地区为基础展开的地方法治评估应在一定范围内伴随法治建设的推进得到推广和应用。

第一，法治评估可以成为当前推进改革和社会管理创新的一个重要手段。随着群众对政府期望值的升高，中国深化改革的难度也在加大。这种情况下，从社会管理的技术层面推进的法治评估，更容易被利益相关方接受，进而有益于推进法治建设。当前社会管理的加强和创新主要体现为从政府的“包打天下”，到注重社会力量的运用和社会合力的形成。法治评估的特点是政府和公民形成合力推动社会“善治”的发展。法治评估如果能避免自测自评式的政绩考核模式，扩大公民参与，就能形成社会管理创新的合力。

第二，法治评估是中国法治建设的重要增长点。法治发展的具体路径有建构式和进化式两种，前者认为法治秩序可以通过人们的理性、主观努力而建构出来；后者认为法治秩序是不能建构的，只能通过社会的自然演进而逐渐成长。实际上，当代中国的实践是二者共存。^①在这两种路径的共同作用中，经济和社会发展水平较高的东部地区，已经开始率先推进区域法治化。在这些先行地区建构法治的过程中，法治评估既为其提供检验依据，也为其不断完善法治建设指明了方向和重点，更为后进地区和整个国家的法治建设提供了重要模版和直接动力。^②

第三，法治评估有利于推动法治政府的建设。政府绩效评估是公共管理学领域的传统议题，从法学角度看，规范政府权力的行使也始终是法治的重要内容，在这方面二者具相同的目标。法治涉及的内容比政府工作考核的范围更广，法治评估可以借鉴公共管理学的已有成果，更全面评估政府绩效，推动法治政府的全面发展。

有学者指出，可以在现行政治考核体系的经济、社会、人文等指标基础上增加一个法治指标，或称“法治GDP”，使得“各级政府和官员就不会唯经济马首是瞻，推行法治就会成为各级官员的理性选择”。^③当一个地方将经济与法治并重时，对“法治GDP”的追求将成为一种新的政绩而受到关注。同时，地方政府在推进法治

-
- ① 孙笑侠、钟瑞庆：《“先发”地区的先行法治化——以浙江省法治发展实践为例》，《学习与探索》2010年第1期。
- ② 钱弘道：《先行法治化与法治评估的创新实践》，孙笑侠等：《先行法治化——“法治浙江”三十年回顾与未来展望》，杭州：浙江大学出版社，2009年，第312—314页。
- ③ 此观点由中国政法大学马怀德教授提出（参见袁祥、王逸吟：《中国法治政府荣获好评 专家吁法治指标列政绩考核》，<http://finance.ifeng.com/roll/20110210/3372826.shtml>, 2011年11月3日）。

评估的过程中应避免和克服“法治政绩工程”。

第四，中国建设法治政府主导模式和县际竞争的现实为法治评估的推广提供了直接的动力。首先，政府主导模式是一定时期内法治实践需要的产物。2004年国务院在《全面推进依法行政实施纲要》中提出：“全面推进依法行政，经过十年左右坚持不懈的努力，基本实现建设法治政府的目标”，地方政府存在推动有效法治建设的现实压力。其次，在GDP考核标准的重要性不断下降、民生标准不断上升的现实竞争体系中，法治评估作为地方竞争优势的一个重要方面，存在不断扩展的可能性。在现实的法治建设中，“县际竞争”的动力机制决定了县、区级国家机关有动力进行法治评估等地方法治创新活动，获取更多的比较优势，从而在推动中国法治建设的同时，为中国法治建设提供更可行的目标和指向。

最后，法治评估为中国法学研究成果应用于实践，以及学术研究领域的拓展提供了新思路。随着法治目标的确立，如何有效推进法治建设成为学术界研究的焦点。地方法治评估为学术界研究理论成果的应用及检验提供了绝佳的平台。随着地方法治评估的不断完善和“试错”，学术研究与实践操作之间的互动有助于提高学术研究的指向性和研究水准。

在具体研究方法上，作为法治评估更客观的量化方式，法治指数研究是法学跨学科研究的一个成果。近年来，中国法学开始走出“政法话语”和法条分析的传统，开始注重实证研究和交叉学科，这一方法论转向可能为中国法学研究的发展提供一个新的突破口。但由于法治评估尚处于探索和“试错”阶段，其科学模式和未来走向最终需要反复的实践给出答案。

〔责任编辑：刘 鹏 责任编辑：张芝梅〕

bond markets in China, the United States, the United Kingdom and Japan. We found that investor preference in the market explains endogenous time-varying systemic risk, with macro-economic changes constituting only an indirect contributing factor. The economic interpretation of the endogeneity of systemic asset risk shows that the CAPM theory has shortcomings in actual application.

(6) Returns to Human Capital and Social Integration of Migrants in China

Xie Guihua • 103 •

This research divides urban workers into four groups according to their *hukou* (household registration) and migration status. From the perspective of self-selection and skill transferability theories, we examine and compare each group's rates of returns to human capital, and then discuss possible economic integration between migrants and local workers. The research finds that positive self-selection works for migrant groups regardless of *hukou* status; migrant workers with non-agricultural *hukou* encounter no obstacles in economic integration, and they even perform better than local urban workers in terms of earnings and rates of returns to human capital; and rural migrant workers have less earnings in the early days after their migration, but after a period of time, this group divides; highly skilled workers can catch up with the earnings of and reach economic integration with local workers; while poorly skilled migrant workers have to stay in inferior income positions. Therefore the integration of migrant workers in urban society is a kind of segregated integration.

(7) The “Publicity” Dilemma in the Social Construction of Contemporary China and How It Can Be Transcended

Li Youmei, Xiao Ying and Huang Xiaochun • 125 •

“Publicity” is one of the important goals and supporting conditions of social construction, and the establishment of publicity-based new social systems provides a foundation for effectively advancing social construction. In recent years, the construction of democratic systems and implementation of policies intended to facilitate the development of social organizations has had clear positive effects on the growth of publicity. However, this remains a major bottleneck for current social construction in China. Apart from some institutional factors, social psychology, the absence of consensus, and the increasingly technical character of management also constitute important obstacles to the development of publicity. Mutual trust between public organs and the population needs to be strengthened in order to promote the growth of publicity and the vigorous participation of different interest groups in social construction.

(8) Assessment of the Rule of Law and Its Application in China

Qian Hongdao et al. • 140 •

Assessment of the rule of law has emerged in response to the demand for assessments of the current state and future development of the rule of law. The Rule of Law Index in the World Bank Global Governance Indicators and that in the World

• 207 •

Justice Project and their related practice have encouraged some countries and regions to conduct quantitative assessments of the rule of law at a broader level. Assessment of the rule of law in China has gone through two stages: assessment and inspection of the legislative and judicial domains, followed by the comprehensive exploration of assessments of the rule of law. These assessments are characterized by “embeddedness,” government dominance, an increasingly broad understanding of the rule of law, and localization. By means of trial and error, the rule of law in China should gradually normalize, systematize and institutionalize local assessment and facilitate the achievement of local government targets with regard to the rule of law as a new growth point for construction of the rule of law at the local level.

(9) The Russian Consul-General and the Qing Customs Intendant at Tianjin: Sino-Russian Negotiations at the Local Level as Seen in Block-printed Materials

Chen Kaike • 161 •

The rare block-printed historical materials, *Arguments in the Case of the Collision of the Futong Tea Ship* and *Letters from the Russian Consul-General* include letters between the Russian Consul-General and the Qing Customs Intendant addressing Sino-Russian disputes during the Tongzhi reign (1862-1874) at the trading port of Tianjin. They provide a panorama of negotiations at the local level. Local negotiations were conducted under the treaty system arrived at through national-level negotiations; they constituted a fundamental component of Sino-Russian relations. Not only did they have distinct modes of operation, they were also necessary for the implementation of the many unequal stipulations in the treaty system. The difference in the negotiating principles of the two parties revealed in the course of specific negotiations demonstrates the different diplomatic status overall of the two nations of China and Russia in the East Asian region.

(10) Fragile High Frontier: The Strategic Dilemma of U.S. Space Deterrence in the Post-Cold War Era

He Qisong • 183 •

In the Cold War period, space assets constituted an important component of the U.S. nuclear deterrence. In the Post-Cold War era, however, space deterrence has been separated from U.S. nuclear deterrence and become a relatively independent deterrent force and amplifier of military capability. The United States has invested heavily in developing its space capacity, which has in turn resulted in a high degree of dependence on space. However, the innate defect of space assets is their relative fragility or vulnerability to attacks. This dependence and vulnerability have brought about the fundamental dilemma of U.S. space deterrence. With the spread of anti-satellite technologies, the United States has become increasingly concerned with forestalling enemy attacks on its space assets and has tried to construct a theory of space deterrence. In terms of practice, it has taken a broad range of measures including establishing international norms for space activities, forming space partnerships with allies, strengthening U.S. flexibility and capacity for fighting in harsh environments, formulating contingency plans, etc. However, U.S. space deterrence faces strategic dilemmas in both theory and practice which will not be easy to resolve.

• 208 •