



《北大法律评论》(2019)

第 20 卷 · 第 2 辑 · 页 311—332

Peking University Law Review

Vol. 20, No. 2, 2019, pp. 311-332

数据与分权：多层治理体系中的法制绩效评估^{*}

凯文·E. 戴维斯^{**} 王美舒^{***} 译

Data and Decentralization: Measuring the Performance of Legal Institutions in Multilevel Systems of Governance

Kevin E. Davis Trans. by Wang Meishu

一、引言

在奥巴马的最后任期，他的政治对手抓住每一个可能的机会，对其任期内的表现进行批评。其中一个有意思的批评责难道，奥巴马政府竟然允许美国在

* 原文载于《明尼苏达法律评论》(*Minnesota Law Review*)2018年第4期(总第102卷)，感谢香港中文大学法学院乔仕彤老师、华东师范大学法学院田雷教授对原作的引介。为便于读者索引，译文保留原文引征格式和层次。——译者按

** 凯文·E. 戴维斯(Kevin E. Davis)，纽约大学法学院教授。作者感谢“分权与发展研讨会”(Conference on Decentralization and Development)的参会者所给予的评议，感谢与费尔南多·德尔加多(Fernando Delgado)、马丁·西布利斯(Martin Syblis)、阮梅恩(Mae Nguyen)等人的交流，感谢阿瓦·哈格海伊(Ava Haghighi)和罗德里戈·梅拉(Rodrigo Mella)所提供的研究辅助工作，并感谢菲洛门·达戈斯蒂诺与马克斯·E. 格林伯格研究基金(Filomen D'Agostino and Max E. Greenberg Research Fund)提供的资助。

*** 王美舒，法学博士，华东师范大学法学院晨晖学者，讲师。

世界银行发布的营商环境便利度排名中下降。《华尔街日报》的一位专栏作家指出,奥巴马总统上任后,美国的综合排名从第三位跌至第八位。^{〔1〕}他还抱怨道:“八年前,市场主体拿到施工许可证需要40天,现在却需要81天之久。布什总统离任时,合同执行仅需300天,可现在,却花了420天。产权登记成本自2009年以来从0.5%上升到了2.4%,翻了近5倍。”^{〔2〕}

世界银行的营商环境指标有很多优点。但显而易见,无论谁当总统,它都不是用来评价美国联邦政府绩效的好方法。原因在于,美国采取多层治理体系,总统只是该体系中的一部分,而营商环境指标主要抓取该体系中其他部分的表现。特别是营商环境评价仅以各国一到两个最大商业城市的数据为基础。^{〔3〕}就美国而言,两个城市被作为样本:纽约市与洛杉矶市。^{〔4〕}美国总统负责的事很多,却无法控制在纽约市或洛杉矶市获得施工许可证,要花多长时间。美国总统也无法控制在纽约或洛杉矶市进行产权登记的具体费用,更无法控制在纽约高等法院或加州高等法院执行合同要花费多少时间。施工许可证由市政府发放;合同执行则在法院职责范围内。至于产权登记的花费,在纽约市则由手续费、税以及国家、市、州三级所设收费项目决定。^{〔5〕}营商环境评价所用的10个国家层面指标中,只有3个与美国联邦政府控制的事项有关,即纳税、跨境贸易和办理破产。^{〔6〕}在这3个指标中,又因纳税指标包含市、州和联邦纳税项目,其得分仍不完全受联邦政府控制。^{〔7〕}

世界银行营商环境指标虽不在其列,但仍有一些评估能够在多层体系中测评各行政单位的表现,并于要求行政单位承担责任、维持和改善绩效中发挥重要作用。没有准确的绩效评估,利益相关者很难坚称某个行政单位,甚或某类行政单位,需要进行改革或被替代。绩效评估适用于内部和外部执行者——若执行者不知道自己的表现如何,那么即使是最富动力的行政官员,也因缺乏理性基础,而不知如何更好履行自己的职责。

在多层级治理体系中对法律制度(legal institutions,又译作“法制”)——

〔1〕 Bret Stephens, *Doomed to Stagnate?*, WALL ST. J., Dec. 20, 2016, at A21.

〔2〕 *Id.*

〔3〕 WORLD BANK GRP., DOING BUSINESS 2017: EQUAL OPPORTUNITY FOR ALL 13 (2017), <https://www.doingbusiness.org/content/dam/doingBusiness/media/Annual-Reports/English/DB17-Report.pdf>.

〔4〕 *Id.* at 163.

〔5〕 WORLD BANK GRP., DOING BUSINESS 2017: EQUAL OPPORTUNITY FOR ALL, ECONOMY PROFILE 2017 UNITED STATES 34—45, 65—68, 103 (2017) (showing data on processes for construction permits, registering property, and enforcing contracts), <https://openknowledge.worldbank.org/bitstream/handle/10986/25643/WP-DB17-PUBLIC-United-States.pdf?sequence=1&isAllowed=y>.

〔6〕 *See id.* at 86—99, 109—15.

〔7〕 *Id.* at 90—92.

颁布、执行规范的制度,进行绩效评估尤为重要。但是由于下述原因,找到准确进行绩效评估的方法并非易事。对体系中的基层法律制度进行准确评估,更是难上加难。毕竟基层机构与体系中引人注目的国家立法机关、宪法法院相距甚远。^{〔8〕}

可能限制法制绩效评估的因素有两个。第一,很难将基层法律制度的影响与其他层级治理的影响区分开来。第二,市场不一定会推动人们费心去建构区分二者的绩效评估方法。分类绩效评估(disaggregated measures)的成本可能与综合绩效评估(aggregated measures)的成本一样高——甚至更高,而且还会在绩效评估的精确性上设限。

如果以上观点皆正确,那么某些形式的多级治理和法制分权(legal decentralization),就有了不可忽视的缺点。许多为多层治理辩护的观点假设,各利益相关者,无论是选民、潜在居民、潜在投资者、高层政府官员,抑或公务员职业网络中的一员,都能够对各行政单位进行绩效评估。没有合适的评估方法,这些利益相关者都无法对政府官员产生积极影响。关于治理分权的学术研究对影响基层政府问责的因素多有关注^{〔9〕},但对法制绩效评估的有效性研究不够。大量研究探讨了建构法制绩效评估所面临的挑战,但迄今,没有专门针对多层治理体系中基层制度绩效评估的研究。^{〔10〕}

二、法制绩效评估的价值

世界上的大多数人都在政府和非政府的多重治理之下。在中国,除了国家政府之外,还有省、市、县、镇、村,五级管理制度的存在。^{〔11〕}在欧盟,单个成员国的政府层级一般少于中国,但还需考虑到欧盟超国家层面的治理。^{〔12〕}事实

〔8〕 For a notable exception, see Todd Mitton, *The Wealth of Subnations: Geography, Institutions, and Within-Country Development*, 118 J. DEV. ECON. 88 (2016) (studying data covering subnational regions from 101 countries).

〔9〕 See generally Dilip Mookherjee, *Political Decentralization*, 7 ANN. REV. ECON. 231 (2015) (providing an overview of the literature on political decentralization).

〔10〕 See, e.g., Kevin E. Davis, *Legal Indicators: The Power of Quantitative Measures of Law*, 10 ANN. REV. L. & SOC. SCI. 37 (2014) (discussing key issues related to legal indicators measuring aspects of a legal system). On the general importance of studying political phenomena at the appropriate geographic level, see Clark C. Gibson et al., *The Concept of Scale and the Human Dimensions of Global Change: A Survey*, 32 ECOLOGICAL ECON. 217, 233—36 (2000).

〔11〕 SEBASTIAN HEILMANN ET AL., CHINA'S POLITICAL SYSTEM 59—60, 76—104 (Sebastian Heilmann ed., 2017).

〔12〕 See *Countries*, EUROPA, https://europa.eu/european-union/about-eu/countries_en (last visited Apr. 26, 2018); *The EU in Brief*, EUROPA, https://europa.eu/european-union/about-eu/eu-in-brief_en (last visited Apr. 26, 2018).

上,世界上任何人的生活都切实地受到国际组织、国外政府域外法权的影响。^[13] 民族国家之外的机构在治理贫穷弱小国家方面,尤其是在治理那些内部治理能力有限的国家方面,发挥着重要作用。比如,世界银行在为基础设施项目设立社会与环境标准中所起的作用。^[14]

法律制度,即公布、实施或执行规范的制度,当然是各级行政单位的应有之义。典型的多层治理系统包含一级以上的立法机构。在印度,除国家议会外,每个邦(state)都有自己的民选立法机关。至少在农村地区,有多达三个级别(县、区、村)的潘查亚特(panchayats)^[15]*。相当大的权力下放给通过选举而成的村级潘查亚特。^[16] 法律由所有层级的行政机关,以及国家警察、邦警察负责实施与执行。^[17] 印度司法制度也区分了层级,最高法院位于各邦高级法院与县法院之上。^[18] 相对非正式的法庭,一般称作“人民法庭”(lok adalats)。它平行于正式的纠纷解决形式,处理日常争端。^[19]

对这些法律制度进行绩效评估意味着什么?这其实是对“治理”的评估,而非对“法治”的评估。不管这些行为是官员自身所欲为的(如,因为他们反对对嫌疑人严刑逼供),还是有意为之(如贿赂的减少、污染水平的降低,或是地价的

[13] For a discussion of how international organizations have impacted international law, see JOSÉ E. ALVAREZ, *INTERNATIONAL ORGANIZATIONS AS LAWMAKERS* (2005).

[14] Benedict Kingsbury, *Operational Policies of International Institutions as Part of the Law-Making Process: The World Bank and Indigenous Peoples*, in *THE REALITY OF INTERNATIONAL LAW* 323, 325 (Guy S. Goodwin-Gill & Stefan Talmon eds., 1999); see also ALVAREZ, *supra* note [13], at 235—41.

[15] See INDIA CONST. art. 243-O, amended by The Constitution (Seventy-Third Amendment) Act, 1992; M. GOVINDA RAO & NIRVIKAR SINGH, *POLITICAL ECONOMY OF FEDERALISM IN INDIA* 41—61, 93—119, 253—79 (2005).

* 根据印度宪法,潘查亚特有三个级别,其中村级潘查亚特基本由各村村民直接选举产生,负责印度农村基层治理的各项事务。每一个村级潘查亚特设有一位主席、五个委员会。每个委员会主管一个关系农村发展的事务,每年召开一次会议,根据选民的要求制定《年度行动计划》。——译者注

[16] INDIA CONST. art. 243-O, amended by The Constitution (Seventy-Third Amendment) Act, 1992; RAO & SINGH, *supra* note [15].

[17] RAO & SINGH, *supra* note [15], at 111—14.

[18] *Id.* at 105—11; Rakesh Kumar Shrivastava, *UPDATE: A Guide to India's Legal Research and Legal System*, N. Y. UNIV. L. GLOBAL (Apr. 2014), http://www.nyulawglobal.org/globalex/India_Legal_Research1.html.

[19] See Marc Galanter & Jayanth K. Krishnan, *Debased Informalism: Lok Adalats and Legal Rights in Modern India*, in *BEYOND COMMON KNOWLEDGE: EMPIRICAL APPROACHES TO THE RULE OF LAW* 96, 108—10 (Erik G. Jensen & Thomas C. Heller eds., 2003).

增加)。^[20] 本文还假设绩效报告以指标形式呈现:以简易形式抓取绩效的某一特定方面,并进行排序的一组数据。^[21] 换言之,本文集中讨论法律指标*。

绩效评估因其影响官员表现的方式,而具有重要的现实意义。这种语境下的“影响”,可以意味着引导官员选择某一规范,调动某种能力,采取某些行为,或将不愿合作的官员替换成愿意合作的官员。对不同影响路径类型化的方法之一,是按影响机制进行分类。^[22] 激励式机制显然依赖于绩效评估。这种机制是依过去的表现进行奖惩,进而预期效果出现的影响机制。为预防起见而以过去工作情况来分配未来行政职责的体系,也依赖绩效评估。其他影响机制,如同侪压力、理性说服、内在驱动,也都可能依赖绩效数据。

另一种类型化绩效评估影响路径的方法,是按照绩效评价的潜在用户来分

[20] See Roderick M. Hills, Jr. & Shitong Qiao, *Voice and Exit as Accountability Mechanisms: Can Foot-Voting Be Made Safe for the Chinese Communist Party?* 48 COLUM. HUM. RTS. L. REV. 158, 180—83 (2017) (arguing that land desirability is “influenced by local officials’ decisions about education, traffic regulation, crime control, and overall governmental efficiency, which are reflected . . . in real estate prices”). Land is not the only good whose value might be highly sensitive to performance of legal institutions. For instance, the prices of illicit drugs might be influenced by drug control laws and how they are enforced. See Jonathan P. Caulkins & Peter Reuter, *How Drug Enforcement Affects Drug Prices*, 39 CRIME & JUST. 213 (2010). In theory, the prices of contracts that pay out when an illegal act occurs—for example a terrorist attack—could convey similar information about law enforcement, but such markets would create perverse incentives.

[21] See generally Kevin E. Davis et al., *Indicators as a Technology of Global Governance*, 46 LAW & SOC’Y REV. 71 (2012) (defining and analyzing indicators in the context of global governance); Davis, *supra* note [10].

* 作者凯文·E.戴维斯在书评《法律指标:法律定量评估的力量》中指出,“Legal indicators, sometimes known as rule of law indicators, governance indicators, or indicators of public service performance, are quantitative measures of the performance of legal systems.”因此,在作者的叙述中,legal indicators 与法治指标(rule of law indicators)、治理指标(governance indicators),或者公共服务绩效指标(indicators of public service performance)同义,都是对法律体系绩效的定量评估。本译文将文中 legal indicators 翻译为“法律指标”。参见 Kevin E. Davis, “Legal Indicators: The Power of Quantitative Measures of Law”, *Annual Review of Law and Social Science*, Vol. 10, 2014, pp. 37-52. ——译者注

[22] Cf. WOUTER VAN DOOREN ET AL., PERFORMANCE MANAGEMENT IN THE PUBLIC SECTOR 118—22 (2015) (classifying uses of performance information as learning, steering and control, and accountability and explaining that accountability for performance exerts reputational, market and political pressure on people in organizations); Gerhard Hammerschmid et al., *Internal and External Use of Performance Information in Public Organizations: Results from an International Survey*, 33 PUB. MONEY & MGMT. 261, 262 (2013) (describing internal and external uses of performance information).

类。^[23] 在民主制国家中,最明显的潜在用户是选民。他们用手中的投票权来促进好的政府行为,防止坏的政府行为。在等级制国家中,政府高层官员运用基层政府的绩效信息,来决定如何行使其监督权。^[24] 政府官员也根据绩效数据,决定对自己的同僚进行表扬还是批评。^[25] 自我激励的政府会运用同样的数据,决定是否改进,以及如何改进其绩效。^[26] 政府债权人会运用与财政相关的绩效数据,来进行信用评估。^[27] 最后,也是最重要的,潜在定居者和投资者会根据过去的绩效数据来预测政府未来的表现,以决定是否将自己置于该政府管辖之下。对那些政府官员有动机去吸引的潜在定居者而言,尤为如此。^[28]

无论绩效评估如何完成,由谁使用,比较数据都会发挥重要作用。比较数

[23] See generally VAN DOOREN ET AL., *supra* note [22], at 96—115 (describing public managers, politicians, members of Parliament, and citizens/customers as users of performance information).

[24] See, e.g., Manuele Citi & Martin Rhodes, *New Modes of Governance in the European Union: A Critical Survey and Analysis*, in HANDBOOK OF EUROPEAN POLITICS 463, 466—67 (Knud Erik Jørgensen ed., 2007) (describing use of indicators to monitor European Union member states' pursuit of objectives established at the European level); Jie Gao, *Pernicious Manipulation of Performance Measures in China's Cadre Evaluation System*, 223 CHINA Q. 618 (2015) (documenting use of performance measures to motivate local officials by Chinese central authorities as well as various ways of gaming of the measures); Jie Gao, *Governing by Goals and Numbers: A Case Study in the Use of Performance Measurement to Build State Capacity in China*, 29 PUB. ADMIN. & DEV. 21 (2009) (documenting use of performance measures to motivate local officials by Chinese central authorities); Jonathan Zeitlin, *Conclusion: The Open Method of Co-ordination in Action: Theoretical Promise, Empirical Realities, Reform Strategy*, in THE OPEN METHOD OF CO-ORDINATION IN ACTION: THE EUROPEAN EMPLOYMENT AND SOCIAL INCLUSION STRATEGIES 447 (Jonathan Zeitlin et al. eds., 2005) (same).

[25] See, e.g., Citi & Rhodes, *supra* note [24], at 469—72 (discussing peer review and peer pressure within the European Union).

[26] See, e.g., U. S. GOV'T ACCOUNTABILITY OFFICE, GAO-14-747, *MANAGING FOR RESULTS: AGENCIES' TRENDS IN THE USE OF PERFORMANCE INFORMATION TO MAKE DECISIONS* (2014) (presenting survey data on use of performance measures within US federal government agencies); VAN DOOREN ET AL., *supra* note [22] (use by public managers, politicians and members of Parliament); Pedro Rubim Borges Fortes, *How Legal Indicators Influence A Justice System and Judicial Behavior: The Brazilian National Council of Justice and "Justice in Numbers"*, 47 J. LEGAL PLURALISM & UNOFFICIAL L. 39 (2015) (describing use of performance measures by Brazil's National Council of Justice).

[27] Ruoying Chen, *Preserving Decentralization with Chinese-Style Markets* 10—11 (Mar. 5, 2017) (working paper) (on file with *Minnesota Law Review*).

[28] ILYA SOMIN, *DEMOCRACY AND POLITICAL IGNORANCE: WHY SMALLER GOVERNMENT IS SMARTER* 119—20 (2013) (discussing how mobile individuals and firms—"foot voters"—have incentives to shop around among governments for the best combination of public services); Hills & Qiao, *supra* note [20], 173—79 (2015) (surveying literature on foot voters); Charles Tiebout, *A Pure Theory of Public Expenditures*, 64 J. POL. ECON. 416 (1956) (seminal article arguing that governments respond to foot voters).

据能够将单个行政单位与其过去绩效,或与其他单位的绩效相关联,进行比较分析。^[29] 作为标尺,比较数据能够洞察到一些在单独评估中无法获得的绩效结果。^[30] 没去过香港的纽约人很难想象,会有比纽约地铁更干净、更安全、更可靠的地铁系统。比较数据还有助于评估政府行为。例如,对治安策略与犯罪率进行跨时或跨区域比较,可以用于发现何种治安策略最佳,以及哪些政府选择这种策略。

不同类型的潜在使用者会对不同的绩效评估感兴趣。像选民、潜在定居者这样的大众群体,通常没有时间,也没有专业知识,对政府绩效进行分析。这些使用者倾向于寻找信息量相对大,但又高度简化的指标。世界银行的营商便利度指数就是高度简化的极佳例子。这个指数意图用一个数字将一国内所有与营商便利相关的法律体系的表现情况呈现出来。^[31]

三、地方绩效评估之实践

本文聚焦于可以用于评估多层治理体系中各法律制度的方法,尤其是地方层面法律制度。自20世纪初以来,各种法律指标层出不穷。但其中最著名的例子,也仅为每个国家打出一个分数。这显然不适合用来评估地方机构的绩效。以世界银行营商便利度指数及构成它的单个指标为例:单个指标测量了国内中小企业完成11个环节的能力,包括注册企业、登记产权、获得施工许可证、执行合同以及纳税。^[32] 指标聚焦于评估完成相关事项所需时间与花费,虽然一些指标也评估法律制度实施的质量。^[33] 用于计算营商便利度指数的营商环境指标,通常根据一国“最大商业城市”的数据,给出该国的得分。^[34] 但为公平起见,11个较大国家(包括美国)的得分,则根据人口比重,加权平均两个最大商业城市的分数而得。^[35]

另一套著名的法律指标是世界正义工程(World Justice Project, WJP)法

[29] See Timothy Besley & Anne Case, *Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition*, 85 AM. ECON. REV. 25, 25 (1995) (discussing “yardstick competition” and how comparative data works in politics); Pierre Salmon, *Decentralisation as an Incentive Scheme*, 3 OXFORD REV. ECON. POL’Y 24, 30—32 (1987) (discussing value of comparative data in evaluating performance of governments).

[30] *Id.*

[31] *Economy Rankings*, DOING BUSINESS, <http://www.doingbusiness.org/rankings> (last visited Apr. 26, 2018).

[32] WORLD BANK GRP., *supra* note [3], at 13—14.

[33] *Id.* at 14.

[34] *Id.* at 13.

[35] *Id.* at 168.

治指数(Rule of Law Index)。^[36] 这一指数旨在衡量“普通人感受到的法治状况”。^[37] 它的数据源有两个:(1) 每个国家的普通人群调查;(2) 国内民商事、刑事诉讼、劳动法及公共健康专业人士调研。^[38] 普通人群调查在每个国家最大的三个城市进行,但国内专家的调研地点不是特定的。^[39]

世界银行营商环境指数与世界正义工程法治指数至少说明它们在各国所刻画的部分。其他法律指标则更为不透明。例如,世界银行创建了一套含6个国家别层面指标的评估体系,称为“全球治理指数”(Worldwide Governance Indicators, WGIs)。^[40] 其指标所评估的治理方面,即治理效率、规制质量、法治,看上去都与法制绩效相关。^[41] 这个评估毫无疑问受到国家与地方政府的影响。但全球治理指数的设计者没有说明哪一层级地方政府是评估的数据源,当然也是因为设计者们不知道。全球治理指数作为一个综合指标体系,基于30多个基本数据源形成评估。其数据来源包括家庭调查、企业调研,以及专家评估。^[42] 被测量的企业与专家并不总是样本国之内的。^[43] 极有可能,全球治理指数有时测量了人们对不同地方机构的认知——如审理国外大公司与国内中小企业纠纷的法院,有时又测量了人们对多个机构的综合认识。在这些情况下,“全球治理指数”的设计者几乎无法说明他们的指标刻画了哪几个机构。

也有对以国内不同地点进行评分的跨国指标体系。^[44] 此类指标最突出者,是由世界银行营商环境项目组,在其国别层面指标的基础上设计而成的。自2005年起,世界银行营商环境项目组就推出65个国家438个地区的地方性

[36] See WORLD JUSTICE PROJECT, RULE OF LAW INDEX 2017—2018, https://www.worldjusticeproject.org/sites/default/files/documents/WJP_ROLI_2017-18_Online-Edition_0.pdf

[37] *Id.* at 156.

[38] *Id.* at 157.

[39] *Id.* at 157—58 (stating that the population survey is a sample from 1000 people in the “three largest cities of each country” while the in-country professionals are selected through directories and referrals and vetted by the World Justice Project).

[40] Daniel Kaufmann et al., *The Worldwide Governance Indicators: Methodology and Analytical Issues*, 3 HAGUE J. ON RULE L. 220, 220 (2011) (stating that the Worldwide Governance Indicators measure “six dimensions of governance”).

[41] See *id.* at 239 (discussing the use of perception data).

[42] WGI Data Sources, WORLD BANK GRP., <http://info.worldbank.org/governance/wgi/#doc-sources> (last visited Apr. 26, 2018) (stating “The WGI compile and summarize information from over 30 existing data sources” including households, firms, and experts).

[43] See Kevin E. Davis, *What Can the Rule of Law Variable Tell Us About Rule of Law Reforms?*, 26 MICH. J. INT’L L. 141, 150 (2004) (stating that sometimes certain indicators contain subjective assessments “made by individuals who are not necessarily resident[s] in the society to which their assessments pertain”).

[44] See generally Mitton, *supra* note [8] (collecting datasets).

报告。^[45] 这些地方性报告通常包含 3—6 个主题。^[46] 一般包括开办企业、办理建筑许可证、产权登记这三个指标。^[47] 其他出现在地方性报告中的指标有: 执行合同、获得电力、纳税、跨境贸易。^[48] 与国家指标不同, 地方性营商环境指标的推出是不定期的, 而非每年都进行设计。^[49]

地方性评估指标也会由国内执行者制定。在巴西, 教育机构热图利奥·瓦加斯基基金会(Fundação Getulio Vargas, FGV)资助设计了巴西司法信心指数(C Confidence in Brazilian Justice Index, 葡萄牙语为 Índice de Confiança na Justiça Brasileira, ICJBrasil)指标。^[50] 这套指标用于测量公众对法院系统的信任程度。巴西司法信心指数(ICJBrasil)包含了两个子指标: 一个评估了人们对法院系统的综合感受, 另一个则评估了人们对法院解决特定纠纷能力的信赖程度。^[51] 指标基于巴西 7 个州(亚马孙、伯南布哥、巴伊亚、米纳斯吉拉斯、里约热内卢、圣保罗和南里奥格兰德), 以及联邦特区城市人口的调研而形成评估结果。^[52] 各州与特区的数据分别进行报告。中国也出现了类似的法治指标(rule of law indicators)。其中包括, 一个名为中国法治研究院(China Rule of Law Research Institute)的智库机构所设计的指标。^[53]

另一个地区性评估指标的重要例子是“越南省际竞争力指标”。这套指标由越南工商总会(Vietnam Chamber of Commerce)与美国国际开发署(United

[45] See *Subnational Reports*, DOING BUSINESS, <http://www.doingbusiness.org/reports/subnational-reports> (last visited Apr. 26, 2018).

[46] In India, seven topics were covered; “Closing a business” was added to the list of topics frequently covered in other subnational reports. See WORLD BANK GRP., DOING BUSINESS IN INDIA 2009 1 (2009), <http://www.doingbusiness.org/~media/WBG/DoingBusiness/Documents/Subnational-Reports/DB09-Sub-India.pdf>.

[47] See, e.g., WORLD BANK GRP., DOING BUSINESS IN KAZAKHSTAN 2017 at 4 tbl. 1. 1 (2017), <http://www.doingbusiness.org/~media/WBG/DoingBusiness/Documents/Subnational-Reports/DB17-Sub-Kazakhstan.pdf>.

[48] See *id.* at 15 tbl. 2. 1.

[49] There is nothing to prevent anyone from producing their own indicators using the Doing Business methodology. In 2016, the Jiangsu province in China “began to evaluate its 96 counties in 13 cities on their implementation of administrative reform and business environment in 2016, according to the ease of doing business index set by the World Bank.” *Premier Li Urges Nationwide Efforts to Deepen Administrative Reform*, The BEIJING NEWS (May 29, 2017), http://english.gov.cn/premier/news/2017/05/29/content_281475670291580.htm.

[50] LUCIANA GROSS CUNHA ET AL., RELATÓRIO ICJBRASIL: 1º SEMESTRE 2016 [Report ICJBrasil: First Half 2016] 2 (2016), http://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/17204/Relatorio-ICJBrasil_1_sem_2016.pdf?sequence=1&isAllowed=y.

[51] *Id.* at 3.

[52] *Id.* at 4—5.

[53] See *China Rule of Law Government Index System Research Project*, CHINA RULE OF LAW RESEARCH INST., <http://www.ruleoflawchina.com/projec-researches/china-rule-of-law-government-index-system-research-project> (last visited Apr. 26, 2018).

States Agency for International Development)共同开发。^[54] 指标报告了越南63个省份和自治市的得分。^[55] 除了综合竞争力指标,还有10个次级指标覆盖了法治或非法治领域(legal and non-legal topics)。^[56] 其中一些子指标与营商环境指标所覆盖的主题类似,包括新设企业的市场进入成本、获得土地及使用权保障、合规时间投入。^[57] 此外,指标中还包括法制综合信心指标和国有实体偏见指标。^[58] 指标数据来源于对每省企业的调研,并且国内外企业是分别进行调研的。^[59]

上述各绩效评估指标都依赖于调研数据。但原则上,它们可以依赖其他数据源。人们最容易想到的是官方数据,包括审计数据。作为负责监督公共资源使用的机构,联邦政府审计局(Controladoria Geral da União, CGU)设计了巴西城市审计模型。^[60] 每年,联邦政府审计局随机选择城市,对其政府开支进行审计,并在网上公布信息。^[61] 虽然联邦政府审计局的审计并不一定与法制运行相关联,但在发现浪费与腐败方面非常有效。^[62] 印度政府则与非政府组织共同开展针对个人扶贫项目的“社会审计”。^[63] 没有什么理由不去利用可比较的审计数据,来获得关于法制的信息。

四、建构复杂法律体系(legal system)绩效评估的挑战

虽然有几个地方性法制绩效评估(legal performance measures)的例子,但它们未必对特定地方的法制绩效进行了较好测量。事实上,现有的绩效评估极少对多层治理体系中法制绩效评估所提出的挑战进行回应。本节内容详细介绍了这些挑战。

[54] EDMUND MALESKY ET AL., VIETNAM CHAMBER OF COMMERCE AND INDUSTRY & USAID, THE VIETNAM PROVINCIAL COMPETITIVENESS INDEX 2016 (2017).

[55] *Id.* at 15.

[56] *Id.* at 17 (listing indicators).

[57] *Id.*

[58] *Id.*

[59] *Id.* at 16 (stating that 10,037 domestic private firms and 1550 foreign firms responded to the PCI survey); see also Trang (Mae) Nguyen, *Grading Regulators: The Impact of Global and Local Indicators on Vietnam's Business Governance*, 88 N. Y. U. L. REV. 2254, 2272—85 (2013) (discussing the impact of the Provincial Competitiveness Index).

[60] See Claudio Ferraz & Frederico Finan, *Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes*, 123 Q. J. ECON. 703, 707 (2008).

[61] *Id.*

[62] See *id.* (looking at how publicly released audits have affected elections in Brazil).

[63] See, e. g., Nidhi Vij, *Collaborative Governance: Analysing Social Audits in MGNREGA in India*, 42 IDS BULL. 28, 30—32 (2011).

(一) 两类法制绩效评估

法制绩效评估可以分为两大类:(1) 直接对法律制度本身进行的评估;(2) 对可能受法律制度影响而产生的效果的评估。^[64] 直接评估是对诸如以下问题进行测量、报告的方法:法律是否允许所有商业纠纷都可进行仲裁? 每级法院的法官有多少,有何资质要求? 对建筑设计图审查人员的建筑学或工程学学历要求? 审查人的资质要求是什么?^[65] 效果评估的例子包括:获得建筑许可证的时长是多少? 公司对公职人员贿赂行为的披露率是多少? 暴力犯罪率是多少? 有多少针对警察的投诉? 空气质量如何?

上述例子表明,法制绩效评估可以与完全不同的法律制度及其效果相关联。对法律制度本身进行的直接评估,与法律规范(如法律是否允许……)以及法律官员(如有多少法官)有关。^[66] 同样的,一些法律效果评估聚焦于法律官员(legal officials)的行为;另一些则聚焦于法制外行动(nonlegal actors)的行为;还有的则聚焦于诸如贿赂这样的共同行为。^[67]

下文讨论了与上述两大类绩效评估相关的挑战。涉及评估特定类型的法律制度或法律效果的问题,本文暂不详述,留待以后分析。

(二) 法律制度评估

对规范或法律官员能力、行为的评估,相对简单,尤其是当所需信息已经具备的时候。规范通常以法律、条例或命令的形式呈现。法律官员的数量及其学历,也往往因行政目的而被记录。^[68] 此外,约束行政许可、行政执法机构的实践的规则,也记载于手册、说明书和指引中。^[69]

以未记录的信息为基础,制定绩效评估指标相对较难。比如,法律官员一般援引法律、条例、指引等书面规范,来决定特定行为应有何种法律后果。但在另外一些情况下,法律官员会视个案情况进行裁决,或援引以往案件中隐含的规范进行裁判。事实上,这正是普通法系法官在案件裁判中所做的工作。而

[64] Davis, *supra* note [10], at 39, 42.

[65] *Id.* at 41—42.

[66] *Id.*

[67] *Id.* at 43.

[68] *See, e.g., About California Courts*, CAL. COURTS., <http://www.courts.ca.gov/2113.htm> (last visited Apr. 26, 2018) (providing information about California courts); *Judicial Directories*, N. Y. COURTS, <http://www.nycourts.gov/judges/directory.shtml> (last visited Apr. 26, 2018) (providing information about judges of the New York State Unified Court System).

[69] *See generally Legal Division*, CAL. DEP'T OF BUS. OVERSIGHT, http://www.dbo.ca.gov/Laws_&_Regs/default.asp (last visited Apr. 26, 2018) (detailing various guidelines and opinion letters for corporations and financial institutions in California); *Private Investigator*, N. Y. STATE, <https://www.dos.ny.gov/licensing/privateinvest/privatei.html> (last visited Apr. 26, 2018) (describing licensing requirements for private investigators in New York).

且,我们也有理由相信大陆法系的法官,或其他法系的裁判官也以同样方式操作。^[70] 这种操作方式通过回顾大量判决文书来获得规范,以此规范解释过去的判决,预测未来的裁判。但至少从现有技术水平看,这种操作方式需要相当大的工作量,且得出的结果容易产生分歧。另一个复杂因素是,负责发放许可证和执法的官员可能不愿披露关于其能力和实践的信息,以保留他们相对于被监管方的策略优势。

除了在评估法律制度特点时遇到的实际操作问题,还有一个挑战是要找出哪些特征需要进行评估。允许所有商业纠纷进行仲裁是否可行? 是否坚持由注册建筑师或工程师检查建筑设计图? 是否随机指派法官进行案件审理,而非根据法官经验多少? 只有当这些规范、能力、实践的效果是理想的,它们本身才是可取的。因此,确定是否是好的绩效评估的唯一方法,就是研究它们是否与理想结果有相关性。最终,建构基于效果的绩效评估就成为必需。

(三) 法律效果评估

只有明确哪些机构对结果负责,以效果为中心的评估才能成为衡量具体法制绩效的有效方法。特别是,如果希望绩效评估能够促进某些有利行动,政府应当因其行动所产生的效果,而非其行动控制之外的效果,而受到奖励或惩罚。那些碰巧产生预期效果的投机行为,不应受到鼓励,反之亦然。

识别为给定结果负责的机构,将会面临两个挑战。首先,需要判断是否某一机构自身就可以完全影响所研究的行为。其次,当不止一个机构产生影响时,就有必要区分单个机构的影响。本文中,我们侧重第二项影响。^[71]

在假想的最简单的世界中,每个结果都与单一级政府中的某个机构相关。如果政府各部分的职能划分得严丝合缝,政府的每部分只能影响特定的一组结果,那么多层级政府也可以满足这种情况。可以设想下面这种情形:市政府负责建设,省政府负责安保,国家政府负责空气污染治理。诸多关于分权的文献,就是假定政府职能以此种方式进行分配。^[72]

真实世界远比假想世界复杂得多:不同层级的政府机构都可影响最终的结果,而结果的变化需要从机构组合变化的角度来解释。当公职人员需根据法律

[70] JOHN MERRYMAN & ROGELIO PÉREZ-PERDOMO, *THE CIVIL LAW TRADITION: AN INTRODUCTION TO THE LEGAL SYSTEMS OF EUROPE AND LATIN AMERICA* 39—55 (3d ed. 2007) (comparing use of statutes and prior judicial decisions by judges in common law and civil law systems).

[71] For a recent discussion of the first challenge, see LAWRENCE M. FRIEDMAN, *IMPACT: HOW LAW AFFECTS BEHAVIOR* 1—6 (2016).

[72] See Marcellin Joanis, *Shared Accountability and Partial Decentralization in Local Public Good Provision*, 107 J. DEV. ECON. 28 (2014) (exploring the theoretical implications of blurred accountability); Mookherjee, *supra* note [9], at 235 (noting that few empirical studies have examined the implications of partially decentralized service provision).

对某项活动进行事前审批时,这种情况就会发生。比如,企业注册可能包括在国家层面的政府机构进行注册,并从市级政府机构处获得许可。多层级政府也可能同时参与某项监管工作,包括监管中的事后处罚工作。比如,法律由一级政府负责颁布,而由另一级政府负责实施。巴西就是一个很好的例子。国家立法中所禁止的事项,如反腐败法中所列事项,可以由州或联邦警察队、多个独立机构,以及联邦或州检察官来进行调查。^[73] 根据调查,独立机构或检察官可以提起民事、行政、刑事诉讼。^[74] 诉讼程序会在州法院或联邦法院进行。^[75] 在复杂案件中,不同层级中的多个机构涉案,已经不再罕见。^[76]

多个治理机构共同导致结果出现的这一事实,并不意味着单个机构所发挥的作用无法被追踪。对于任何一个结果,足够深入的历史分析可以揭示问题的根源——如,哪个排放标准过于宽松,哪个政府机构在排污控制的监督上失职,或是哪个法院拖延了对贪污犯的诉讼。^[77] 在对少量结果进行原因分析时,这种分析方法是切实可行的。可是当需要分析的结果变多时,这种方法就迅速变得不切实际。

只要有充足数据,就有可能将涉及不同机构组合的结果进行比较,并推断出各机构所发挥的作用。当下面三个条件满足时,这种比较分析法尤为简单明了。首先,机构的影响范围完全根据地理范围来界定。其次,每个机构的运作,对辖区内人口的影响是一致的。最后,在多层治理体系中,基层机构的影响范围,应完全囊括于高层级机构的影响范围内。当上述条件都具备时,给定的地理范围内的所有结果,都会以相同的方式,受到相同制度/机构组合的影响。而

[73] See Mariana Mota Prado & Lindsey Carson, *Brazilian Anti-Corruption Legislation and Its Enforcement: Potential Lessons for Institutional Design*, 4 J. SELF-GOVERNANCE & MGMT. ECON. 34, 47—49 (2016) (discussing institutional multiplicity in corruption investigations in Brazil).

[74] See *id.* at 49—52.

[75] See Lucas Dotto Borges et al., *Bribery and Corruption 2018: Brazil*, GLOBAL LEGAL INSIGHTS, <https://www.globallegalinsights.com/practice-areas/bribery-and-corruption-laws-and-regulations/brazil> (last visited Apr. 26, 2018) (explaining that “[t]he judicial branch (divided into Federal and State jurisdictions according to the nature of the matters being discussed) is formed by first instance courts, appellate courts, and Superior courts”).

[76] See Prado & Carson, *supra* note [73] (discussing the various Brazilian institutions investigating corrupt actions). For a discussion of a particularly complex case, see Kevin E. Davis et al., *Transnational Anticorruption Law in Action: Cases from Argentina and Brazil*, 40 L. & SOC. INQUIRY 664 (2015).

[77] On process tracing, see ALEXANDER L. GEORGE & ANDREW BENNETT, CASE STUDIES AND THEORY DEVELOPMENT IN THE SOCIAL SCIENCES 205—32 (2004); David Collier, *Understanding Process Tracing*, 44 PS: POL. SCI. & POL. 823 (2011); James Mahoney, *The Logic of Process Tracing Tests in the Social Sciences*, 41 SOC. METHODS & RES. 570 (2012).

且,相邻区域的结果,虽会受到不同基层机构的影响,但实际上受到相同高层级机构的影响。因此,相邻区域间的结果比较,可以揭示基层机构变化带来的影响。例如,开办企业所花费的时间,可能受国家政府和市级政府的共同影响。但由于所有城市都以同样方式受国家政府的影响,当开办企业所花费时间发生城际变化时,即可解释为不同市级政府的表现所带来的影响。

不幸的是,通过区域间的结果比较,进而推断特定法律制度的绩效所需的一系列条件,在现实世界中并不具备。根本原因在于,法律制度的影响范围并不由领土范围界定。域外法律适用就是最明显的例子^[78],而其中最为典型的,就是关于行贿公职人员的立法。每个国家都已经或承诺将行贿定为犯罪行为。^[79]越来越多的国家,包括经济合作与发展组织(Organization for Economic Co-Operation and Development, OECD)成员国,已经将其国民在境外行贿外国公职人员的行为定为刑事犯罪。^[80]因此,当一名公职人员受到经合组织成员国国民的行贿时,该贿赂行为至少违反了两个国家的法律。在根据行贿率评估反腐败绩效时,就必须考虑到上述情况。

除了根据国籍原则,而非领土原则适用的法律,还有一些法律仅对特定宗教团体成员适用。印度继承法(India's inheritance law)就是一个例子:印度教徒、佛教徒、耆那教徒和锡克教徒受该法约束;而其他宗教的人,如穆斯林、基督徒、帕西人和犹太人,则受其他法律的约束。^[81]因此,印度继承法在国家层面

[78] Other examples, not discussed in this Article, involve legal institutions which apply exclusively to actors within a specific territory but which regulate phenomena with extraterritorial effects, such as migratory species or pollutants that move across borders. Scholars have devoted considerable attention to determining the appropriate geographic level at which to study and regulate these kinds of phenomena. See, e. g., ORAN R. YOUNG, *THE INSTITUTIONAL DIMENSIONS OF ENVIRONMENTAL CHANGE: FIT, INTERPLAY, AND SCALE* (2003); David W. Cash et al., *Scale and Cross-Scale Dynamics: Governance and Information in a Multilevel World*, 11 *ECOLOGY & SOC'Y* 81 (2006). The discussion in the text also ignores the complications that can arise when multiple legal institutions operate collaboratively across territorial boundaries. See Donald P. Moynihan et al., *Performance Regimes Amidst Governance Complexity*, 21 *J. PUB. ADMIN. RES. & THEORY*, 141, 145—47 (2007) (describing challenge of formulating performance measures for government agencies that engage in collaborative governance).

[79] G. A. Res. 58/4, ch. 3, art. 15 (Oct. 31, 2003) (adopting the Convention against Corruption which requires the 183 state parties to criminalize bribery of national public officials).

[80] This kind of legislation is required by international law. See Convention on Combating the Bribery of Foreign Public Officials in International Business Transactions, OECD, art. 4, Nov. 21, 1997.

[81] See Hindu Succession Act, No. 30 of 1956, <http://indiacode.nic.in> (outlining application of the Act in Chapter 2).

的约束力,就无法得到评估。^[82]

当人们可以选择自己受哪些法律制度治理时,另一个复杂情形就出现了。一些政府允许其国民受另一政府法律制度的约束。例如,美国和欧盟的商人有相当大的自由,选择自己的商业实体适用哪个国家管辖权下的公司法。^[83] 同样,在契约关系中,当事人可选择适用较大商业中心的法律,如纽约州法或英国法。^[84] 合同当事人也可选择外国法院来解决纠纷。^[85]

即使可以判断哪个政府负责治理某一特定活动,识别其所涉及的专门机构仍不是易事。政府有时也会对不同主体进行区别待遇。一类重要的例子就是,在有些国家政府设立的特别经济区里,诸如允许雇员组织工会、罢工的立法会被废除。^[86] 有时,国家政府还会根据与外国公司签订的投资协议,或与投资公司所在国签订的投资协议,为外国投资者设立特殊的法律制度。^[87] 较高层级

[82] Andrew Ellul et al., *Inheritance Law and Investment in Family Firms*, 100 AM. ECON. REV. 2414, 2426 n. 16 (2010) (analyzing the effect of inheritance law on corporate governance, but reporting that in some countries, inheritance law is too complex to capture in a simple indicator).

[83] See Zoe Adams & Simon Deakin, *Freedom of Establishment and Regulatory Competition*, in THE OXFORD HANDBOOK OF EUROPEAN UNION LAW 537, 538 (Anthony Arnall & Damian Chalmers eds., 2015). For recent empirical studies, see Ofer Eldar & Lorenzo Magnolfi, *Regulatory Competition and the Market for Corporate Law* (Yale Law & Econ., Research Paper No. 528), <https://ssrn.com/abstract=2685969> (describing the incorporation choices within the United States); Carsten Gerner-Beuerle et al., *Why Do Businesses Incorporate in Other EU Member States? An Empirical Analysis of the Role of Conflict of Laws Rules* (ECGI, Working Paper No. 361/2017), <https://ssrn.com/abstract=3012139> (discussing the incorporation choices within the European Union).

[84] See Gilles Cuniberti, *The International Market for Contracts: The Most Attractive Contract Laws*, 34 NW. J. INT'L L. & BUS. 455 (2014) (describing and explaining attractiveness of English and Swiss law among parties to international agreements providing for arbitration); Theodore Eisenberg & Geoffrey P. Miller, *The Flight to New York: An Empirical Study of Choice of Law and Choice of Forum Clauses in Publicly-Held Companies' Contracts*, 30 CARDOZO L. REV 1475 (2008) (finding that major commercial contracts frequently designate New York law and explaining why).

[85] See generally REGULATORY COMPETITION IN CONTRACT LAW AND DISPUTE RESOLUTION 427—31 (Horst Eidenmüller ed., 2013).

[86] See GOKHAN AKINCI & JAMES CRITTLE, SPECIAL ECONOMIC ZONES: PERFORMANCE, LESSONS LEARNED, AND IMPLICATIONS FOR ZONE DEVELOPMENT, FOREIGN INVESTMENT ADVISORY SERVICE 40 (2008), <http://documents.worldbank.org/curated/en/343901468330977533/Special-economic-zone-performance-lessons-learned-and-implication-for-zone-development> (describing labor standards issues documented by the International Labour Organization and the International Confederation of Trade Unions).

[87] See, e. g., Keith Molkner, *A Comparison of the Legal Regimes for Foreign Investment in Russia, Kazakhstan, and Kyrgyzstan*, 11 INT'L TAX & BUS. L. 71 (1993) (providing an overview and comparison of legal regimes for foreign investment).

政府也可采取影响范围较小的区别待遇。例如,中国政府有时会在试点城市进行制度改革(legal reform),以此作为先期试验。^[88]最近的一篇文章指出,一系列以市场为导向的城市住房改革,最初都是中央政府在地方发起的试验。^[89]

甚至表面上中立的法律规范,也会因执法的不统一,而形成不同效果。例如,肯尼亚2015年在其《公司法》(Companies Law)中,对企业开办程序作出全国统一的规定。^[90]这个程序的诸多方面由县级政府负责。^[91]根据《肯尼亚营商环境报告》,开办企业的时间、成本出现县际变化。而这是各县级机构在营业许可证申领上的时间要求、收费标准的差异所造成的。^[92]但有些差异则是由国家机构造成的。^[93]开办企业者需要向全国企业注册中心提交申请,而这个注册中心只在内罗毕设有办事机构。^[94]当面提交申请需要一到两周的时间。^[95]因此,到内罗毕旅费的差异,就造成各县企业注册费用的不同。^[96]

基于以上原因,行为的物理位置不是评估法律制度影响的必要指向标。这就意味着,通过比较不同地理位置的结果,来区分各法律制度影响的方法并不牢靠。跨区域结果的变化,可能是由名义上负责相关区域的机构不同所引起的,但也可能是由治外法权、自愿选择管辖,或是更高层级政府的区别对待所产生的。比如,在不同城市开办企业,所花费的时间有差异,这既可以解释为市级政府间的不同所引起的,也可以解释为国家立法差异所引起的。企业披露行贿的比例在省际有差异,这可以解释为省级政府在实施反贿赂法上的不同,也可以解释为受域外反贿赂法规制的企业比例不同所引起的。

[88] See Ciqi Mei & Zhilin Liu, *Experiment-Based Policy Making or Conscious Policy Design? The Case of Urban Housing Reform in China*, 47 POL'Y SCI. 321, 329 (2014).

[89] See *id.* at 327–29 (2014). Other commentators suggest that these experiments are driven by subnational governments. See, e.g., Chenggang Xu, *The Fundamental Institutions of China's Reforms and Development*, 49 J. ECON. LIT. 1076, 1079 (2011). Xu also reports that each level of subnational government negotiates performance targets with subordinate levels of government, implying that different subordinate governments might owe different obligations to the higher-level government. See *id.* at 1093.

[90] See WORLD BANK GRP., *DOING BUSINESS IN KENYA 2016 21* (2016), [http://www.doingbusiness.org/~media/WBG/DoingBusiness/Documents/Subnational-Reports/DB16-Sub-Kenya.pdf](http://www.doingbusiness.org/~/media/WBG/DoingBusiness/Documents/Subnational-Reports/DB16-Sub-Kenya.pdf).

[91] See *id.* at 5.

[92] See *id.* at 7.

[93] See *id.* at 30.

[94] See *id.* at 21.

[95] See *id.* at 24.

[96] See *id.* at 24–25.

五、法制绩效评估市场

鉴于第三部分所阐述的原因,在多层级治理系统中进行法制绩效评估将会极具挑战性。在一些情况下,无论投入多少资源,挑战都无法克服。在另一些情况下,问题在于有谁可以为形成有效的绩效评估投入资源。下面的前两部分内容,侧重于审视市场力量,也就是“私”主体的利己选择,是否能促进良好的绩效评估的形成。最后一部分内容则讨论,我们是否可以期待非营利组织或公共部门去弥补绩效评估市场中的缺陷。

(一) 绩效评估的用户需求

为了解自利主体是否会进行绩效评估,我们首先要计算绩效的需求水平。评估者的收益来自其需求的满足程度。所以,需求水平将决定评估者的动机强度。

正如本文第一部分所指出的,对法制绩效评估的需求可能来自不同的潜在用户群体。首要的潜在用户是那些关注法律制度如何影响他们生活,以及重要决策(如居住地,是否创业、在哪里创业、如何创业等)的人群。这些人群尤其关注:开办企业所需时长,获得经营许可是否需要行贿,或是晚间周边环境的安全程度,诸如此类的信息。换句话说,他们关心由政府行为所产生的效果,包括不同机构组合所产生的结果。

出于利己考虑的人们,可能不需要将他们关心的结果与特定机构对结果的贡献相联系。为什么准企业家或准定居者一定要关心国家政府还是市政府对企业准入、腐败或预防犯罪负责?原则上,仅有关于结果的信息就完全足够了。但是,关于特定机构绩效的信息,会在预测未来结果时有实际价值。假设,某个最近营商较为便利的省份,其政府可能会做出一些改变。那么省政府的改变,是否会影响营商环境?为了回答这个问题,需要知道以往结果是由省政府还是国家政府造成的。

对法制绩效评估的需求,也会来自在法律官员之上行使权利的人群,即选民、更高级别的官员。^[97] 这些用户会特别关注特定机构的绩效。^[98] 公职人员的同僚也会对绩效监测感兴趣。例如,警察会关注邻近辖区同事的表现,以评估他们自己的表现,或确定新招募人员,或借鉴学习新的工作方法。他们也会希望获得特定机构的精准信息。

[97] See, e. g., Davis et al., *supra* note [21], at 92—95 (discussing how the Doing Business indicators are used in decision making).

[98] See *id.*

对法制绩效评估的总需求,取决于相关信息对每个潜在用户的价值及用户总量。^[99] 绩效评估对每个用户的价值,取决于评估相对于用户个人知识,能提供多少有效信息。这依赖几个方面的因素,包括感知的精确度、评估频率以及评估者身份。^[100] 精确度的价值是不言而喻的。可得数据的数量,在指标用于比较目的的情况下,也会成为相关因素。评估的数量越大,可比较的范围就越广。最后,绩效评估用户可能会用评估者的“出身”或名声,来推测评估结果的质量。^[101]

聚焦于特定机构绩效的评估,其价值取决于替代者的可得性。^[102] 必要时,用户会根据他们已经了解,或他们自认为了解的相似的机构绩效,来推断评估对象的绩效。如果相似制度/机构的绩效相差较多,这种策略就会导致较大误差。^[103] 例如,一个考虑在墨西哥投资的商人,会倾向于根据世界银行《营商环境报告》中的数据,来预测在墨西哥境内经商的便利程度。毕竟,理论上讲,相对于各国之间的差异,墨西哥境内的营商环境差异会小很多。但如果投资计划是建造仓库,那么世界银行的数据会误导这个商人。当他试图于2016年在科利马,一个位于太平洋海岸的墨西哥小州,获得建筑许可证。那么科利马的办事效率会远远落后于全球最好的城市,也会落后于墨西哥在此项上的得分。因为墨西哥在该项上全球第87名的排名,主要得益于墨西哥城和蒙特雷的分数。^[104]

最后,对特定绩效评估的总需求,会随着用户数量的增加而增加。用户数量又反过来取决于受相关机构影响的人群、受监管的人群,或是自认是公职人员职业共同体一员的人群。受利己因素驱动的用户,对影响高风险决策的法制绩效评估需求最大。由此可以推定,至少对出于利己目的的用户而言,影响给定结果的机构越多,对单个机构绩效评估信息的需求就越少。^[105]

[99] See, e. g., KEVIN E. DAVIS & BENEDICT KINGSBURY, INDICATORS AS INTERVENTIONS: PITFALLS AND PROSPECTS IN SUPPORTING DEVELOPMENT INITIATIVES 17–18 (2011) (discussing determinants of the demand for indicators).

[100] See *id.* at vii.

[101] See *id.* at 27–29.

[102] See *id.* at 28.

[103] See Richard Snyder, *Scaling Down: The Subnational Comparative Method*, 36 STUD. COMP. INT'L DEV. 93, 98 (2001) (discussing risks of “meanspirited analysis”).

[104] See WORLD BANK GRP., DOING BUSINESS IN MEXICO 5 (2016); see also *Ease of Doing Business in Mexico*, DOING BUSINESS, <http://www.doingbusiness.org/data/exploreeconomies/mexico> (last visited Apr. 26, 2018).

[105] See, e. g., CHRISTOPHER R. BERRY, IMPERFECT UNION: REPRESENTATION AND TAXATION IN MULTILEVEL GOVERNMENTS 57–61 (2009) (explaining that most voters in multilevel governments have little incentive to pay attention to institutions with narrow policy functions).

对绩效评估的需求,并不必然转化为对其支付的意愿。一旦公布,绩效评估数据就可以极低成本进行复制和传播。当他们手握便宜的复制版本时,潜在用户就不愿为绩效评估的信息支付费用。这就意味着,对营利性评估者而言,绩效评估的有效需求相对较低。

(二) 绩效评估的供给

法制绩效评估市场分析的下一步,就需考虑等式另一边的供给方。评估形成的过程具体来说有以下几步:(1) 评估的设计与概念化操作;(2) 收集相关数据;(3) 分析数据并计算分数;(4) 公布分数。^[106] 每一步的成本都会很高。成本包括:绩效评估设计和数据分析的专业人士所花费的时间;从已有数据库购买数据所支付的费用;从熟悉相关制度形式与效果的人士那里直接收集数据所支付的费用;推广和出版绩效评估产品的花费。^[107]

在很多情况下,绩效评估成本是固定的这一假设是合理的。这意味着评估成本不会随评估价值的变化而变化。首先,一般来说,绩效评估的设计和推广成本,似乎与所评估机构的影响力或重要程度无关。而且,就法制数据本身而言,从法律条文和法律实践中收集数据所花费的成本,也与所评估机构的影响力或重要性无关。更大的管辖区域,并不必然对应更长的法律文本:印度尼西亚的人口是马来西亚的8倍之多,但是马来西亚的宪法文本长度却是印度尼西亚的10倍。^[108] 阅读和编辑印度尼西亚宪法的成本,并不会比用在马来西亚宪法上的多。

收集或分析未经编纂的法律实践或法治效果数据的成本,通常取决于需评估的行为或结果的数量。一般而言,需收集和分析数据的行为和结果越多,成本就越高。^[109] 因此,这一结论是可信的:相对有影响力的机构,其数据收集的

[106] See Kevin E. Davis et al., *Introduction: The Local-Global Life of Indicators: Law, Power, and Resistance*, in THE QUIET POWER OF INDICATORS: MEASURING GOVERNANCE, CORRUPTION, AND RULE OF LAW 1, 10–17 (Sally Engle Merry et al. eds., 2015).

[107] See *id.*

[108] *Constitution Rankings*, COMPARATIVE CONSTITUTIONS PROJECT, <http://comparativeconstitutionsproject.org/ccp-rankings> (last visited Apr. 26, 2018) (reporting the length of Indonesia's and Malaysia's constitutions as 5915 words and 64,080 words, respectively); *Open Data: Population*, WORLD BANK, https://data.worldbank.org/indicator/SP.POP.TOTL?locations=ID-MY&-name_desc=false (last visited Apr. 26, 2018) (reporting total populations of Indonesia and Malaysia in 2016 as 261,115,465 and 31,187,265, respectively).

[109] This is true to the extent that data are collected through interviews or surveys, since these methods of data collection require effort on the part of each respondent, and sometimes an interviewer as well. See FLOYD J. FOWLER, *SURVEY RESEARCH METHODS* 69–85 (4th ed. 2009) (describing costs per respondent for different modes of survey research).

成本也会较高。但数据收集成本与结果量之间,并不必然成正比。比如,获得行贿率调研数据的成本,会随着人口规模的增加而提高。因为人口规模越大,在一定的置信区间,评估行贿率所需的样本量就越大。但是,样本规模并不必然随着人口基数的增加而增加。例如,预测 500 万人的平均结果,需要 700 人的样本量;而对 1000 万人口进行评估,则需要 1000 人的样本量。但是,如果人口规模更大,其异质性更高,且需要对人口内部变化进行评估,而非简单测量典型结果,情况就会不同。^[110]

如果在计算机的帮助下进行数据分析,额外数据分析所增加的成本就会很小。^[111] 但同样的,如果相关结果的成分复杂,越大的数据组就需要越多的分析工作。决定成本高低的关键因素,更多的是被评估对象内部的异质程度。

鉴于上述原因,试图以最低成本,展开最有价值测评的评估者们会发现,评估相对有影响力的法律制度会更诱人。这就意味着,他们需要对以显著且相似方式,影响最广泛范围人口、活动的法律制度进行评估。类似评估的综合价值相对较高,而其产品成本,并不明显高于其他法律制度评估的成本。这表明,市场力量不会对多层级治理体系中的最基层制度进行绩效评估。即使对基层制度进行评估,也只覆盖影响最大、最突出的那部分。因为潜在用户会希望从少量机构绩效出发,推测相似机构的表现。

(三) 非营利组织与“公”主体

目前为止,我们已经考虑了利益最大化的利己主体是否会进行法制绩效评估。他们并非是仅有的评估供给者。事实上,非营利组织与公共部门在现有法制绩效评估中,也发挥了重要作用。^[112] 世界银行[它发布了《营商环境报告》(Doing Business Reports)]是国际组织,热图利奥·瓦加斯基金会[它发布了巴西司法信心指数(ICJBrasil)]是私人基金会^[113],而越南工商总会与美国国际开发署[发布了《越南省际竞争力指数》(Provincial Competitiveness Index)]分别是非营利贸易协会与政府机关^[114]。这些主体可能因关注社会福利,或希望

[110] See Snyder, *supra* note [103], at 94.

[111] See Davis et al., *supra* note [21], at 85.

[112] For an analysis of entities playing a role in producing performance measures, see Tim Büthe, *Beyond Supply and Demand: A Political-Economic Conceptual Model, in GOVERNANCE BY INDICATORS: GLOBAL POWER THROUGH QUANTIFICATION AND RANKINGS* 29, 37—40 (Kevin E. Davis et al. eds., 2012).

[113] See *Mission and History of FGV*, FGV, <http://portal.fgv.br/en/institutional#> (last visited Apr. 26, 2018).

[114] See *PROVINCIAL COMPETITIVENESS INDEX*, <http://eng.pcivietnam.org> (last visited Apr. 26, 2018).

自我提升,而进行绩效评估。但他们在这之中并不关心利益最大化。没有这些主体的介入,是否进行地方性机构绩效评估就变得不明朗。

六、结论

绩效评估是保持和改善法律制度表现的重要工具。通过多种渠道,包括投票、激励性薪酬、绩效竞争上岗,绩效评估既可以引导公职人员提升工作表现,也可将表现更好的人员替换上岗。在大多数情况下,将大量数据压缩成一个指标的比较式评估,可能更具有价值。

构成现有法制绩效评估的法律指标,都是将其测量现象的一方面进行彻底简化而形成的:他们忽略了这样一个事实,即法律制度是在一个多层治理体系中运作,而这个体系中的每个层级都包含多个治理机构。很多指标都试图评估国家层面的政府绩效,但极少去关注地方绩效对国家机构表现的影响,也很少去试图评估地方机构自身的绩效。

目前法律指标中的这个缺漏,是一个很大的缺点。为了实现其潜在价值,法制绩效评估必须在多机构治理体系中,区分出特定机构的绩效。建构这样一套绩效评估体系并非易事,特别是当该评估依赖于测量相关机构的社会或经济效果时。其主要挑战来自将特定机构与特定效果相匹配。与流行观点相反,地理位置与法律适用之间没有直接的对应关系。法律适用范围全部或部分被非领土要素决定,如国籍、宗教、合意,或是表面上的统一适用而内里的区别对待。

即使能充分建构起法制绩效评估体系,在市场运作下,也无法保证它们能得到实现。在市场的一边,绩效评估的需求很大。选民、投资者、债权人、监管者、公职人员的同僚——所有这些群体都有理由重视机构表现的信息。围绕影响范围广、活动价值高的机构所展开的绩效评估,其需求更高。对市场的供给方来说,评估高影响力机构的成本,与评估低影响力机构的成本相近。这就意味着,营利性评估者会聚焦于测量更有影响力的机构。而对影响力较低的机构评估的需求,将由非营利组织或公共部门满足。这可能意味着,相关需求无法被满足。

分析的结果是,潜在评估者不能或不愿为特定法律制度提供绩效评估。这个结论对围绕多层治理绩效的讨论有诸多重要启示,而该讨论则反过来与制度设计中的重要问题有关。在法律语境下,经典讨论围绕超国家组织、联邦、州和市政府之间的分权问题展开。普遍认为,各法律机构的表现,一定程度上取决于是否有让它们承担责任的机制。因此,这些机构的表现将取决于绩效评估的有效性。

最后,上述讨论也涉及了一系列围绕法律指标建构与运用的文献。^[115] 此类研究发展迅速。研究表明,在建构抓取机构表现的法律指标的过程中,我们应当给予多层治理体系所提出的挑战以更多关注。

(审校编辑 邵博文)

[115] See generally Davis, *supra* note [10] (surveying the literature on the production and use of legal indicators).