

中国刑事法治指数的创制研究

曾 贇

摘要:为创制一种全面、客观、科学的刑事法治指数,首先,需要从刑事法治专家评估和普通公众评估两个层面建立全面的研究构想和假设。根据研究构想和假设,创设刑事法治专家评估和普通公众评估两种类型调查问卷;采用整群随机抽样调查法,收集刑事法治专家调查有效样本1312个,普通公众调查有效样本1289个。其次,在对数据进行无量纲化基础上,采用验证性因素分析法,创建刑事法治专家评估指数、普通公众评估指数结构方程模型和二阶单因素模型;通过检验,刑事法治专家评估10个研究构想和理论假设中,有5个被验证有效;普通公众评估6个研究构想和理论假设中,有4个被验证有效。再次,根据被有效验证的研究构想和理论假设,创建刑事法治专家评估指数、普通公众评估指数二阶单因素计算模型;通过对模型的计算客观得出指标权重和二级指数权重。最后,采用加权算术平均法创制中国刑事法治指数。

关键词:刑事法治;法治指数;因素分析

DOI: 10.19836/j.cnki.37-1100/c.2021.01.006

法治指数犹如一件测量器具,可据以客观测量一个国家、地区或城市的法治程度。当法治成为法治社会中人们生活的一部分时,法治测量可以以人们对于法治的认同度为对象。人们关于法治生活的认同来自人们的直接经验,虽然这种经验具有主观性,但我们可以通过客观的方法认识它。对刑事法治程度的测量,国内外研究中目前没有创制专门的刑事法治指数,而仅将其视作法治测量或法治评估内容的一个部分。例如,世界正义工程在法治指数体系中构设刑事法治二级指数;又如,中国法治满意度评估在法治指标体系中构设刑事法治测量指标。显然,前者由于设有刑事法治分级指数,故指数构设较全面,同时亦可为人们提供一扇管窥一个国家或地区刑事法治状况的窗口;后者由于将刑事法治测量指标分散设立于不同测量内容之中,故刑事法治测量指标构想有些片面,且亦难具有刑事法治程度测量功能。有鉴于此,创制一种准确、全面、客观的刑事法治指数,对于科学测量我国刑事法治程度,显得尤为重要。

一、刑事法治指数研究构想、理论假设与调查

指数有广义和狭义两种含义。一般来说,狭义上的“指数”(index)与“指标”(indicator)是两个不同的概念^①。根据这一区分,法治指数和法治指标亦为两个不同的概念。基于此,刑事法治指数可界定为一种反映复杂的、异质的刑事法律现象综合变化和及相关影响因素总体变化的相对数;刑事法治指标是指一种反映简单的、同质的刑事法律现象总体数量、特征的概念和数值。刑事法治指数是一套定型的、客观的、科学的旨在测量和评估一个国家、地区刑事法治程度的标准化数值,它不能通过对指标得分的直接加总得出,而只能通过对指数模型的创制与计算得出。兹从我国刑事法治实践出发,通过对现有文献的梳理,提出中国刑事法治指数的研究构想和理论假设;在此基础上,设计调查问卷,展开实证调查。

作者简介:曾贇,广州大学法学院教授,博士生导师(广州 510006; zengyun70@163.com)。

^① 刘汉良:《统计学教程》,上海:上海财经大学出版社,1995年,第269页。

(一) 研究构想

法治指数研究构想应贯彻全面性原则。具体来说,一是测量内容的全面性,二是测量维度的全面性,三是测量维度全面贯穿于每一测量内容之中。基于此,可从形式法治和实质法治两个角度构设中国刑事法治指数测量内容。形式法治测量涵盖警察适用法律平等、司法公开、司法廉洁、司法效率、程序公正、司法机关独立行使职权等内容;实质法治测量涵盖禁止刑讯逼供、无罪推定、司法公正和权利保障等内容。同时,从刑事法治专家评估和普通公众评估两个层面构想中国刑事法治指数测量指标。刑事法治专家评估指数构想涵盖 10 项二级指数和 43 个指标(x1-x43),普通公众评估指数构想涵盖 6 项二级指数和 25 个指标(y1-y25)。

基于被调查对象缺乏立法明确性方面的直接经验,故研究构想未设计立法明确性方面的测量指标。当然,立法明确性问题也反映在司法适用中,故通过对正确适用法律方面的调查,亦可实现对立法明确性的测量。另外,刑事法治普通公众评估指数测量未涵摄刑讯逼供、无罪推定、监狱罪犯教育改造效率和刑事司法效果等 4 项二级指数。理由是:一方面,本次普通公众调查对象为监狱中正在服刑的罪犯,故为保持调查的客观、中立性,针对监狱服刑人员的调查不宜涉及前述四个方面;另一方面,即便所调查对象为非特定普通公众,即社会中的普通公众,他们对前述四个方面内容既缺乏感觉经验,也缺乏相关专业知识,故对社会普通公众的调查亦不宜涉及前述四个方面。鉴于司法救济在刑事法治实践中的重要地位,从我国学者苗连营关于司法救济权相关论述出发^①,创设嫌疑人获得及时公正审理的权利(x10)、被告人获得及时公正审理的权利(x11)、审判期间法律援助(x18)、侦查期间法律援助(x19)四个测量指标。

(二) 理论假设

司法公正既是法治国家的首要要求^②,也是法治的最重要内容^③,因此刑事司法公正对刑事法治的测量具有特殊重要意义。据此,提出“假设 4,刑事司法公正对刑事法治具有最强解释力”。如果这一假设成立,那么我们可以通过对刑事司法公正诸影响因素的测量来创制刑事法治指数。基于刑事法治指数研究构想,提出如下假设:假设 1,禁止刑讯逼供对刑事司法公正具有显著性正向影响;假设 2,无罪推定对刑事司法公正具有显著性正向影响;假设 3,刑事司法廉洁对刑事司法公正具有显著性正向影响;假设 5,刑事司法效率对刑事司法公正具有显著性正向影响;假设 6,警察平等执法对刑事司法公正具有显著性正向影响;假设 7,监狱罪犯教育改造效率对刑事司法公正具有显著性正向影响;假设 8,刑事司法效果对刑事司法公正具有显著性正向影响;假设 9,刑事司法公开对刑事司法公正具有显著性正向影响;假设 10,刑事司法机关独立行使职权对刑事司法公正具有显著性正向影响。

(三) 调查问卷设计

根据刑事法治指数研究构想和假设,设计两种类型调查问卷——刑事法治专家调查问卷(Qualified Respondent Questionnaires,以下简称 QRQ)和普通公众调查问卷(General Population Poll,以下简称 GPP)。例如 QRQ1:为获取口供,侦查人员会对嫌疑人身体造成轻微伤害吗?又如 GPP1:你认为我国公安民警中有没有腐败行为?然后采用极值法对调查问题各题项进行无量纲化^④,从而得到指标标准值。例如 QRQ1 调查题项设计为:(1)很可能,(2)可能,(3)不可能,(4)很不可能;又如 GPP1 调查题项设计为:(1)没有一人,(2)有些人,(3)大部分人,(4)所有人。然后采用极值法对指标各题项进行无量纲化。前者各题项计量标准为:很可能(0),可能(0.333),不可能(0.667),很不可能(1);后者各题项计量标准为:没有一人(1),有些人(0.667),大部分人(0.333),所有人(0)。据

① 苗连营:《公民司法救济权的入宪问题之研究》,《中国法学》2004 年第 5 期。

② 甘雯:《关于司法公正的几个基本问题》,《中国法学》1999 年第 5 期。

③ 王利明:《法治的社会需要司法公正》,《人民司法》1998 年第 2 期。

④ 马立平:《统计数据标准化——无量纲化方法》,《北京统计》2000 年第 3 期。

此,指标标准值取值均在 0-1 之间;相应地,指数取值亦均在 0-1 之间。调查问卷涵盖二级指数(f1-f10)和指标(x1-x43、y1-y25),测量内容包括:

1. 禁止刑讯逼供(f1)。根据我国学者的研究,刑讯逼供的行为对象主要为嫌疑人^①,故刑事法治专家调查问卷仅涉及对犯罪嫌疑人的调查。测量指标包括:x1 轻微逼供;x2 严重逼供;x3 重罪逼供。

2. 无罪推定(f2)。无罪推定是嫌疑人、被告人依法享有的基本权利。专家评估测量指标包括:x4 审判阶段假定无罪率;x5 侦查阶段假定无罪率;x6 被告挑战不利证据率。

3. 刑事司法廉洁(f3)。刑事司法廉洁即刑事司法工作人员职务行为的廉洁性。专家评估与公众评估指标包括:警察廉洁(x7、y1);检察官廉洁(x8、y2);法官廉洁(x9、y3)。

4. 刑事司法公正(f4)。刑事司法公正测量既包括对正确适用法律等诸实体公正的测量,又包括对获得法院及时审理等诸正当程序的测量。刑事司法公正专家评估指标包括:x10 嫌疑人获得及时公正审理;x11 被告人获得及时公正审理;x12 法院独立行使审判权;x13 法官没有偏见;x14 囚犯权利保障。公众评估指标包括:y4 法院公正审理;y5 公安机关公正执法;y6 检察机关公正适法;y7 被告人权利保障;y8 嫌疑人权利保障。

5. 刑事司法效率(f5)。刑事司法效率专家评估指标包括:x15 刑事侦查部门间信息资源共享;x16 公、检、法、司间信息资源共享;x17 办案检察官数量;x18 审判期间法律援助;x19 侦查期间法律援助。公众评估指标包括:y9 入室盗窃犯罪案件侦破的及时性;y10 入室盗窃犯罪案件起诉、审理的及时性;y11 抢劫犯罪案件侦破的及时性;y12 抢劫犯罪案件起诉、审理的及时性。

6. 警察平等执法(f6)。警察平等执法所测量的是机会上的平等,即适用法律平等。警察平等执法专家评估与公众评估指标包括:平等对待贫穷的嫌疑人(x20、y13);平等对待妇女嫌疑人(x21、y14);平等对待少数民族嫌疑人(x22、y15);平等对待有宗教信仰的嫌疑人(x23、y16);平等对待具有外国国籍的嫌疑人(x24、y17);平等对待同性恋嫌疑人(x25、y18)。

7. 监狱罪犯教育改造效率(f7)。监狱罪犯教育改造效率是指在监狱刑罚执行资源总量不变的情况下,将监狱罪犯教育改造资源配置到最优状态。专家评估指标包括:x26 监狱罪犯之间的暴力冲突;x27 监狱在罪犯再社会化与再犯预防方面的投入;x28 监狱单独羁押“顽危犯”的设施;x29 监狱监管安全。

8. 刑事司法效果(f8)。刑事司法效果专家评估指标包括:x30 刑事案件最终判决结果受工作压力影响比率;x31 刑事案件最终判决结果受某些腐败行为影响的比率;x32 错控率;x33 错判率;x34 错捕率;x35 青少年犯罪率。

9. 刑事司法公开(f9)。刑事司法公开专家评估指标包括:x36 法律权利普及;x37 采用各民族通用语言文字公开出版发行基本法律;x38 刑法普及。公众评估指标包括:y19 公安机关案件信息公开;y20 检察机关案件信息公开;y21 法院案件信息公开;y22 司法机关案件信息公开。

10. 刑事司法机关独立行使职权(f10)。我国司法机关独立行使职权原则一般仅适用于审判机关、检察机关职权范围。鉴于世界正义工程刑事法治指数构想将独立行使侦查权纳入其中,为了便于比较,结合我国刑事诉讼法的有关规定,本调查把侦查人员在侦查过程中不受不当干预的情形,也纳入调查范围。司法机关独立行使职权专家评估指标包括:x39 警察办案不受犯罪组织影响;x40 检察官办案不受犯罪组织影响;x41 法官办案不受犯罪组织影响;x42 最高人民法院审理不受不当干预比率;x43 地方各级法院审理不受不当干预比率。公众评估指标包括:y23 法官独立办案;y24 侦查人员独立办案;y25 检察官独立办案。

(四)实证调查

2019年3-9月,笔者采用整群随机抽样方式,从中国东北地区的黑龙江,东部地区的浙江、广东,中部地区的湖南,西部地区的四川5个省抽取哈尔滨、杭州、宁波、温州、广州、佛山、长沙、永州、成都、绵阳

^① 孙长永、王彪:《审判阶段非法证据排除问题实证考察》,《现代法学》2014年第1期。

10个城市进行问卷调查。调查采用随机拦截、集中调查和电话访问三种方式进行。通过调查,共收集2601个有效调查样本,其中刑事法治专家调查样本1312个,普通公众调查样本1289个。根据描述统计,专家调查样本中男、女性各占样本总体的53.4%、46.6%;从业时间在5年以下者占样本总体的88.3%,5-10年者占样本总体的7.1%,10年以上者占样本总体的4.6%;律师占调查样本总体的89.5%,其余法律职业占样本总体的10.5%。公众调查对象为正在监狱中服刑的罪犯,男、女性别样本数按照监狱押犯性别比抽取,分别约占样本总体的75.7%、24.3%;失业者样本数约占样本总体的29.7%。

总的来说,中国刑事法治指数调查是一次比较全面、客观的调查。首先,本研究调查样本容量较大,抽样误差较小,样本代表性较好。与世界正义工程法治指数调查相比较,2019年,其仅选取中国东部地区的3个城市进行问卷调查,从中仅收集刑事法治专家调查样本30多个,普通公众调查样本508个^①。其次,本研究关于刑事法治指数的构想较为全面,且被调查对象具有关于刑事法治全过程的直接经验。再次,本调查获得的数据比较客观、真实。如前所述,刑事法治专家调查样本的89.5%来自律师,公众调查样本来自监狱中正在服刑的罪犯。相较于其他司法工作人员,律师介入了刑事案件的全过程,故以律师为调查对象所获得的调查数据更加客观、全面。监狱罪犯由于曾处于被追诉地位,且现时正遭遇惩罚,故其对我国刑事法治的感觉经验更加客观、真实。当然,其对刑事法治的认同程度也可被视为普通公众对我国刑事法治信任和信赖程度的最低反映。

二、刑事法治专家评估指数模型的创建

一般地,国内外法治指数或法治评估研究均直接根据研究构想来计算法治指数或指标得分^②。然而,由于人的思维的主观性,故并非所构想的所有指标均可被用来创制法治指数。就此而言,创制刑事法治指数的关键,是如何科学地发现所构想的指标究竟有哪些可据以测量刑事法治。那么,刑事法治专家评估指数研究所构想的43个指标中,究竟有哪些可被用来创制刑事法治专家评估指数?为解决这一问题,兹根据实证调查数据,采用验证性因子分析法,运用mplus8.0统计软件,创建刑事法治专家评估指数结构方程模型和二阶单因素模型。

(一)刑事法治专家评估指数结构方程模型

基于刑事法治专家调查数据,依据研究构想和假设1-3、假设5-10,可初步创建刑事法治专家评估指数结构方程模型(后文简称为“初始模型”)。鉴于初始模型拟合不够理想,故需要根据初始模型参数估计对其进行修正。经修正,模型组成包括:一是 x_1-x_{41} 共41个观测变量(评估指标)。由于潜变量“刑事司法机关独立行使职权(f_{10})” \rightarrow 观测变量“法官办案不受犯罪组织影响(x_{40})”的标准化载荷系数绝对值高于1.00,说明模型存在多重共线性。在删除 x_{42} 、 x_{43} 这2个观测变量后,模型多重共线性问题消失。基此,模型的观测变量为41个(x_1-x_{41})。二是 f_1-f_{10} 共10个潜变量(因素)。三是共7条路径走向。根据初始模型标准化参数估计结果,“禁止刑讯逼供($P=0.195$)”“警察平等执法($P=0.243$)”“刑事司法公开($P=0.169$)”“刑事司法机关独立行使职权($P=0.762$)”4个潜变量对“刑事司法公正”潜变量不具有直接的显著性正向影响,假设1、6、9、10未得到有效验证。基此,需检验其是否通过中间变量对刑事司法公正产生影响。经修改,模型路径走向为:路径1: $f_1 \rightarrow f_8 \rightarrow f_4$;路径2: $f_6 \rightarrow f_8 \rightarrow f_4$;路径3: $f_{10} \rightarrow f_8 \rightarrow f_4$;路径4: $f_9 \rightarrow f_2 \rightarrow f_4$;路径5: $f_3 \rightarrow f_4$;路径6: $f_5 \rightarrow f_4$;路径7: $f_7 \rightarrow f_4$ 。

^① 2019年,世界正义工程对中国法治指数调查的地点为上海、北京、广州。2019年,其在全球126个国家和司法区域共收集法律专家调查样本3800个。参见The World Justice Project, *The WJP Rule of Law Index 2019*, p. 161, 164, <https://worldjusticeproject.org/our-work/wjp-rule-law-index>, 访问日期:2020年3月31日。

^② 曾赞:《法治评估的有效性和准确性——以中国八项法治评估为检验分析对象》,《法律科学》2020年第2期。

在根据 M. I. 指标值对模型进行四次修正后,卡方值(3005.112)与自由度(745)之比为 4.03,符合小于 5 的标准要求。卡方值显著性水平 $P=0.0000$,符合小于 0.0001 的标准要求。 $RMSEA$ 为 0.048,符合小于 0.08 的标准要求; $SRMR$ 为 0.058,未满足小于 0.05 的标准要求。 CFI 、 TLI 值分别为 0.929、0.922,符合大于 0.9 的标准要求。经修正后,刑事法治专家评估指数结构方程模型拟合理想。根据模型参数估计标准化结果与路径分析,可得出以下结论:

一是刑事司法廉洁对刑事司法公正具有很大影响力。在模型的 7 条路径中, $f_3 \rightarrow f_4$ 的路径系数最高,为 0.648 ($P=0.00$),说明在其他条件不变时,“刑事司法廉洁(f_3)”潜变量每提升 1 个单位,“刑事司法公正”潜变量就提升 0.648 个单位。据此,假设 3 得以验证。

二是刑事司法效率、刑事司法效果对刑事司法公正具有较大影响力。 $f_5 \rightarrow f_4$ 、 $f_8 \rightarrow f_4$ 的路径系数分别为 0.25 ($P=0.00$)、0.127 ($P=0.000$),说明在其他条件不变时,“刑事司法效率(f_5)”“刑事司法效果(f_8)”潜变量每提升 1 个单位,“刑事司法公正”潜变量分别提升 0.25、0.127 个单位。据此,假设 5、8 得以验证。

三是监狱罪犯教育改造效率对刑事司法公正具有较明显影响力。 $f_7 \rightarrow f_4$ 的路径系数为 0.093 ($P=0.000$),说明在其他条件不变时,“监狱罪犯教育改造效率(f_7)”潜变量每提升 1 个单位,“刑事司法公正”潜变量提升 0.093 个单位。据此,假设 7 得以验证。

四是无罪推定对刑事司法公正具有一定程度的影响力。 $f_2 \rightarrow f_4$ 的路径系数为 0.06 ($P=0.000$),说明在其他条件不变时,“无罪推定(f_2)”潜变量每提升 1 个单位,则“刑事司法公正”潜变量提升 0.06 个单位。据此,假设 2 得以验证。不过,与上述 4 个因素相比,其影响力要小一些。

五是禁止刑讯逼供、警察平等执法、刑事司法机关独立行使职权、刑事司法公开通过中间变量对刑事司法公正具有间接影响力。根据修改后模型参数估计, $f_1 \rightarrow f_8$ 、 $f_6 \rightarrow f_8$ 、 $f_{10} \rightarrow f_8$ 等 3 条路径系数分别为 0.257 ($P=0.00$)、0.146 ($P=0.00$)、0.301 ($P=0.00$),因此“禁止刑讯逼供”“警察平等执法”“刑事司法机关独立行使职权”等 3 个潜变量通过“刑事司法效果”这一中间效应对“刑事司法公正”潜变量产生间接影响。 $f_9 \rightarrow f_2$ 的路径系数为 0.166 ($P=0.000$),故“刑事司法公开”通过“无罪推定”这一中间效应对“刑事司法公正”潜变量产生间接影响。根据计算,禁止刑讯逼供、警察平等执法、刑事司法机关独立行使职权、刑事司法公开等 4 个因素对刑事司法公正具有间接影响,其效应分别为:0.033 (0.257×0.127)、0.019 (0.146×0.127)、0.038 (0.301×0.127)、0.01 (0.166×0.06)。

(二)刑事法治专家评估指数二阶单因素模型

从刑事法治专家评估指数结构方程模型检验结论出发,如果能证明假设 4 成立,即证明刑事司法公正对刑事法治具有最强解释力,那么我们就可以通过对刑事司法公正影响因素的测量来创制刑事法治专家评估指数。基此,可创建刑事法治专家评估指数二阶单因素模型。模型的组成包括:一是 38 个观测变量(评估指标)。首先,采用验证性因素分析方法,以 x_1-x_{41} 共 41 个评估指标为观测变量,以 f_1-f_{10} 共 10 个二级指数为一阶潜变量,以“刑事法治(f_0)”为二阶潜变量,初步创建刑事法治专家评估二阶单因素模型。其次,根据初步创建的模型标准化参数估计结果,“刑事司法机关独立行使职权(f_{10})”潜变量 \rightarrow “法官办案不受犯罪组织影响(x_{40})”观测变量的标准化载荷系数高于 1.00,说明模型存在多重共线性问题。在删除 f_{10} 潜变量所涵盖的 5 个观测变量($x_{39}-x_{43}$)后,模型多重共线性消失。基此,刑事法治专家评估指数二阶单因素模型的观测变量由 x_1-x_{38} 共 38 个观测变量组成。二是 9 个一阶潜变量,包括 f_1-f_9 。三是 1 个二阶潜变量(f_0),即假设 9 个一阶潜变量可以共同解释一个更高级的变量“刑事法治(f_0)”。模型的路径走向为 f_0 指向 f_1-f_9 。基此,可创建刑事法治专家评估指数二阶单因素模型。

经信度检验,刑事法治专家评估指数二阶单因素模型所涵盖的 38 个观测变量的 Cronbach's Alpha 系数为 0.841,说明模型的组合信度较高。结合模型拟合信息,在根据 M. I. 指标值对模型进行修正后,模型卡方值(3038.693)与自由度(651)之比为 4.67,符合小于 5 的标准要求。卡方检验的显

著性水平 $P=0.000$,符合小于 0.0001 的标准要求。 $RMSEA$ 为 0.053,符合小于 0.08 的标准要求; $SRMR$ 为 0.074,未满足小于 0.05 的标准要求。 CFI 、 TLI 值分别为 0.924、0.918,符合大于 0.9 的标准要求。据此,刑事法治专家评估指数二阶单因素模型拟合理想。根据刑事法治专家评估指数二阶单因素参数估计标准化结果,可得出如下结论:

一是刑事司法公正对刑事法治具有最强解释力。根据路径系数分析,在模型的 9 条路径中,“刑事法治(f_0)”指向“刑事司法公正(f_4)”的路径系数最高,为 0.994($P=0.000$),说明“刑事司法公正”一阶潜变量有效反映了“刑事法治”二阶潜变量的意义,刑事法治专家评估指数可以通过对刑事司法公正的测量来创制。据此,假设 4 得以有效验证。

二是刑事司法廉洁、刑事司法效率对刑事法治具有强解释力。根据路径系数分析,“刑事法治(f_0)”指向“刑事司法廉洁(f_3)”“刑事司法效率(f_5)”的路径系数分别为 0.941($P=0.000$)、0.897($P=0.000$),说明“刑事司法廉洁”和“刑事司法效率”2 个一阶潜变量,有效反映了“刑事法治”二阶潜变量的意义,刑事法治专家评估指数可以通过对刑事司法廉洁、刑事司法效率的测量来创制。

三是监狱罪犯教育改造效率、刑事司法效果对刑事法治具有较强解释力。根据路径系数分析,“刑事法治(f_0)”指向“监狱罪犯教育改造效率(f_7)”“刑事司法效果(f_8)”的路径系数分别为 0.471($P=0.000$)、0.421($P=0.000$),这说明“监狱罪犯教育改造效率”和“刑事司法效果”2 个一阶潜变量,有效反映了“刑事法治”二阶潜变量的意义,刑事法治专家评估指数可以通过对监狱罪犯教育改造效率、刑事司法效果的测量来创制。

四是禁止刑讯逼供、无罪推定、警察平等执法、刑事司法公开 4 个因素对刑事法治具有一定程度的解释力。根据路径分析,“刑事法治(f_0)”指向“禁止刑讯逼供(f_1)”“无罪推定(f_2)”“警察平等执法(f_6)”“刑事司法公开(f_9)”等 4 条路径的系数分别为 0.263($P=0.000$)、 -0.04 ($P=0.193$)、0.239($P=0.000$)、0.012($P=0.726$),均未满足大于 0.3 的标准要求,且其中两条路径系数的显著性水平大于 0.05,因此“禁止刑讯逼供”等 4 个一阶潜变量对“刑事法治”二阶潜变量有一定程度的解释力。相较上述五个因素,其解释力要弱一些。

综合上述,刑事法治专家评估指数结构方程模型检验结论与二阶单因素模型检验结论一致。对刑事司法公正具有直接显著性正向影响的因素,包括刑事司法廉洁、刑事司法效率、监狱罪犯教育改造效率、刑事司法效果,可被用来创制刑事法治专家评估指数;比较而言,对刑事司法公正影响小或不具有直接影响的因素,可以不被用来创制刑事法治专家评估指数。根据结构方程模型和二阶单因素模型检验结果,刑事司法廉洁、刑事司法公正、刑事司法效率、监狱罪犯教育改造效率、刑事司法效果 5 个因素所涵盖 23 的个指标可被用来创制刑事法治专家评估指数。当然,这并不是说没有被用来创制刑事法治专家评估指数的因素,对刑事法治不具有理论和实践意义,而是说刑事法治专家评估指数可以不通过对这些潜变量的直接测量来创制。这是因为,首先,刑事司法公正潜变量已涵盖嫌疑人获得及时公正审理(x_{10})、独立行使审判权(x_{12})、法官没有偏见(x_{13})测量指标,因此,刑事法治专家评估指数实际上已涵盖了对刑事司法机关独立行使职权和司法平等、无罪推定的测量。其次,根据测量结论,禁止刑讯逼供等其他 4 个潜变量通过中间变量而对刑事司法公正具有间接影响,故刑事法治专家评估指数的创制,实际上涵盖了对前述 4 个潜变量的间接测量。再次,在现代法治条件下,司法公开实践日趋完善,因此对刑事法治专家评估测量来说,因其超越了测量的界值,故可能不再具有测量上的意义。

三、普通公众评估指数模型的创制

在如何科学创建刑事法治指数上,同样必须解决的一个关键问题是:普通公众评估指数研究所构想的 25 个指标究竟有哪些可被用来测量刑事法治? 基此,笔者采用与前述相同的方法,创建刑事法治普通公众评估指数结构方程模型和二阶单因素模型。

(一)刑事法治普通公众评估指数结构方程模型

基于普通公众调查数据和理论假设,笔者采用验证性因素分析方法,创建刑事法治普通公众评估指数结构方程模型。模型组成包括:一是25个观测变量(y_1-y_{25});二是6个潜变量,包括“刑事司法廉洁(f_3)”“刑事司法公正(f_4)”“刑事司法效率(f_5)”“警察平等执法(f_6)”“刑事司法公开(f_9)”“刑事司法机关独立行使职权(f_{10})”;三是5条路径走向,包括 $f_3 \rightarrow f_4$ 、 $f_5 \rightarrow f_4$ 、 $f_6 \rightarrow f_4$ 、 $f_9 \rightarrow f_4$ 、 $f_{10} \rightarrow f_4$ 。经信度检验,公众评估指数结构方程模型组合信度较高,25个观测变量Cronbach's Alpha系数为0.846。结合模型拟合信息,在根据M.I.指标值对模型进行修正后,模型卡方值(827.203)与自由度(255)之比为3.24,符合小于5的标准要求。RMSEA为0.042,符合小于0.08的标准要求;SRMR为0.049,符合小于0.05的标准要求。CFI、TLI值分别为0.963、0.957,符合大于0.9的标准要求。据此,刑事法治普通公众评估指数结构方程模型拟合理想。根据模型参数标准化估计结果,可得出如下结论:

一是刑事司法机关独立行使职权对刑事司法公正具有很大影响力。在模型的5条路径中, $f_{10} \rightarrow f_4$ 的路径系数最高,为0.684($P=0.000$),说明在其他条件不变时,“刑事司法机关独立行使职权”潜变量每提升1个单位,“刑事司法公正”潜变量提升0.684个单位。据此,假设10得以有效验证。

二是刑事司法廉洁、警察平等执法对刑事司法公正具有较大影响力。 $f_3 \rightarrow f_4$ 、 $f_6 \rightarrow f_4$ 的路径系数分别为0.207($P=0.000$)、0.185($P=0.000$),这说明在其他条件不变时,“刑事司法廉洁”“警察平等执法”潜变量每提升1个单位,“刑事司法公正”潜变量分别提升0.207、0.185个单位。据此,假设3、6得以有效验证。

三是刑事司法效率对刑事司法公正具有一定程度的影响力。 $f_5 \rightarrow f_4$ 的路径系数为0.07($P=0.001$),说明在其他条件不变时,“刑事司法效率”潜变量每提升1个单位,“刑事司法公正”潜变量提升0.07个单位。据此,假设5得以有效验证。但与上述3个因素相比,其影响力要小一些。

四是刑事司法公开对刑事司法公正通过中间变量发挥间接影响力。 $f_9 \rightarrow f_4$ 的路径系数为0.19($P=0.379$),未通过显著性水平,说明“刑事司法公开”潜变量对“刑事司法公正”潜变量不具有直接的显著性正向影响。据此,假设9未得到有效验证。为检验刑事司法公开对刑事司法公正是否具有间接影响,笔者对普通公众评估指数结构方程模型路径进行了修改。根据路径修改后模型标准化参数估计结果,“刑事司法公开”潜变量通过“警察平等执法”潜变量对“刑事司法公正”具有间接影响($P=0.002$)。根据计算,其所具有的间接影响效应为0.0076(0.101×0.075)。

(二)刑事法治普通公众评估指数二阶单因素模型

从刑事法治普通公众评估指数结构方程模型测量结论出发,如果能证明刑事司法公正对刑事法治具有最强解释力,则可通过对刑事司法公正具有显著性正向影响因素的测量来创制刑事法治指数。据此,可创建刑事法治普通公众评估指数二阶单因素模型。模型创建步骤包括:一是21个观测变量的确定。根据初始模型检验,在将“刑事司法公开(f_9)”纳入模型时模型无法拟合,故将其所涵盖的4个指标予以删除。据此,普通公众评估指数二阶单因素模型的观测变量有21个,包括 y_1-y_{18} 、 $y_{23}-y_{25}$ 。二是21个观测变量的重新聚类。为确定21个观测变量的聚合状况,笔者采用探索性因素分析法进行聚类,把“刑事司法公正(f_4)”所涵盖的5个观测变量和“刑事司法机关独立行使职权(f_{10})”所涵盖的3个观测变量聚合为同一个因素,其余观测变量的聚合则与普通公众指数结构方程模型中观测变量的聚合一致。根据法治一般原理,司法机关独立行使职权是司法公正的重要前提和保证,因此可将新的聚合因素(f_4 和 f_{10})仍然命名为“刑事司法公正(f_4)”。据此,组成普通公众评估指数二阶单因素模型的一阶潜变量有4个,包括“刑事司法廉洁(f_3)”“(f_4)刑事司法公正(包括刑事司法机关独立行使职权)”“刑事司法效率(f_5)”和“警察平等执法(f_6)”。模型路径走向为“刑事法治(f_0)”二阶潜变量分别指向4个一阶潜变量,一阶潜变量指向21个观测变量。

经信度检验,刑事法治普通公众评估指数二阶单因素模型组合信度较高,21个观测变量的Cronbach's Alpha系数为0.861。结合模型拟合信息,在根据M.I.指标值对模型进行修正后,模型

卡方值(895.049)与自由度(181)之比为 4.95,符合小于 5 的标准要求。*RMSEA* 为 0.055,符合小于 0.08 的标准要求;*SRMR* 为 0.049,符合小于 0.05 的标准要求。*CFI*、*TLI* 值分别为 0.946、0.937,符合大于 0.9 的标准要求。据此,刑事法治普通公众评估指数二阶单因素计算模型拟合理想。根据模型参数估计标准化结果,可以得出如下结论:

一是刑事司法公正对刑事法治具有最强解释力。根据路径分析,在模型的 4 条路径中,“刑事法治(f0)”指向“刑事司法公正(f4)”的载荷系数最高,为 0.836($P=0.000$),说明“刑事司法公正”一阶潜变量有效反映了“刑事法治”二阶潜变量的意义,刑事法治公众评估指数可以通过对刑事司法公正的测量来创制。据此,假设 4 得以有效验证。

二是刑事司法廉洁、警察平等执法对刑事法治具有较强解释力。根据路径分析,“刑事法治(f0)”潜变量指向“刑事司法廉洁(f3)”“警察平等执法(f6)”潜变量的载荷系数分别为 0.571($P=0.000$)、0.437($P=0.000$),这说明“刑事司法廉洁”“警察平等执法”2 个一阶潜变量有效反映了“刑事法治(f0)”二阶潜变量的意义,刑事法治普通公众评估指数可以通过对刑事司法廉洁、警察平等执法的测量来创制。

三是刑事司法效率对刑事法治具有一定程度的解释力。根据路径分析,“刑事法治(f0)”潜变量指向“刑事司法效率(f5)”潜变量的载荷系数为 0.21($P=0.000$),未能满足大于 0.3 的标准要求,因此,“刑事司法效率(f5)”一阶潜变量能在一定程度上反映“刑事法治(f0)”二阶潜变量的意义,其解释力比上述 3 个因素要弱一些。

综合上述,刑事法治普通公众评估指数结构方程模型与二阶单因素模型的检验结论一致。对刑事司法公正具有直接显著性正向影响的因素,包括刑事司法机关独立行使职权、刑事司法廉洁、警察平等执法,均可被用来创制刑事法治公众评估指数;比较而言,对刑事司法公正影响小或不具有直接影响的因素,可以不被用来创制刑事法治公众评估指数。据此,根据结构方程模型和二阶单因素模型测量结论,刑事司法廉洁、刑事司法公正、刑事司法机关独立行使职权、警察平等执法 4 个因素所涵盖的 17 个指标可被用来创制刑事法治公众评估指数。需要指出的是,刑事司法效率、刑事司法公开 2 个因素所涵盖的 8 个指标可以不被用来创制刑事法治普通公众评估指数,这一结论不是否定刑事司法效率、刑事司法公开的必要性。一方面,这只是一种基于司法测量的结果。另一方面,前文已指出,“刑事司法公开”通过“警察平等执法”对刑事司法公正发挥间接影响力。

四、刑事法治指数的创制

综合前文,刑事法治指数可以通过对刑事司法廉洁、刑事司法公正、刑事司法效率、刑事司法效果、监狱罪犯教育改造效率、警察平等执法、刑事司法机关独立行使职权 7 个因素的测量来创制。那么,可据以测量刑事法治的因素和指标的权重究竟如何确定呢?国内法治指数相关研究大多采用专家主观打分法直接分配指标权重。显然,主观权重算法的科学性及其结论的准确性明显低于客观权重算法。本文试图通过刑事法治专家评估指数和普通公众评估指数二阶单因素计算模型的创建,来客观计算指标权重和二级指数权重。在此基础上,运用加权算术平均法创制刑事法治专家评估指数和普通公众评估指数,然后按刑事法治专家评估指数和普通公众评估指数各占 50%的比率,合成刑事法治指数。

(一)刑事法治专家评估指标权重及二级指数权重计算

根据刑事法治专家评估指数二阶单因素模型检验结果,运用可据以测量刑事法治的指标来创建刑事法治专家评估指数二阶单因素计算模型(图 1)。计算模型由以下因素组成:一是“刑事法治专家评估指数(f0)”1 个二阶潜变量(一级指数);二是“刑事司法廉洁(f3)”“刑事司法公正(f4)”“刑事司法效率(f5)”“监狱罪犯教育改造效率(f7)”“刑事司法效果(f8)”5 个一阶潜变量(二级指数);三是由 x_7-x_{19} 、 $x_{26}-x_{35}$ 组成的 23 个观测变量(指标)。经信度检验,模型 Cronbach's Alpha 系数为 0.827,

说明模型组合信度较高。根据模型拟合信息,卡方值(1054.105)与自由度(217)比为 4.86,符合小于 5 的标准要求。模型卡方检验的显著性水平 $P=0.000$,符合小于 0.0001 的标准要求。 $RMSEA$ 为 0.054,符合小于 0.08 的标准要求; $SRMR$ 为 0.042,符合小于 0.05 的标准要求。 CFI 、 TLI 值分别为 0.96、0.953,符合大于 0.9 的标准要求。据此,中国刑事法治专家评估指数二阶单因素计算模型拟合理想。根据计算模型标准化参数估计结果,可算出 23 个刑事法治专家评估指标和 5 项二级指数的权重,具体计算方法如下。

一是评估指标权重计算。根据 5 个一阶潜变量(二级指数)指向各自所涵盖观测变量(评估指标)的因素载荷系数(图 1)计算指标权重。指标权重计算方法为:各二级指数名下某一指标载荷系数与该二级指数名下各指标载荷系数和之比。例如:“警察廉洁(x7)”指标权重计算: $0.792/0.792+0.85+0.858=0.729/2.5=0.32$ 。据此,可算出 23 个专家评估指标的权重^①。

二是二级指数权重计算。根据“刑事法治专家评估指数(f0)”二阶潜变量指向 5 个一阶潜变量的载荷系数,计算各二级指数权重。计算方法为:某二级指数载荷系数与二级指数载荷系数和之比。例如:“刑事司法廉洁(f3)”权重为: $0.929/0.929+0.977+0.917+0.491+0.407=0.929/3.721=0.25$ 。据此,可算出 5 项刑事法治专家评估二级指数的权重^②。

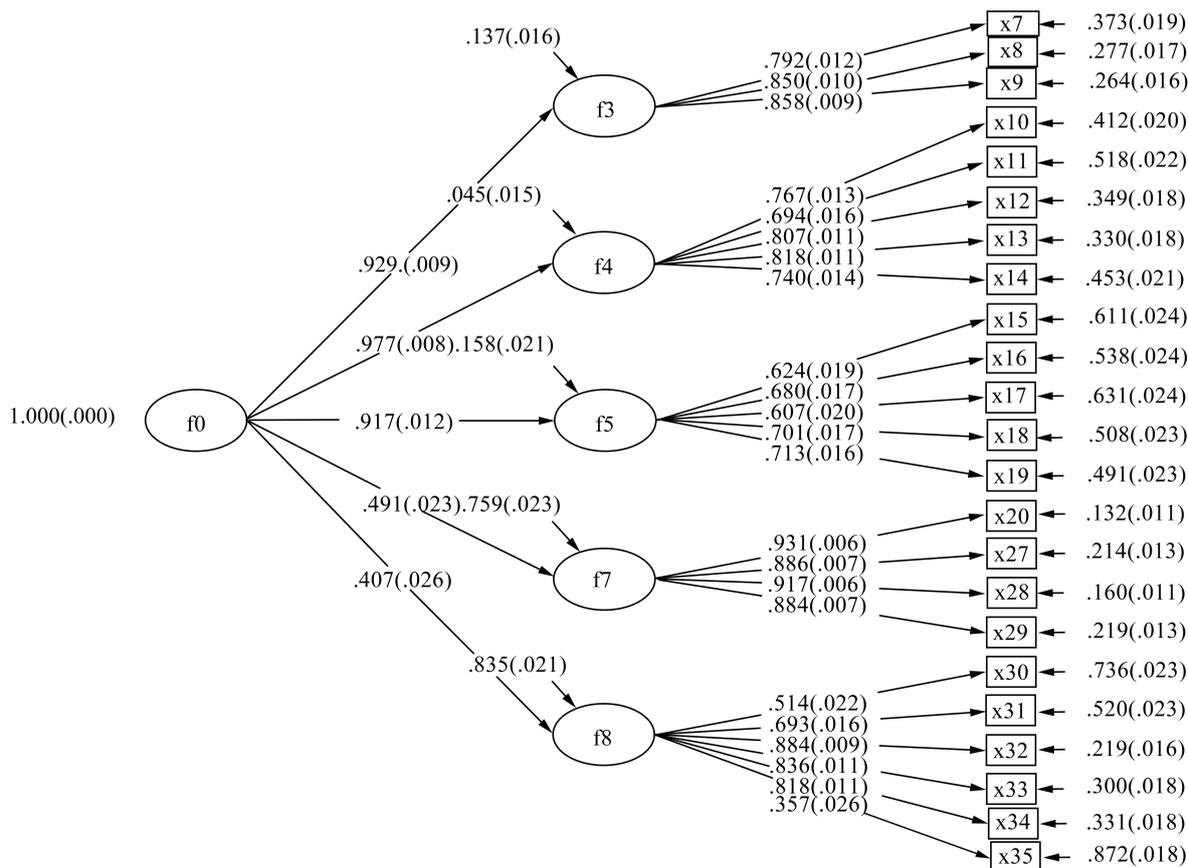


图 1 刑事法治专家评估指数二阶单因素计算模型

① x7、x8、x9 的指标权重为 0.32、0.34、0.34；x10、x11、x12、x13、x14 的指标权重为 0.20、0.18、0.21、0.21、0.19；x15、x16、x17、x18、x19 的指标权重为 0.19、0.20、0.18、0.21、0.21；x26、x27、x28、x29 的指标权重为 0.26、0.24、0.25、0.24；x30、x31、x32、x33、x34、x35 的指标权重为 0.13、0.17、0.22、0.20、0.20、0.09。

② 刑事法治专家评估刑事司法廉洁(f3)、刑事司法公正(f4)、刑事司法效率(f5)、监狱罪犯教育改造效率(f7)、刑事司法效果(f8)二级指数的权重为 0.25、0.26、0.25、0.13、0.11。

(二)普通公众评估指标权重及二级指数权重计算

采用上述相同方法,用可据以测量刑事法治的17个普通公众评估指标创建刑事法治普通公众评估指数二阶单因素计算模型。计算模型由如下因素组成:一是“普通公众评估指数(f0)”1个二阶潜变量,二是“刑事司法廉洁(f3)”“f4刑事司法公正(包括司法机关独立行使职权)”“警察平等执法(f6)”等3个一阶潜变量,三是由 $y_1 - y_8$ 、 $y_{23} - y_{25}$ 、 $y_{13} - y_{18}$ 组成的17个观测变量。经信度检验,模型Cronbach's Alpha系数为0.875,说明模型组合信度较高。根据模型拟合度信息表,卡方值(524.793)与自由度(109)之比为4.81,符合小于5的标准要求。模型卡方检验的显著性水平 $P=0.000$,符合小于0.0001的标准要求。RMSEA为0.054,符合小于0.08的标准要求;SRMR为0.044,符合小于0.05的标准要求。CFI、TLI值分别为0.963、0.954,符合大于0.9的标准要求。基此,中国刑事法治普通公众评估指数二阶单因素计算模型拟合理想。根据计算模型标准化参数估计结果,采用前述相同方法,可算出17个普通公众评估指标权重和3项二级指数权重^①。

(三)刑事法治指数的制成

依据计算得出的指标权重、二级指数权重和指标标准值,采用加权算术平均法,即可制成刑事法治指数,具体步骤包括:

一是二级指数的计算。二级指数计算方法为:各指标标准值与其权重乘积之和的算术平均值。例如:我国刑事法治专家评估刑事司法廉洁二级指数 $=0.4(x_7 \text{ 标准值的均值}) \times 0.32(x_7 \text{ 权重}) + 0.47(x_8 \text{ 标准值的均值}) \times 0.34(x_8 \text{ 权重}) + 0.47(x_9 \text{ 标准值的均值}) \times 0.34(x_9 \text{ 权重}) = 0.45$ 。据此,可算出5项刑事法治专家评估二级指数和3项普通公众评估二级指数。

二是一级指数的计算。一级指数计算方法为:各二级指数与其权重乘积之和。经计算,中国刑事法治专家评估指数 $=0.45(\text{刑事司法廉洁专家评估指数}) \times 0.25(\text{权重}) + 0.43(\text{刑事司法公正专家评估指数}) \times 0.26(\text{权重}) + 0.43(\text{刑事司法效率专家评估指数}) \times 0.25(\text{权重}) + 0.46(\text{监狱罪犯教育改造效率专家评估指数}) \times 0.13(\text{权重}) + 0.63(\text{刑事司法效果专家评估指数}) \times 0.11(\text{权重}) = 0.46$ 。中国刑事法治普通公众评估指数 $=0.64(\text{刑事司法廉洁普通公众评估二级指数}) \times 0.33(\text{权重}) + 0.60(\text{刑事司法公正普通公众评估二级指数}) \times 0.45(\text{权重}) + 0.87(\text{警察平等执法普通公众评估二级指数}) \times 0.22(\text{权重}) = 0.67$ 。

三是刑事法治指数的合成。按照刑事法治专家评估和普通公众评估各占50%的原则,合成刑事法治指数。经计算,中国刑事法治指数 $= (0.46 + 0.67) \div 2 = 0.57$ 。

综上,通过实证调查和刑事法治指数结构方程模型、二阶单因素模型创建而制成的中国刑事法治指数,是测量人们对于刑事法治认同程度的量具,可用来测量全国或各省、市的刑事法治状况。根据本研究测量结论,人们对我国刑事法治的认同程度为0.57。通过比较,本研究测得的中国刑事法治指数(0.57)高于2014-2019年世界正义工程关于中国刑事法治指数的5年均值(0.46)11个百分点^②。基此,根据本研究测量结论,我国刑事法治程度处于世界中等水平的上端,而根据世界正义工程法治指数测量结论,我国刑事法治程度处于世界中等水平的下端。若以2019年世界正义工程法治指数测

① y_1, y_2, y_3 的指标权重为0.31、0.36、0.33; $y_4, y_5, y_6, y_7, y_8, y_{23}, y_{24}, y_{25}$ 的指标权重为0.11、0.11、0.14、0.12、0.12、0.13、0.13、0.15; $y_{13}, y_{14}, y_{15}, y_{16}, y_{17}, y_{18}$ 的指标权重为0.11、0.17、0.18、0.2、0.17、0.17。公众评估刑事司法廉洁(f3)、f4刑事司法公正(含司法机关独立行使职权)、警察平等执法(f6)二级指数的权重为0.33、0.45、0.22。

② 参见The World Justice Project, *The WJP Rule of Law Index 2014*, p. 80; *The WJP Rule of Law Index 2015*, p. 76; *The WJP Rule of Law Index 2016*, p. 69; *The WJP Rule of Law Index 2017-2018*, p. 69; *The WJP Rule of Law Index 2019*, p. 58, <https://worldjusticeproject.org/our-work/wjp-rule-law-index>, 访问日期:2020年3月31日。

量为参照,根据本测量结论,我国刑事法治程度将提升 23 名,至世界前 34 位^①。

为检验本测量结论的准确性,兹将依据客观数据计算得到的刑事裁判准确度与本测量结论进行对比。基此,笔者收集了 1988 年至 2017 年共 30 年的刑事上诉、再审改判发回案件数与一审结案数^②。经计算,30 年间刑事上诉、再审改判发回案件数与一审刑事案件结案数之比的均值为 0.039,如果将未改判发回案件数与一审刑事案件审结数之比作为裁判准确度的测量值,则我国 30 年间刑事裁判准确度绝对数平均为 0.961。按照极值法对数据进行无量纲化处理,我国 30 年间刑事裁判准确度相对数均值为 0.83。根据本测量结论,法院刑事裁判准确度(x33)为 0.72(为刑事法治专家评估指标中得分最高的一个)。虽然前者高于后者 11 个百分点,但未改判发回案件数与一审刑事案件审结数之比,显然不能完全反映裁判准确度这一客观真实。总的来说,通过比较世界正义工程关于中国刑事法治指数的测量结论和基于客观数据调查与计算得出的结论,本研究制成的中国刑事法治指数具有较高的准确性。

A Study of the Creation of the Criminal Rule of Law Index in China

Zeng Yun

(Law School, Guangzhou University, Guangzhou 510006, P. R. China)

Abstract: In order to create a comprehensive, objective and scientific index of criminal rule of law in China, first of all, it is necessary to establish a comprehensive research constructs and theoretical hypotheses from two aspects which are the evaluation of legal experts and the general public. Therefore two types of questionnaires are created, namely, expert assessment and public evaluation. Then empirical surveys are carried out by cluster random sampling, and 1312 survey samples are collected from questionnaires of legal experts and 1289 survey samples from questionnaires of the general public. Secondly, the sample data was non-dimensionalized by extreme value method. The structural equation model and the second-order single factor model of criminal law index of the evaluation of legal experts and the general public are created by the confirmatory factor analysis method. According to the test results, 5 among the 10 research constructs and theoretical hypotheses from the evaluation of legal experts have been validated; 4 of the 6 research constructs and theoretical hypotheses from the evaluation of the general public have been validated. Thirdly, the second-order single factor calculation models of criminal law index of the evaluation of legal experts and the general public are established. Then weights of indicators and the second-order indices are objectively determined through the model calculation. At last index of the criminal rule of law of China is created by weighted arithmetic average method.

Keywords: The criminal rule of law; Rule of law index; Factor analysis

[责任编辑:李春明]

^① 2019 年,我国刑事法治指数(0.47)在世界正义工程所调查的 126 个国家和司法区域中排名第 57 位。参见 The World Justice Project, *The WJP Rule of Law Index 2019*, p. 29, <https://worldjusticeproject.org/our-work/wjp-rule-law-index>, 访问日期:2020 年 3 月 31 日。

^② 1988 年-2017 年刑事一审结案数和二审、再审改判发回案件数来自各年度的《中国法律年鉴》。