

量刑失衡的克服:模式与选择

蔡曦蕾*

摘 要 在世界范围内,量刑失衡的克服模式可以分为传统型、指南型与信息型三种,而后两者无疑在克服量刑失衡问题上被寄予厚望。在指南型模式中,美国联邦及各州、加拿大、英国英格兰和威尔士、韩国以及新西兰的量刑指南实践及实效值得关注;在信息型模式中,加拿大四省、澳大利亚新南威尔士州、英国苏格兰的量刑信息系统的结构及实效也应加以重视。在克服量刑失衡的进路上,中国应当选择一条以指南型模式为基的渐进式进路。本文以33512份样本组成的数据库为分析对象,通过对最高人民法院量刑规范化改革进行实证评估,证明了量刑指南在中国的可行性,并就构建中国量刑指南的量刑起点与量刑情节调节基准刑的经验根基进行了实证研究。

关键词 指南型 信息型 量刑规范化改革

前 言

量刑失衡,是一种违背公正的量刑现象。对于中国量刑失衡的严重程度,我国已有学者进行了开创性研究。该学者给予一百余名审判人员150个盗窃犯罪的案例,要求他们不臆造其他与量刑有关的情节,仅根据每个案件中给定的事实作出量刑,结果显示:如果以和平均量刑值的正负误差是否在1年以内作为评价量刑失衡的标准,那么,有30%左右的量刑与均值的正负误差在1年以上,属于失衡的量刑。^{〔1〕}由此可见,中国司法实践中量刑失衡现象已经相当严重。然而,在量刑失衡问题上,中国的情况绝非个例。早在1989年,荷兰学者博格惠斯

* 北京警察学院讲师。本文受2014年教育部人文社会科学研究青年基金项目(项目号:14YJC820002)以及2014年北京市社会科学基金青年项目(项目号:14FXC022)资助。

〔1〕 参见苏惠渔、张国全、史建三:《量刑与电脑:量刑公正合理应用论》,百家出版社1989年版,页160—179。

(Berghuis)就对荷兰 19 个地区法院的 7 类犯罪^{〔2〕}的量刑进行过实证研究,并认为那些对某一类犯罪习惯于判处更严重刑罚的法院在对其他犯罪类型上,也同样习惯用重刑,量刑失衡普遍存在于这些法院的量刑之中;^{〔3〕}在德国,“虽然德联邦 16 个州之间适用的是同一部刑法,但实证研究揭示出,在量刑上各地区之间存在着明显的差异。比如,非法持有毒品在法兰克福和科隆一般都会免除处罚,但是同样的案情在柏林和慕尼黑则至少会被判处罚金”,^{〔4〕}也正因此,托马斯·魏根特(Thomas Weigend)教授才笃定地认为:“在德国不同的城市以及区域中,量刑标准存在着非常显著的差异”;^{〔5〕}在美国,学者威廉·奥斯汀(Willam Austin)和托马斯·A. 威廉姆斯三世(Thomas A. Williams III)于 1975 邀请了 47 名弗吉尼亚州地区法官参与量刑实验,每位法官都被分发到了一本小册子,上面包括了五个不同的案件的具体案情,并要求每位法官根据案情给出适当的定罪与量刑结论。以第四个案件的反馈为例:该起案件为盗窃案,共有 41 名法官(占全体法官的 87.2%)认为构成犯罪,法官在定罪上的一致性较高。但是,对被告人的量刑却有 9 种不同的形式,12 名法官(占 41 名法官的 29.3%)对被告人判处有期徒刑以及缓期监禁刑;7 名法官(占 41 名法官的 17.1%)对被告人单处缓期监禁刑;7 名法官(占 41 名法官的 17.1%)对被告人判处监禁刑,但缓期执行部分刑期,并认为被告人应执行的实际刑期从 10 天到 90 天不等;5 名法官(占 41 名法官的 12.2%)对被告人判处有期徒刑加监禁刑,程度从 25 美元罚金加 10 天监禁到 150 美元罚金加 60 天监禁不等;剩下的 8 位法官则判除了剩下的 5 种刑罚方式。^{〔6〕}

由上可知,量刑失衡是一个世界性难题,在克服量刑失衡的进路上,中国并不孤单。其他国家或地区在克服量刑失衡的过程中,形成了很多颇有成效的方法,只要经过适当调试,这些方法在世界范围内具有共通性。正因此,各国克服量刑失衡的方法已引发了国内外学者的广泛关注,其中的集大成者为美国的麦克·托尼(Michael Tonry)教授,他曾将西方国家的克服量刑失衡的主要方法进行过归纳,将其分为“强制法定刑标准”、“假定数字化司法指南”、“假定数字化行政指南”、“建议型以及信息刑量刑指南”、“上诉指南裁决书”、“法定量刑规则”六种具体方案,并逐个进行了详细地评介,^{〔7〕}我国也有学者对其中的一种或几种进行过介绍。然而,零散地对这些方案进行评介,既缺乏系统性,也不利于从整体上进行采纳与吸收。从宏观视角来看,如果能够从模式化的高度,将量刑失衡具体的克服方案按照运作理念、环境、方式进

〔2〕 该 7 类犯罪分别为殴打罪、毁坏罪、一般盗窃罪、加重盗窃罪、伪造罪、无资格驾驶罪以及毒品犯罪,该 7 类犯罪占据了荷兰司法实践发案率的 2/3。

〔3〕 转引自 Josine Junger-Tas, “Sentencing in the Netherlands: Context and Policy”, 7 *Federal Sentencing Reporter*, 295(1995)。

〔4〕 Cornelius Nestler, “Sentencing in Germany”, 7 *Buffalo Criminal Law Review*, 110(2003)。

〔5〕 Thomas Weigend, “Sentencing in West Germany”, 42 *Maryland Law Review*, 81—82(1983)。

〔6〕 参见 Willam Austin, Thomas A. Williams III, “A Survey of Judges’ Responses to Simulated Legal Cases: Research Note on Sentencing Disparity”, 68 *The Journal of Criminal Law & Criminology*, 307—308 (1977)。

〔7〕 参见 Michael Tonry, “Punishment Policies and Patterns in Western Countries”, in Michael Tonry, Richard S. Frase ed., *Sentencing and Sanctions in Western Countries*, Oxford University Press, 2001, pp.21—24。

行区分,将其分为几大模式,对其拔高到模式论的层面上进行探讨,这对于我国选择何种路径来应对量刑失衡现象将大有帮助。

从模式化的高度,笔者将主要国家或地区量刑失衡克服模式分为如下三类,即:传统型模式、指南型模式与信息型模式。后两种模式将在下文详述,此处暂且按下不表。传统型模式寄望于在不对现有量刑体系进行大幅度变革的前提下,通过完善现有的方法来达到克服量刑失衡的目的。这些方法主要有:在成文法以及判例法上明确规定量刑基准、完善量刑说理制度与量刑建议制度、加强上级法院对下级法院量刑的审查等。这些克服方法普遍存在于世界各国司法体制中,并不为某一国家或地区所独有。

虽然传统型模式是克服量刑失衡三极中的一元,但本文中却欲略去对其的探讨,理由有二:一方面,传统型模式中各方法对于量刑失衡的克制效应并不显著,早已饱受诟病。以量刑基准制度为例,在德国,虽然存在着以“责任原则、矫正被告人原则、规范确认原则、特殊预防以及一般预防原则”构建成的一整套量刑基准制度,但其对于司法实践却影响甚微。魏根特教授对此毫不讳言:“多数量刑原则的研究者已经意识到了,量刑原则仅仅只是一个纯学术价值的主题。德国法官在他们日常的量刑裁量中对应根据何种量刑原则毫不在意”;〔8〕再以量刑说理制度为例,其本应成为克服量刑失衡的利器,但在实践中却退化成“鸡肋式”的制度。在荷兰,“法官们在裁决文书中经常以一些标准段落加以应付,比如:本量刑与犯罪行为的严重性程度、实施犯罪行为时的情境以及被告人的人格相适应”。〔9〕在德国,“书面式的量刑说理演化成机械式的套话,比如‘在充分考虑了犯罪行为的本质以及被告人的人格后,如下的量刑是充分且必要的’,或者,法官仅仅只是简要的将量刑情节加以例举,但对于这些量刑情节与量刑结果之间的关联,却不作任何说明”,〔10〕这样的量刑说理制度对量刑失衡的克制作用是有限的。另一方面,近些年国内学界对于传统型模式下的诸制度之介绍与研究已较充分,故本文中对传统型模式予以省略。

一、指南型克服模式

指南型克服模式一般通过建立独立的或附属于司法机关之下的量刑委员会,由委员会通过对法官量刑经验的实证研究,制定规定所有犯罪或部分犯罪各种量刑情节的刑罚幅度,形成体系化的量刑指南,以此来指导司法实践中法官的量刑。综观世界范围内指南规范的指引方式以及所统摄的犯罪范围,可以将该模式区分为两类,一类可称为数字格状量刑指南,一类可称为文字叙述量刑指南。前者采取数字格状(Numerical Grid)的形式,将罪刑关系加以精确数字化,并试图对所有犯罪的量刑进行规定,其代表为美国联邦及其多数州的量刑指南;而后

〔8〕 Thomas Weigend, *Supra* note 5, at 80.

〔9〕 Peter J.P. Tak, “Sentencing in the Netherlands: Discretion and Disparity”, 7 *Federal Sentencing Reporter*, 301(1995).

〔10〕 Thomas Weigend, *Supra* note 5, at 67.

者则采取文字叙述的形式,其代表为加拿大量刑委员会、韩国、英格兰和威尔士、新西兰等的量刑指南。

(一)数字格状量刑指南

数字格状量刑指南的杰出代表当属美国,但具体在量刑指南的建立上,又可以区分成联邦与州两个阵营。前者根据《美国法典》第18章“犯罪”(USC—Title 18—Crimes)的规定,管辖联邦法官对联邦犯罪的量刑,但是,归联邦量刑指南管辖的案件在全美整体发案率上所占比重较小,一般在10%左右,^{〔11〕}而各州量刑指南所起的实际作用更大。下文将对两者分别介绍。

1. 州量刑指南

(1)发展历程。在美国建立美国量刑委员会并着手制定联邦量刑指南之前,各州在假释指南的影响下,就已经开始了州量刑指南的实践。^{〔12〕}1979年1月,犹他州开始在实验层面上实施量刑指南;紧接着,到1980年1月,阿拉斯加州也开始了量刑指南的实践,但两者都没有形成固定的量刑委员会,而且其内容也与之之后成熟的指南大相径庭;美国第一部完备的量刑指南来自明尼苏达州(下文简称明州),其生效时间为1980年5月,其也成为美国历史上第一部标榜为假定指南的强制量刑指南,区别于之前各州的建议量刑指南。在此之后,美国量刑指南就存在假定与建议两个流派。^{〔13〕}在明州指南生效不到两年,1982年7月,宾夕法尼亚州量刑指南生效,虽然没有明确表明自己假定量刑指南的身份,但是,由于其也存在低限度的上诉法院审查制度,因此,其亦属于假定量刑指南的范畴。根据美国学者凯文·利兹(Kevin R. Reitz)的统计,截止到2008年,陆续又有20个州(包括华盛顿特区)已经或即将建立正式的量刑指南体系。^{〔14〕}

(2)杰出代表。在全美乃至世界范围内,明州量刑指南产生的影响最大。它存在以下几个方面的显著进步:首先,它设立了一个永久性的量刑指南委员会,由其承担量刑指南的制定与修改工作;其次,量刑指南的建立依赖于对刑事司法实践中大量生效判决以及法官量刑经验的

〔11〕 2000年,全美共有983823起重罪案件定罪,其中属于联邦的占到了59123起,仅占整体比重的6%左右;2003年,全美共有近1470045个人被定罪,属于联邦管辖的人数为173059人,仅占总数的11.8%。参见Marc L. Miller, “A Map of Sentencing and a Compass for Judges: Sentencing Information Systems, Transparency, and the Next Generation of Reform”, 105 *Columbia Law Review*, 1353(2005)。

〔12〕 参见Michael R. Gottfredson, “Parole Guidelines and the Reduction of Sentencing Disparity”, 16 *Journal of Research in Crime and Delinquency*, 205(1979)。

〔13〕 在1975年到1980年的时间里,建议量刑指南是量刑改革的主流,到了1980年,一项统计显示大多数州都开始发展或正在酝酿建议量刑体系。参见Michael Tonry, “Structuring Sentencing”, in Michael Tonry ed., *Crime and Justice*, The University of Chicago Press, 1988, p.280。

〔14〕 它们分别是:马里兰州(1983年7月)、佛罗里达州(1983年10月)、华盛顿州(1984年7月)、威斯康辛州(1985年)、特拉华州(1987年10月)、俄勒冈州(1989年11月)、田纳西州(1989年11月)、堪萨斯州(1993年7月)、阿肯色州(1994年1月)、北卡罗来纳州(1994年10月)、弗吉尼亚州(1995年1月)、俄亥俄州(1996年7月)、密苏里州(1997年3月)、马萨诸塞州(1998年2月)、犹他州(1998年10月)、密歇根州(1999年1月)、阿拉巴马州(2006年10月)、华盛顿特区(2008年)、爱荷华州(筹划阶段)、南卡罗来纳州(审议阶段)。参见Kevin R. Reitz, “The Disassembly and Reassembly of U.S. Sentencing Practices”, in Tonry et al., ed., *Supra* note 7., pp.226—227。

实证调查。除了对被告人的犯罪行为进行考察外,还重点考量被告人的罪前纪录,从而建立与犯罪以及犯罪人成比例的量刑,这一方法也被之后联邦及各州的指南奉为圭臬;再次,指南涉及假释、上诉、量刑与监狱容量的关系^[15]等许多其他方面的内容;^[16]最后,其建立了适度的上诉法院审查(Moderate Appellate Review)制度。

在明州1980年的第一版量刑指南中,“被控罪行的严重程度”(Severity Levels of Conviction Offense,下文简称“罪行程度”)共分为严重性程度依次递增的1~10级,而“刑事历史记录分数”(Criminal History Score,下文简称“历史分数”)也按照从轻到重分为0~6级(6级以上的记录被归入6级之中)。其中,属于“罪行程度”第1级的所有“历史分数”等级、属于“罪行程度”第2级的0~5级“历史分数”等级、属于“罪行程度”第3级的0~3级“历史分数”等级、属于“罪行程度”第4级的0~3级“历史分数”等级、属于“罪行程度”第5级的0~2级“历史分数”等级、属于“罪行程度”第6级的0~2级“历史分数”等级所对应的量刑均为确定刑。而在此整个由确定刑所构成的区域中,法官一般会先考虑适用非监禁刑,只有当被告人违反了缓刑或非监禁刑的管理规定的情况下,才会撤销之前的缓刑或社区矫正等非监禁刑罚,并按照指南规定的确定刑量刑。除上述区域外,其他区域在量刑指南上均具有一个量刑幅度。除此之外,委员会还明确规定了法官享有偏离指南规定量刑的权力,但是,基于其指南规定的假定合理性,只有当实质性的、紧迫性的(Substantial and Compelling)情状出现时,偏离量刑才被允许。当法官选择偏离量刑时,必须对委员会在量刑指南中明确表述的量刑目的与原则予以充分关注,并且给出为何做出偏离指南量刑的决定更合理的书面理由。同时,对于偏离量刑的理由,指南中规定:被告人的种族、性别、教育程度、雇佣历史等就不能成为偏离量刑的理由;而被害人过错、被害人由于特殊原因而处于特别易受侵犯的地位、犯罪手段残忍等,则是潜在的可以作为偏离量刑的理由。^[17]

自1980年始,明州量刑指南经过数次修改,当前有效的指南于2010年8月起生效。本版指南仅在具体内容上稍作了修改:其一,将“罪行程度”按照严重性程度依次递增的顺序分为1~11级,较第一版多出了一个级别;其二,在相邻“罪行程度”与“历史分数”的交叉格上保持“尾”与“首”相互重叠,^[18]从而给予法官更大的自由裁量权,这明显是受到联邦量刑指南的影响。

[15] 由于立法机关不会再批准建立新的监狱,因此,根据指南所导致的对被告人的行刑规模应当能够与现有的监狱容量相吻合,而不必要导致出现根据指南进行量刑的刑期超出现有监狱容量而不得不求助于安全阀措施(Safety-valve Measures),比如启动紧急释放程序。因此,在设定指南的量刑幅度时应当能够预测出由此产生的总的监禁期限,从而与现有监狱容量相符合。参见 Albert J. Reiss, "New Sentencing Policies and Practices in the USA", in Hélène Dumont ed., *Sentencing*, Les Éditions Yvon Blais Inc., 1987, pp. 140-142.

[16] 参见 Richard S. Frase, "State Sentencing Guidelines: Diversity, Consensus, and Unresolved Policy Issues", 105 *Columbia Law Review*, 1196(2005).

[17] 请参考 Minnesota Sentencing Guidelines, 1980, pp. 30, 31, 32, 38.

[18] 比如,在“罪行程度”同为第10级时,当“历史分数”等级为第2级时,指南规定的幅度为153-216月,而当“历史分数”等级为第3级时,指南规定的幅度为166-234月。

2. 联邦量刑指南

(1) 发展历程。在联邦层面,长期以来对法官量刑缺乏指引性规定,比如,在《美国法典》第18章对联邦犯罪的规定中,基本没有法定最低刑的量刑限制。针对此状况,1974年,美国国家司法研究会开启了针对建立联邦量刑指南可行性的研究计划,该研究主要由佛蒙特、丹佛以及科罗拉多州的法院承担,之后,又陆续支助丹佛、纽瓦克、新泽西、芝加哥以及菲尼克斯刑事法院进行指南工程的研究。大概在同时,国家司法研究会也进行了多司法区量刑指南项目的研究,探讨在人口及文化差异不同州中发展量刑指南的可行性。^[19]在上述诸项研究得出乐观结论后,1981年《量刑改革法》问世,其目标之一即是“建立一个针对重罪被告人的量刑体系,该体系既能够规范,同时也并不减损法官进行量刑的自由裁量权”。在充分协商后,1984年《量刑改革法》正式出台,该法提出了量刑改革的五个目标,其中最引人关注的莫过于提出通过建立联邦量刑委员会,授予该委员会审查联邦量刑程序的权力,并着手制定联邦量刑指南。委员会最终于1987年4月13日向国会提交其量刑指南草案,经国会审议后于同年11月1日生效,由此拉开了规制联邦法官量刑的帷幕。

(2) 新的创制。1987年生效的联邦量刑指南在基本运行思路,与之前的州量刑指南完全一致,其也通过“犯罪行为级别”(Offense Level,下文简称“行为级别”)以及“刑事历史记录级别”(Criminal History Category,下文简称“记录级别”)的综合来对犯罪的严重性做总的评价。但是,其仍体现出如下创新之处:

其一,犯罪行为程度在层级上更为细密。虽然就“记录级别”而言,其实质上仍只有六个级别,但在“行为级别”上,量刑表格的纵轴对应的“行为级别”共有43级,明显较之前各州的指南更为细密;其二,在法定刑幅度上,联邦指南明显更小。在联邦指南中,当应判处监禁刑时,从相对比例上,该监禁刑幅度的最高限不得超过其最低限的25%(包括本数)。从绝对数量上,该监禁刑幅度的最高限最多只能超过其最低限6个月以下(包括本数)。这种绝对数值以及比例数值的双重限制,给法官提供了较各州指南更小的操作空间;其三,联邦指南在其第三章中规定了一个单独的罪行等级调整程序。当案件被害人、被告人在犯罪中的地位,被告人是否妨碍司法以及被告人是否接受刑事责任等方面有特定因素时,在指南规定的基本“行为级别”之外,可以再进行调整;^[20]其四,严格上诉审查制度(Intensive Appellate Review),指南明确规定:当法官量刑超出指南规定的量刑幅度时,案件被上诉后,上诉法院将重点关注法官作出此种偏离量刑之理由是否恰当。^[21]这在很大程度上促使法官重视指南的规定,也与明州指南的适度上诉审查制度相区别。

(3) 合宪性危机。建立联邦量刑指南在美国从未获得一致认同,早在指南出台之前,许多司法界人士即在1986年秋季量刑委员会对指南草案举行的听证会中以及1987年4月委员会

[19] 参见 Michael Tonry, *Supra* note 13.at 279—280.

[20] 比如说,根据联邦量刑指南3A1.1的规定:“如果被告人知道或应当知道被害人由于年龄、生理或心理上的状况而异常脆弱(Vulnerable),或者被害人对该特定犯罪特别易受损时,罪行等级增加2级”

[21] 参见 Kevin R. Reitz, *Supra* note 14.at 226.

向国会提交指南的第三稿时持反对意见。在指南出台之后,从1988年中期开始,陆续出现了一系列针对指南违宪的诉讼。在1989年1月,最高法院在 *Mistretta v. United States* 案中,以8比1的裁决结果,判定量刑改革法和量刑委员会的存在是不违宪的。在量刑指南合宪性危机暂时解除10年之后,该问题又被重提,在 *Apprendi v. New Jersey* 案中,指南的合宪性受到了极大的挑战;之后的一个案件是 *Blakely v. Washington*,紧接着, *United States v. Booker* 案发生,并于2005年1月12日由联邦最高法院作出最终裁决,其可以视为美国联邦量刑指南性质归属上(从假定合理型到建议型)的转折点。^[22] 在此之前,联邦量刑指南一直延续明州量刑指南的思路,认为指南的规定是假定合理的。然而,在 *Booker* 案中,基于一旦认为这种具有假定合理性的指南条款具有强制力,就涉及到违反宪法第六修正案的理由,以布瑞耶(Breyer)为首的五位大法官认为联邦量刑指南整体上是建议性的,作为一种建议性条款,法官可以自由选择是否适用。这样一来,虽然量刑指南合宪性危机得到彻底解决,但却以放弃指南的假定合理性为代价。^[23] 由此,在 *Booker* 案之后,在地区法院一级已有法院赞成在量刑时不必考虑指南所规定的量刑幅度之限制。但是,绝大多数地区法院在量刑时,还是会首先考虑指南中规定的量刑幅度以及偏离指南的量刑是否有指南所规定的理由,而后再决定考虑美国法典第18章第3553(a)所列举的那些因素能否给予这种偏离指南规定幅度之量刑以正当性支撑。^[24] 在巡回法院一级,有六个巡回法院(第四、第五、第六、第七、第八以及第十巡回法院)认为正确依据指南所做出的量刑被当然地假定为合理,只有一个巡回法院依据指南量刑格的计算所得出的量刑不是假定为合理。^[25] 其他巡回法院则以实际行动表明了下级法院依据指南量刑的合法有效性。及至2006年,还仅只有 *United States v. Lazenby* 这一个依据指南而量刑的判决被巡回法院撤销,而且,在撤销理由上,并非因为量刑不合理,而是因为法院没有充分考虑一些能够减轻量刑的因素。^[26]

总的看来,虽然 *Booker* 案生效后联邦法官们并未就此抛弃量刑指南,但 *Booker* 案对量刑实践的影响还是不容小觑的。美国联邦量刑委员会曾对这种变化以时间与地区两个坐标为标准,揭示了 *Booker* 案之后量刑实践的种种变化。从时间上,其列出了《保护法案》(Protect Act)生效之前、《保护法案》生效期间以及 *Booker* 案之后三个时间段,相应变化有如下表所示,自 *Booker* 案

[22] 虽然在 *Booker* 案之后,又发生了另一起涉及量刑指南属性问题的 *Rita v. United States* 案,但在本案中,最高法院并没有在之前 *Booker* 案的基础上做进一步延伸,而是认为联邦巡回法院可以假定在指南之内的量刑合法有效,但是,这种假定同样承受合理性标准的评价。参见 Max M. Schanzenbach, Emerson H. Tiller, "Reviewing the Sentencing Guidelines: Judicial Politics, Empirical Evidence, and Reform", 75 *The University of Chicago Law Review*, 738(2008).

[23] 当然,虽然 *Booker* 案确定联邦量刑指南仅仅是建议性的,但是,根据布瑞耶大法官的判词,联邦法官同样必须考虑指南的幅度,继续根据指南而发现案件中相对应的事实,举行量刑听证,证明偏离的正当性与否。参见 *Ibid.*, at 738.

[24] 参见 United States Sentencing Commission, "Final Report on the Impact of United States v. Booker on Federal Sentencing", p. iv.

[25] 参见 *Ibid.*, at V.

[26] 参见 Schanzenbach et al., *Supra* note 22, at 738.

之后,联邦法官适用指南的符合率下降,偏离率相应上升。同时,在刑种严重性程度上,Booker 案之后未有显著变化,但在平均量刑长度上却出现了上升趋势。^[27]

表一

	PA 生效前	PA 生效后	Booker 案后
指南符合率	90.6%	93.7%	85.9%
向上偏离率 ^[28]	8.6%	5.5%	12.5%
向下偏离率 ^[29]	0.8%	0.8%	1.6%

从地区上,在对全美共 12 个巡回法院量刑实践的对照研究后,发现所有 12 个巡回法院在符合指南规定量刑的比重上均呈下降趋势。其中,下降幅度最小的第 11 巡回法院从《保护法案》生效后的 74.7%下降到 Booker 案发生后的 69.9%,下降了 4.8 个百分点。而下降幅度最大的第 1 巡回法院从《保护法案》生效后的 79.6%下降到 Booker 案发生后的 65.3%,下降了 14.3 个百分点。^[30]

(二)文字叙述量刑指南

1.加拿大量刑指南

在应对量刑失衡问题上,加拿大司法、行政部门自上世纪 80 年代起,即开始陷入一股量刑改革的狂潮中。加拿大量刑委员会、法律改革委员会、联邦政府、刑法改革协会、律师协会等部门都提出了自己的量刑改革方案,其中最完备者,当属 1987 年加拿大量刑委员会在其最终报告——《量刑改革:加拿大的方法》中所提倡的量刑指南体系。

加拿大量刑委员会的量刑指南体系的核心主要有如下四个部分——刑种选择、刑期选择、指南格式以及加重和减轻的量刑情节。在刑种选择上,指南试图为法官在选择监禁刑还是非监禁刑问题上提供帮助,当指南规定为“假定的入或出”(Presumptive In or Out)时,则该类案件都必须判处监禁刑种或者非监禁刑种,当指南规定为“有条件的入或出”(Qualified In or Out)时,则该类案件在大多数情况下应当判处监禁刑种或者非监禁刑种;在刑期选择上,对于指南中涉及到的犯罪,当需要判处监禁刑时,指南规定了一个相对较宽的刑期幅度,以允许法官根据犯罪的严重性程度以及被告人的犯罪记录以自由裁量刑期;在加重和减轻的量刑情节上,在量刑指南表格每个犯罪的判例部分,委员会提供了该犯罪所具有的加重与减轻情节,并且对每种情节都辅以判例,以供法官参考。同时,委员会也重申,量刑情节在指南中并未被穷尽,当特殊情况出现时,法官可以将其作为偏离指南的量刑情节。^[31]

[27] 参见 United States Sentencing Commission, *Supra* note 24.at 46.

[28] 即法官运用偏离量刑的权力,判处被告人比指南规定量刑更高的刑罚。

[29] 即法官运用偏离量刑的权力,判处被告人比指南规定量刑更低的刑罚。

[30] 参见 United States Sentencing Commission, *Supra* note 24.at 86.

[31] 参见 The Canadian Sentencing Commission, *Sentencing Reform: A Canadian Approach*, Minister of Supply and Services Canada,1987,pp.271,316—322.

在该体系中,最有特色的部分当属其指南格式,委员会建议量刑指南采取文字叙述而非数字格状式,并将所有文字精炼至一张表格之中,以便量刑法官使用。在该表格中,还给出了犯罪在实践中的量刑信息,以供法官参考。以抢劫罪为例,指南规定的格式如下^[32]:

表二

I.罪名	抢劫罪 SS.302/303《刑事法典》					
II.法定刑最高限	9 年					
III.假定刑种选择	监禁(入)					
IV.量刑指南						
构成要件等级	罪状			刑期选择		
一级抢劫罪 (加重)	使用暴力或者威胁使用暴力， 持械抢劫银行、商铺或者私人处所。			2～4 年		
二级抢劫罪 (普通)	持械抢劫没有保护的商业场所， 但并未对被害人造成实质身体伤害； 包括了抢夺钱包的行为。			4～16 月		
建议信息						
V.当前实践	资料来源	25%	50%(中值)*	75%	90% * *	
	矫正量刑计划	1 年	2 年	3.5 年	6 年	
	量刑委员会		2 年		7 年	
VI.判例	略(具体加重情节:9 个;具体减轻情节:8 个)					

* 中值量刑意味着,在对司法实践中抽样的所有抢劫罪案例中,当被告人被判处监禁刑时,一半的案例量刑在 2 年以下,一半的案例量刑在 2 年以上。

* * 90%量刑意味着,在对司法实践中抽样的所有抢劫罪案例中,当被告人被判处监禁刑时,90%的案例量刑在 6 年以下(根据矫正量刑计划的调查资料),或者 90%的案例量刑在 7 年以下(根据量刑委员会的调查资料)。

2. 英国英格兰和威尔士量刑指南

英格兰和威尔士在应对量刑失衡问题上,主要经历了四个阶段:上诉法院阶段、量刑建议小组阶段、量刑指南委员会阶段以及量刑委员会阶段,前两个阶段主要是依靠量刑判例指南来应对量刑失衡,这种方法尚属传统型模式之列。由于量刑判例指南的固有缺陷,^[33]经过量刑指南委员会阶段的过渡,2009 年,《法医与司法条例(2009)》(Coroners and Justice Act 2009)获得颁行,该条例的主要目的之一即是推进量刑的透明性与一致性,保持司法的独立性。根据该条例的第四部分,量刑委员会被最终建立,由此进入了量刑委员会阶段。

量刑委员会承担的职责任如下:制定量刑指南、公布起草量刑指南所依赖的各种资料、管理

[32] 参见 Ibid., at 319.

[33] 参见 Andrew Ashworth, Elaine Player, "Criminal Justice Act 2003: The Sentencing Provisions", 68 *Modern Law Review*, 827(2005).

量刑指南的运行、准备制定新指南的各种资料、推进量刑以及量刑实践的公开化、公布与量刑实践等相关事项实际效果的年报。在上述职责中,制定量刑指南是重中之重。与美国联邦量刑委员会不同,该委员会并不是将所有犯罪统一规定在一部单独的量刑指南中,而是主要以具体某一类或某一种犯罪为标准,分别建立单独的量刑指南,到目前为止,该委员会已经公布了共 19 部单独的量刑指南。以《谋杀罪》的指南为例,首先,根据《刑事行为条例(1981)》的规定,该罪最高刑为无期徒刑,指南对此予以列明;其次,依据严重性程度不同,本罪被细化为三种罪状,每种罪状下根据犯罪结果又分为三种情形,这样,本罪就一共有 9 个起刑点与 9 个量刑幅度。以本罪最重的罪状等级 1 为例,指南根据犯罪结果又分出了三组,第一组为“严重和长期的生理与心理损害”,起点刑为 30 年监禁,量刑幅度为 27~35 年的监禁;第二组为“一定程度的生理与心理损害”,起点刑为 20 年监禁,量刑幅度为 17~25 年的监禁;第三组为“很小程度或没有造成生理与心理损害”,起点刑为 15 年监禁,量刑幅度为 12~20 年的监禁。在起刑点与量刑幅度之外,还专门就本罪规定了 5 个加重量刑情节以及 4 个减轻量刑情节,由此对本罪量刑进行了全面的指引。

3. 韩国量刑指南

2006 年 12 月,韩国国民大会通过调整《法院组织条例》,从而同意建立量刑委员会,韩国量刑委员会在最高法院内部得以建立。通过一系列努力,最终于 2009 年 7 月 1 日正式开始实施第一部量刑指南,该指南涉及到故意杀人罪、强奸罪、受贿罪、伪证与诬告罪、挪用与侵占罪以及抢劫罪等几个最常见犯罪。^[34]

根据韩国量刑指南,法官主要根据指南上对犯罪规定的犯罪类型、量刑情节等级选择合适的量刑幅度,以作出量刑结论。具体而言,在犯罪类型上,指南根据严重性程度的不同,将每个犯罪都分成三种类型;而在量刑情节等级上,量刑情节则被分为具体与普通两类,^[35]在对量刑所起作用力上,具体量刑情节要高于普通量刑情节。同时,该两类量刑情节又被细分为基本、加重与减轻三个等级,分别对应着基本、加重与减轻三个量刑幅度。法官在量刑时,当认为加重情节的影响力更大,则可以根据加重幅度量刑,相反,则根据减轻幅度量刑,而如果认为两者的影响力相似时,则根据基本幅度量刑。除此之外,还有一个特别规定,当在某个案件中具体加重情节的数量要比具体减轻情节的数量多两个或以上时,则可以提高加重幅度上限的 50% 作为最终的量刑上限,相反,如果某个案件中具体减轻情节的数量要比具体加重情节的数量多两个或以上时,则可以降低减轻幅度下限的 50% 作为最终的量刑下限。^[36]以谋杀罪为例,法官量刑时所接受的指南指引如下表所示:

[34] 参见 Hyungkwan Park, “The Basic Features of the First Korean Sentencing Guidelines”, 22 *Federal Sentencing Reporter*, 262, 265 (2010).

[35] 对于何谓具体加重与减轻、普通加重与减轻量刑情节,指南进行了详细的规定。比如,对于谋杀罪,与犯罪行为相关的具体减轻量刑情节有:超过正常限度的防卫行为、非常弱的犯罪意图、被害人的强烈挑衅;而与犯罪人相关的普通减轻量刑情节则有:合适的赔偿、心理疾病(对犯罪有过失)、真诚悔罪。

[36] 参见 Hyungkwan Park, *Supra* note 34, pp. 265, 267—268.

表三

犯罪类型	减轻幅度	基本幅度	加重幅度
第一类型	3~5 年	4~6 年	5~7 年
第二类型	6~9 年	8~11 年	10~13 年
第三类型	8~11 年	10~13 年	12~15 年,无期徒刑,死刑

4. 新西兰量刑指南

新西兰量刑指南建立的法律依据为《新西兰量刑委员会条例(2007)》,其在新西兰法律委员会与司法界代表团队的合作下得以问世。在指南的效力问题上,根据上述《条例》第 27 条 A 的规定,制定的量刑指南应当具有约束力,只有在法官认为适用指南的条款量刑将违背公正时,才可以不予适用,而且此种偏离指南的量刑必须给出充分的理由,并且在整体上偏离量刑的案件不应超过 20%;在指南的表达形式上,新西兰量刑以英格兰和威尔士的量刑指南为蓝本,针对不同犯罪、不同情况分别制定了共 60 个单独的量刑指南文件;在指南的具体内容上,其主要可分为两类,一类是专门针对特定犯罪或犯罪不同罪状的指南,被称为“罪行指南”,另一类是那些能够适用于所有犯罪的指南,被称为“一般指南”。就罪行指南而言,对每一种具体犯罪,都根据犯罪严重性程度分为三个层次,并分别规定了不同的量刑幅度,法官根据具体犯罪的加重或减轻情节,在相应的幅度内量刑。就一般指南而言,有专门针对有罪答辩的指南,规定被告人存在有罪答辩时可以降低的量刑幅度。

二、信息型克服模式

信息型克服模式通过收集以往法官量刑的案例,根据一套案例的分解机制,将那些影响法官量刑的情节设定为一定的变量,通过电子计算机技术,使得法官能够在该系统中搜索到与待量刑案件最相似的生效判例,然后通过查看这些生效判例的量刑分布状况,以及所有判例的具体案情描述,最终得出最适合待量刑案件量刑结论。就目前来看,正在实施或曾经实施过量刑信息系统的代表主要有:加拿大四省(英属哥伦比亚、马尼托巴、萨斯喀彻温、纽芬兰)、澳大利亚三州(新南威尔士州、维多利亚州、昆士兰州)以及英国苏格兰地区。

(一) 加拿大四省量刑信息系统

加拿大是最早尝试以量刑信息系统方式来应对量刑失衡的国家。上世纪 80 年代中期,加拿大学者约翰·霍噶斯(John Hogarth)在英属哥伦比亚省进行了一次被称为电子化数据量刑系统的尝试,但受限于当时电子计算机硬件技术的落后,该尝试归于失败。在此之后,于上世纪 80 年代末期,在加拿大学者安东尼·杜布(Anthony Doob)和帕克(Park)的倡议下,加拿大四省(英属哥伦比亚、马尼托巴、萨斯喀彻温、纽芬兰)正式启动为期六年的量刑信息系统实验项目,其中,以英属哥伦比亚省的量刑信息系统最为突出。

英属哥伦比亚的量刑信息系统从 1987 年开始,到 1992 年结束。这个量刑信息系统

包括了如下5个组成部分:①来自于英属哥伦比亚省的省、郡、高级法院的120000份判决书中的量刑数据,这可以使法官通过输入某些罪行或被告人的特质而查看相对应的量刑分配;②来自英属哥伦比亚上诉法院的包括了1110份报道决定(Reported Decisions)的数据库;③获得英属哥伦比亚上诉法院认可的关于量刑加重或减轻情节的清单,同时附上得出这些加重或减轻量刑情节案件的全文作为参考资料;④由学者约翰·霍噶斯汇编的由量刑判例法与实体法组成的量刑法律数据库;⑤英属哥伦比亚省内能够为量刑信息提供帮助的机构名目。^[37]

(二)澳大利亚新南威尔士州量刑信息系统

在澳大利亚,新南威尔士州、维多利亚州以及昆士兰州均试图通过制定量刑信息系统来克服其量刑失衡现象,其中尤以新南威尔士州的实践最为突出。1986年,根据新南威尔士的《司法官员条例》,其建立了司法委员会,由其负责量刑信息系统的组建工作,1996年,该系统得以竣工。

量刑信息系统由以下部分组成:裁判文书数据库、案件总结数据库、规范数据库、数据资料数据库、新南威尔士以及联邦立法数据库、机构数据库、精选案例数据库、出版物数据库。在裁判文书数据库中,法官通过输入相关字段,如输入“疏忽杀人罪”,就能够得出与此相关的所有案例;案件总结数据库包括了对案例的一些简单描述,比如起诉罪名、辩诉交易情况、被告人的背景、最终量刑结论等,法官据此能够更加便捷的发现相类似的以往案例;规范数据库对所有量刑立法以及量刑原则加以总结,从而给法官量刑时提供及时的规范指导;机构数据库提供这样的服务,当法院对被告人判处非监禁刑时,特别是考虑到矫正被告人目的时,法官经常面对选择特定的机构来帮助被告人更好的社会化。因此,机构数据库就提供一些相关机构的信息,比如毒品麻醉品辅导办公室,赌博、假释办公室,临时监管办公室,少年司法办公室等;精选案例数据库中的所有案例都来自高等法院,高等法院的裁决对全澳大利亚具有强制约束力;出版物数据库包括了新南威尔士量刑委员会的研究部门定期公布的《量刑趋势》以及《司法办公室公告》的内容,以供法官量刑时查阅。^[38]

当然,在所有数据库中,最为重要的还是数据资料数据库。该数据库包括了更为详细的量刑实践数据,其来源的法院主要有三类:治安法院、上级法院(包括了地区法院和高级法院)以及未成年法院。本数据库中的量刑数据由新南威尔士总检察长之下的犯罪统计与研究局收集并电子化,再交给委员会使用。到1998年,数据资料数据库已经有了自1992年11月起的400000个来自治安法院的案件量刑资料以及自1990年12月起的20000个来自上级法院的量刑资料。如果量刑法官想知道上级法院故意伤害罪的量刑实践时,就可以从三类法院类型列表中选择上级法院类别,再选择故意伤害罪罪名,最终,相关的量刑信息就会以柱状图的形式

[37] Neil Hutton, Cyrus Tata, John N. Wilson, "Sentencing and Information Technology: Incidental Reform", 2 *International Journal of Law and Information Technology*, 259(1994).

[38] 参见 Ivan Potas, David Ash, Murali Sagi, Stephen Cumines, Natalie Marsic, "Informing the Discretion: The Sentencing Information System of the Judicial Commission of New South Wales", 6 *International Journal of Law and Information Technology*, 107-110(1998).

式显示出来。横轴代表该类犯罪被量处的不同类型的刑罚种类,纵轴代表该种刑罚种类所具有的案例个数,在代表每种刑罚种类的柱状体上方,还标示出了该种刑种在整体中所占的比重。如果使用者还想进一步了解每种刑种具体的刑期(数量)分布,只要点击代表该刑种的柱状体,电脑屏幕上就显示出相关信息。^{〔39〕}

(三)英国苏格兰量刑信息系统

苏格兰量刑信息系统虽然倍享盛誉,但其建立的初衷并非是为了克服量刑失衡,而是应对政治压力。^{〔40〕}“在经历从1993年到2002年断断续续的研究后,最终于2002年决定将量刑信息系统运用至高等法院的量刑之中,同时,数据的收集整理输入的工作由法院的书记员接手。到2002年,该系统已经拥有了自15年以来的共15000个案例的详细量刑资料(包括这些案件上诉后的量刑资料)”。^{〔41〕}“苏格兰的量刑信息系统主要由两个子系统组成:一个被称为‘数据获取子系统’,通过这个系统能够获取到以往量刑的相关信息;另一个被称为‘数据输入子系统’,通过这个系统,能够将与量刑相关的信息输入系统中以备之后所需”。^{〔42〕}

使用“数据获取子系统”获得需要的量刑信息可以有两种方法,一种称为“修正方法”,另一种称为“主要罪行方法”。现假设这样一个案例:一位男性被告人挥舞着匕首在一个商店进行抢劫,犯罪没有造成任何伤亡,该被告人没有任何前科。当法官想要通过用“修正方法”查询与本案相似案件的量刑时,在“数据获取子系统”的界面中,法官最先在“犯罪分类”的选区内选择抢劫罪,同时,由于本案件中的抢劫罪具体罪状属于“加重型抢劫”的情形,因此,在“罪状等级”的选区内选择加重型抢劫,最后,点击“修正方法”的按钮,从而进入“修正方法”的操作界面。此时,如果法官仅仅只是想了解加重型抢劫的量刑情况,就可以在出现的新界面中点击“查看量刑图表”,程序就会自动显示以往法官对加重型抢劫的所有量刑数据,当然,这些量刑不是相同或相似的,因为即使同为加重型抢劫,也会在具体行为方式、犯罪结果、被告人特质上存在极大的差异,从而量刑当然会有高低不同。因此,对于大多数法官而言,还需要继续细化犯罪,以寻求精致性程度更高的同案,从而,在选择“修正方法”的操作界面并进入新界面时,就不会立即点击“查看量刑图表”按钮,而是在犯罪前科、辩诉交易、被告人性别、被告人年龄、犯罪行为类型、犯罪结果等选区内选择与本案最适合的选项,同时,如果法官不清楚某些情节,比如,法官不清楚被告人年龄究竟为何,则在被告人年龄选区内可以选择“忽略”选项。当对所有选区内的选项选择完毕后,此时再点击“查看量刑图表”按钮,就进入另一个新界面,在界面的左上方有一个数字,表示与待量刑案件相似的已量刑案件的数量,再下方就显示出与待量刑案件最相似的案件的量刑情况表。该表实际上是一个柱状图,横轴表示的是各种刑种,而每种刑

〔39〕 参见 Ibid., pp.111—113。

〔40〕 上世纪90年代初期,苏格兰的保守党大臣们基于对量刑实践的不满,提出引入包含了法定最低刑的“两击模式”的法律,试图以此限制法官的自由裁量权。在此种政治压力下,苏格兰法院系统开始酝酿量刑领域的变革,并最终选择建立量刑信息系统作为回应。参见 Neil Hutton, Cyrus Tata, “A Sentencing Exception? Changing Sentencing Policy in Scotland”, 22 *Federal Sentencing Reporter*, 275(2010)。

〔41〕 Ibid., at 274.

〔42〕 Neil Hutton 等, *Supra* note 37, at 268.

种的柱状体颜色各不相同,在整个图表的右边也会以各种颜色的小方块各自对应着各种刑种,纵轴表示每种刑种的出现频率。如果法官想查看其中某一种刑种的刑期排列情况,则只需要点击某种刑种的柱状体,就会出现另一个新的界面,上面显示该刑种刑期的排列情况。比如,对于前述的加重型抢劫案,如果在第一个柱状图界面中,法官发现在以往量刑中,大约有 60% 的案件都是量处的监禁刑,因此,认为监禁刑是对此种案件的合适量刑,就可以继续点击代表着监禁刑的柱状体,进入新界面,在该界面中,同样也会出现一个柱状图,横轴代表刑期的具体数额(以年为计量单位),纵轴代表每种刑期的出现频率,比如,对于前述的加重型抢劫案,法官会发现在以往量刑中,主要的量刑情况为:判处 2 年以下的有约 2%,判处 2~4 年的有约 8%,判处 4~6 年的有约 35%,判处 6~8 年的有约 12%,判处 8~10 年的有约 1.5%,判处 10~12 年的有约 1.5%。不仅是简单的量刑情况,当法官希望具体查看某个案件更细节的情况时,还可以选择进入“个案具体情况表格”,从而可以查询到由高级法院存档的某个案件的全部审判文书,从而更细致地了解某案件最终量刑的具体理由。“主要罪行方法”的使用流程与“修正方法”几乎完全相同,唯一不同之处在于,当法官认为某个情节对量刑的作用不大,或者认为如果选择某些量刑情节,最终得出的同案数量将非常少或根本没有时,法官就可以在具体的某个量刑情节的选区内选择“忽视”选项,排除该特定量刑情节,从而查看不完全相同的准“同案”的量刑情况。^[43]

而通过‘数据输入子系统’,相关人员可以将一个生效的刑事案件按照系统中提示的变量,将该案件进行降解并逐项输入量刑信息系统中,其会自动保存,在以后法官使用“数据获取子系统”时,该新输入的量刑信息将成为整体量刑信息的一部分,供法官量刑时参考。

三、模式之实效评价

前文对指南型与信息型两种量刑失衡克服模式进行了介绍,但两种模式的优劣如何,仅进行形而上的探讨难以作出公正的判断,还需要对各自在司法实践中的运作实效予以描述。

(一)指南型克服模式之实效

在实施数字格状量刑指南的美国,从联邦到各州,其量刑委员会都会定期对量刑状况进行评估。在联邦的报告中,法官量刑处于指南规定幅度内的比重(下文简称合规性比重)一直在 80% 以上,即使在 Booker 案之后,该比重亦没有明显下降;在明州的报告中,自其第一版指南实施以来,法官量刑的合规性比重就保持在至少 80% 以上,在 1981 到 1991 年,该比重保持在 90% 以上,1991 年以后,该比重稍有回落,在近 6 年内,该比重维持在 87% 左右,在 2009 年度,法官量刑的合规性比重为 87.3%;^[44]在宾夕法尼亚州的报告中,从 1985 年起,法官量刑的合

[43] 参见 Ibid., pp.269—278.

[44] 参见 Minnesota Sentencing Guidelines Commission, “Sentencing Practices: Annual Summary Statistics for Felony Offenders Sentenced in 2009”, p.69.

规性比重保持在85%以上,从2005年开始该比重一直在90%以上,2009年度该比重为90%。^{〔45〕}不仅在联邦和上述两州中,即使是在实施纯粹建议量刑指南的弗吉尼亚州,法官量刑的合规性比重仍然非常高,“在1995年到1998年,法官量刑的合规性比重大约在75%左右,在1999年到2001年稳步提升,到2002年又略有下降。在2009年之前的6年内,法官量刑的合规性比重保持在80%以上。在2009年度,有79.9%的案件量刑符合量刑指南的规定”。^{〔46〕}由于量刑指南的制定过程遵循统一的量刑目的,因此,遵循量刑指南的法官在量刑上基本保持相同的量刑尺度,^{〔47〕}可以说,实践量刑符合量刑指南的规定,就大体上可以认定该量刑保持了量刑的平衡性,因此,美国以数字格状量刑指南为载体的量刑改革实践证明:量刑指南是现代国家克服量刑失衡的一种重要手段,量刑指南的问世,极大地降低了美国刑事司法实践中量刑失衡的程度。

在实施文字叙述量刑指南形式的国家或地区,其中:尽管加拿大量刑委员会的量刑指南体系非常完备,且备受推崇,但其总归只停留在建议阶段,未被当局采纳并推行,其实效无从判断;虽然《新西兰量刑委员会条例》已付诸实施,并且相关的指南草案也基本完成,但是,2008年11月上台的新政府认为推行量刑指南时机尚不成熟,其实效亦无从谈起。就真正开始实践量刑指南的英国英格兰和威尔士地区而言,在一项由量刑委员会所公布的研究报告中,通过对治安法官、地区法官、司法建议官(Legal Advisors)的实证调查,发现:绝大多数被调查者均认可由于量刑指南明确规定了刑罚幅度,因此有利于实现同案同罚,总之,从英格兰和威尔士司法实践所反映的情况来看,其量刑指南极大地统一了法官的量刑尺度,对于克服量刑失衡提供了强劲助力;韩国对其指南的实践进行了细致的评估,“根据韩国最高法院行政办公室的报告,从2009年7月1日到本月底,共有2920个与指南规定犯罪相关的案件,其中法官按照指南进行量刑的比例为89%,量刑失衡的案件大幅度降低。”^{〔48〕}基于对第一部量刑指南取得效果的乐观评估,韩国已经开始着手进一步的量刑指南计划,在第二部量刑指南中,将涉及到诈骗罪、盗窃罪、伪造罪、绑架罪、毒品犯罪、妨碍公务罪等犯罪。^{〔49〕}

(二)信息型克服模式之实效

信息型量刑失衡克服模式立足于为法官提供量刑信息,其不像量刑指南一样,试图提供一个被限定好的“量刑结论”,其不会被过多地诠释成限制法官自由裁量权的制度。同时,尽管其不具有强制性,但正像英国学者姆巴赫(Murbach)与依恩(Nonn)所说,“量刑信息系统具有一定程度的间接影响。当法官看到量刑信息系统所展示出来的相对客观的量刑差异的证据,而不是像以前一样仅仅是从一些道听途说的途径得来的量刑差异的证据,法官将会切实形成自

〔45〕 参见 Pennsylvania Commission on Sentencing, “Annual Reports 2009”, p.41.

〔46〕 参见 Virginia Criminal Sentencing Commission, “Annual Reports 2009”, p.18.

〔47〕 值得注意的是,并非所有偏离量刑的案件就一定属于量刑失衡的范畴,不能否定在比重不大的偏离量刑案件中,其中绝大部分都是有合理的偏离量刑理由,因此,偏离量刑结论并无不当。

〔48〕 Hyungkwan Park, *Supra* note 34, at 269.

〔49〕 参见 Ibid., at 269.

己的量刑方法存在问题的印象,从而适当的修正自己的量刑模式”,^[50]因此,量刑信息系统正在成为解决量刑失衡的新出路。

就加拿大四省的量刑信息系统而言,其为期六年的量刑信息系统计划并未如约实施下去,有的省完全停滞,有的省虽未停止,但也只是使用该系统为上诉法院提供量刑信息,从而基本宣告失败。对此结局,其首倡者安东尼·杜布和帕克认为:“其一,加拿大的法官对目前实践中的量刑信息不感兴趣。他们既不习惯于以当前的这种数据化的方式使用量刑信息,法律传统上也没有对当前的量刑实践予以关注;其二,对量刑的限制主要来自于上诉法院的审查,量刑信息系统没有任何的强制力”;^[51]就澳大利亚新南威尔士州的量刑信息系统而言,“在司法委员会自己看来,其在出台后受到了使用者的好评,其用户也呈稳定增长趋势”,^[52]新南威尔士的法律改革委员会在其官方评价中,也给予了高度肯定,认为“量刑信息系统对法官在现有案件的量刑过程中是一个非常有价值的辅助工具”。^[53]而受到新南威尔士州量刑信息系统的启发,澳大利亚法律改革委员会也提出建议:“为了推进联邦犯罪量刑的一致性,澳大利亚法律改革委员会建议各司法区发展一套量刑信息系统,为量刑法官提供全面的、最新的以往犯罪的量刑信息,包括对犯罪适用的刑罚种类、刑罚期限、与犯罪行为 and 犯罪行为人相关的量刑因素以及各种因素对量刑的影响程度。”^[54]虽然官方给予一致好评,但从其运用率来看,结果并不乐观。“在1994年到1995年实施系统一年半的时间里,系统中的量刑法律系统访问次数每月在1000次以下……而对量刑数据系统的访问次数则每月不足200次”。^[55]因此,实际上该系统的使用率并不高,其对量刑失衡的克服作用显然是微弱的;就苏格兰的量刑信息系统而言,由于其对法官运用量刑自由裁量权不做任何限制,因而倍受法官青睐,^[56]“通过对运用了量刑信息系统量刑法官的调查,结果显示,苏格兰法官认为此套系统非常实用”。^[57]在苏格兰1994年的政府白皮书《稳定与公平:提高苏格兰的司法公正》中,政府也明确表示“欢迎量刑信息系统的使用,并且指出如果这套系统能够成功运行的话,类似的系统也可以考虑在下级法院启用。”^[58]虽然表面形势一片大好,但前文已提及,苏格兰量刑信息系统建立的首要目的在于缓解政治上对司法界的压力,当其被建立,政治上限

[50] Neil Hutton 等, *Supra* note 37.at 283.

[51] *Ibid.*, at 259.

[52] *Ibid.*, at 260.

[53] New South Wales Law Reform Commission, “Sentencing(Discussing paper 33) 1996”, p.248.

[54] Australian Law Reform Commission, “*Sentencing of Federal Offenders (Issues Paper 29)* 2005”, p.29.

[55] Austin Lovegrove, “Statistical Information Systems as a Means to Consistency and Rationality in Sentencing”, 7 *International Journal of Law and Information Technology*, 34(1999).

[56] 参见 Neil Hutton 等, *Supra* note 37.at 282.

[57] Julian V.Roberts, “Sentencing Research in Canada”, 41 *Canadian Journal of Criminology*, 230 (1999).

[58] Neil Hutton 等, *Supra* note 37.at 282.

制司法自由裁量权的呼声渐弱后,正如有学者已指出的:“在没有急迫的政治压力下,司法界高层开始对系统置若罔闻”,^[59]这也给该系统未来的发展蒙上阴影,其对量刑失衡的克服作用也可能在此种阴影下被逐渐蚕食。

四、中国的选择

(一)以指南型模式为基的渐进式进路

中国的量刑失衡克服模式显然更接近于传统型模式,然而,传统型模式下的诸种制度,在目前中国的刑事量刑体系中却并未被较好的贯彻,判决书不说理、量刑基准上各行其是,导致同案不同刑的情况屡见不鲜。^[60]由此可见,对中国量刑体制进行变革势在必行。而在变革的进路上,究竟采取一种怎样的失衡克服模式,这是待解决的首要问题。正如本文开头所言,即使在其他具有较好量刑传统与制度规范的大陆法系国家中,传统型模式克制量刑失衡的无能性也在日益暴露,因此,中国如果要继续目前的传统型模式,即使在具体制度上做足功夫,也难免会步入死胡同。因此,笔者赞同中国走上全局性的量刑变革道路,而这就意味着,我们的待选项将是指南型与信息型两种模式。

就信息型模式而言,它的精髓在于量刑的先例。美国联邦最高院卡多佐大法官曾在谈论先例时轻描淡写地说到:“法官(在审判中)第一步就是考察和比较先例。如果先例清楚明了并且契合案件,那么法官就无需做更多的事了”,^[61]由此可见先例在英美法系的显要地位。在中国理论界,早就存在对先例的关注,有学者甚至提出了建立先例制度的具体举措,^[62]在实务界,早在2002年8月,郑州市中原区人民法院经过一年的试行,正式推出了“先例判决制度”,虽然这一制度更直接针对的是定罪活动,但对于量刑的间接作用也是存在的。而江苏省姜堰市人民法院在其《量刑规范化指导意见》中,则直接针对量刑提出了“先例参照”制度,强调了先例判决对同案的量刑指导作用。基于理论与实务界的双头并进,2010年7月和11月,最高检和最高法先后出台了各自的《关于案例指导工作的规定》(下文简称“规定”),由此,案例指导制度在中国得以初步建立。这是否可以视作信息型模式在中国本土化的一种表现形式呢?笔者对此持否定态度,它们的区别主要体现在:

首先,两者在运作理念上存在本质差别。在量刑领域内,根据最高院的上述“规定”:各级法院审判类似案件时应当参照指导性案例,由此,指导案例这一先例的量刑对于法官具有强制力;与此不同的是,“量刑信息系统的运作在本质上具有自愿性。法官并不被强迫着去使用,其

[59] Neil Hutton 等, *Supra* note 40, at 275.

[60] 笔者曾进行过一次量刑实验,将一个故意伤害的案例交由全国42个法院共209名法官量刑,实验结果发现:全部法官对完全相同案件的量刑存在30%左右的失衡情况。参见拙文:“克服量刑失衡二元体系之构建——基于对我国量刑失衡现象的实证分析”,《政治与法律》2013年第11期。

[61] (美)本杰明·卡多佐:《司法过程的性质》,苏力译,商务印书馆2000年版,页8。

[62] 参见邓修明:“我国刑罚裁量模式与刑事判例机制”,《现代法学》2006年第1期。

也并未被强迫根据系统展示出来的量刑分配而选择一种具有普遍性的量刑结论”,^[63]量刑信息系统中的先例并不会对法官的量刑“指手画脚”,运作该系统的结果并非某一确定的量刑结论,而是向操作者展示司法实践中其他法官在面对类似案件中的量刑经验,从而能够让不同地区、不同时间的法官通过系统这一中介而相互“沟通”,以逐步实现量刑平衡。

其次,两者在运作规模上存在显著差异。案例指导制度是以个案对整体司法实践的指导为表现形式,而量刑信息系统则表现为司法实践整体对个案量刑的引导。由此可见,在制度设计上,前者显然只包括了数量较少的经典案例,而后者则包括了大样本的生效量刑先例。从实践运作情况来看,两者规模差异更加明显,仅以英国的苏格兰地区为例,其在不到10年的时间里,量刑信息系统的案例容量即已达到15000之多。以此类推,中国量刑信息系统的案例容量将是巨大的,“国家级的模范判例库应当具有相当规模。以我国刑法四百多个罪名、每年几十万个刑事案件来说,每个常见犯罪至少一万个模范判例的库容,应该不是过分的要求”;^[64]而反观指导案例的构建,从2010年两高发布“规定”开始,最高检只公布了三个案例,最高法则尚未公布案例。两种制度在运作规模上的显著差异,由此可见一斑。

最后,两者在运作方法上存在繁易不同。案例指导制度存在使用繁琐的缺陷,对此,加拿大量刑委员会早已认识到:“许多量刑法官需要了解的关于上诉法院所作出的量刑信息都存在于一些公开出版的报告中,但问题是,法官工作太忙而根本无法花费几个小时的时间去寻找,进而理解这一些公开出版资料上的内容”。^[65]对此,我国有学者亦深有体会;^[66]而相比案例指导制度,量刑信息系统则在使用上更为易捷,其随着电子计算机的普及而异军突起,成为传统型及指南型模式之外的第三选择,在使用电脑已经成为普通公众基本生活方式的现今时代,要求作为社会精英群体的法官们学会使用量刑信息系统,通过轻点鼠标以获得其他法官类似情况下的量刑经验,完全是不过分的。

总之,从以上三点可以看出,案例指导制度并非量刑信息系统的中国化表现形式,这也就意味着,即便在实施案例指导制度的当下,如果决心选择信息型模式,我们仍然还有从零开始构建量刑信息系统的必要。当然,这并不意味着在现阶段中国应当立即开始步入信息型模式的发展进路,在笔者看来,我国当下并不存在适合信息型模式生存的司法环境,原因在于:

第一,估堆量刑法使得量刑信息系统的构建成为不可能。信息系统的原理即在于将整个案件事实进行降解拆分,对于每个部分对最终量刑结论所起的作用力都一一标明,而在估堆量刑法下,法官只知道案件的最终量刑,但对于得出量刑结论的案件基础事实以及各量刑情节各自对量刑所起的具体作用力,不需要法官说清,实际上也无法说清,据此得出的量刑先例,根本

[63] Cyrus Tata, Neil Hutton, “Beyond the Technology of Quick Fixes: Will the Judiciary Act to Protect Itself and Shore Up Judicial Independence? Recent Experience from Scotland”, 16 *Federal Sentencing Reporter*, 68(2003).

[64] 白建军:“裸刑均值的意义”,《法学研究》2010年第6期。

[65] The Canadian Sentencing Commission, *Supra* note 31, at 62.

[66] 参见舒洪水:“建立我国案例指导制度的困境和出路——以刑事案例为例”,《法学杂志》2012年第1期。

无法按照量刑信息系统的运作机理拆解输入数据库内,即使笼统入库,之后供法官参考的也只能是案件的整体量刑,而这样的量刑信息系统在作用上将大打折扣。

第二,量刑隐性自变量的作用过大,导致不合理量刑经验上升为量刑先例,其参考价值自然受损。判决书量刑不说理已经成为中国刑事司法实践的潜规则,这种“不讲理”不仅表现在对于判决书上写明的影响量刑的事实和情节仅作简单罗列现象之上,更表现在基于不能明说的理由而在判决书上对那些事实上影响法官作出量刑结论的事实和情节加以省略的现象之上。对于这些列明的事实与情节,笔者称之为量刑的显性自变量,而对于这些未列明的事实与情节,则可称之为量刑隐性自变量。笔者曾在一次实证研究中发现,在数据库判决书列明的显性自变量对量刑结论这一因变量的 R^2 为 54%,这也就意味着法官在判决书上列明的这些影响量刑作出的事实与情节只在 54% 的程度上解释了量刑结论的变化,而剩下的对量刑结论起 46% 的作用力则源自隐性自变量。^{〔67〕} 量刑隐性自变量在司法实践中对于法官量刑所起作用过大,一旦将这样的案例群纳入系统中,据此所得出的量刑经验显然是不合理的,参考不合理的量刑经验只能导致量刑不合理的延续而非断绝。

第三,虽然并不能将此绝对化,但缺乏判例法传统必将在一定程度上影响量刑信息系统的使用。量刑信息系统是以先例作为构建基础的,而在深受成文法影响而缺乏先例认同的中国司法实践,量刑信息系统的引入恐怕无法一帆风顺。通过前文对信息型模式的实效评价可知,在实施该模式的法域下(多具有判例法传统),对于量刑信息系统的使用率仅维持在较低的水平上,如何提高系统的使用率,成为信息型模式采纳者的共同难题,而在案例指导制度这一判例法制度刚刚起步的中国,目前还远不到启动信息型模式的时候。

由此可见,在当下中国的现实司法环境下,信息型模式的构建确实操之过急,在作别传统型模式之时,我们应当从选择指南型模式开始中国的量刑克服进路,这样一条进路,在笔者看来,应当具有渐进性的特征。

渐进式进路的第一步,是选择一条单维的指南型模式进路,建立中国的量刑指南体系。然而,单维的指南型模式并非理想量刑体系的终点。马克思·韦伯从形式与实质(Formal/Substantive)、合理与不合理(Rational/Irrational)这两条轴展开对法律决定的分类,从而,法律决定可以分成四种类型:形式合理性、实质合理性、形式不合理性、实质不合理性。而在现代法治国家中,运用韦伯的上述法律决定类型理论来选择量刑决定类型,显然,形式与实质不合理的量刑决定类型是不恰当的,而只有形式合理性与实质合理性才能契合现代法治国家的需求。就形式合理性的量刑决定类型而言,其通过事先建立起明确的量刑规范,将法官的量刑过程统一在量刑规范辐射范围之内,采取统一的、分步骤的量刑方法来实现量刑目的,据此我们将倾向于建立指南型量刑失衡克服模式;就实质合理性的量刑决定类型而言,法官在公正理念的统摄下,依据良心确信准则,运用其认为可靠的量刑方法,最终实现量刑目的,信息型量刑失衡克

〔67〕 参见拙文:“论犯罪过失类型与刑罚轻重的关联”,《中国刑事法杂志》2009年第4期。无独有偶,白建军教授也曾发现最高院数千个示范性案例的 R^2 值仅为 48.8%。参见白建军:“论法的确定性与公正的可检验性”,《中国法学》2008年第2期。

服模式与此更接近。实际上,指南型与信息型两种模式也各有其短,指南型过于程式化,其虽然能够指导司法实践中绝大多数典型案件之量刑,但在少数非典型的极端个案上,很难或无法实现量刑公正;而信息型则过于动态化,其虽然在少数非典型的极端个案上能够更大程度的实现量刑公正,但对于普遍性的典型案件,则由于不具有外在规则制约而显得标准不一。要克服两者之短,指南型与信息型的合作是关键,由此需要迈进渐进式进路的第二步。

渐进式进路的第二步,是当指南型模式运转良好,当信息型模式不适应中国状况的上述问题都已通过指南型模式与案例指导制度的贯彻而被消解后,最终,我们可以适时地推出中国的量刑信息系统,化解适用指南型模式而无法解决的种种问题。〔68〕

(二)指南型模式的可行性

建立起中国的量刑指南体系显然是渐进式进路的关键。然而,虽然通过量刑指南克制量刑失衡现象已经在其他国家或地区获得成功,但“橘逾淮为枳”,指南型模式在中国同样会产生积极效应吗?量刑指南在中国的可行性问题,将是下文所关注的重点。

1. 最高院规范化改革的三个阶段

面对中国量刑失衡现象,中国最高司法机关显然是有所作为的。最高院早在2005年10月《人民法院第二个五年改革纲要》中,即提出了要改革和完善人民法院量刑工作机制,进一步规范量刑活动,并具体到要研究并制定人民法院量刑指导意见,由此,拉开了最高院量刑规范化改革的序幕。总体来看,其可以分为三个阶段:第一阶段,2008年8月,最高院在前期调研的基础上,将形成的与量刑相关的实体与程序上的两个量刑规范化文件在12个法院展开试点工作(4个中级法院、8个基层法院);在取得积极效果之后,2009年6月1日,最高院开始扩大规范化改革试点法院的范围,将包括了交通肇事、故意伤害、抢劫、盗窃以及毒品犯罪这五个常见犯罪在有期徒刑范围内的量刑指南文件《人民法院量刑指导意见(试行)》(下文简称“指导意见”)及其配套的程序文件《人民法院量刑程序指导意见(试行)》交由120余个法院试点,由此进入了第二个阶段;2010年10月1日,在前两次小规模范围的试点工作取得积极反响之后,最高院正式向社会公布量刑规范化改革的实体文件:《人民法院量刑指导意见(试行)》(下文简称“量刑指导意见”)以及程序文件《关于规范量刑程序若干问题的意见(试行)》,前者在之前五类犯罪的基础上,增加了强奸、诈骗、寻衅滋事、抢夺、窝藏赃物、敲诈勒索、非法拘禁、妨害公务、职务侵占、聚众斗殴罪10个新罪名。同时,还将之前的有期徒刑扩大为拘役及有期徒刑,此为最高院量刑规范化改革的第三阶段。

2. 最高院规范化改革之实证评估

最高院量刑规范化改革,实质上是量刑指南体系在中国司法实践的一次实验,其功效如

〔68〕 当然,当存在于同一司法体系时,两种模式应当是有主次之分的,量刑指南为主,量刑信息系统为辅,后者的功能主要有三:其一,保证量刑指南的合理适用率。当量刑信息必须被系统记录并保存,法官在量刑时就会正视指南的规定,并将其作为自己量刑的依据;其二,为个体法官借鉴其他法官的量刑经验提供便捷渠道,从而在非典型案例的量刑上将更为合理;其三,为指南主管机构提供实践量刑资料,从而为指南的完善提供数据支持。

何,直接决定了中国在未来能否采取指南型量刑失衡克服模式。抛开最高院量刑规范化改革中量刑指南所涉及的量刑方法、实体与程序上的规定与前文中介绍的诸量刑指南方案之差异处不谈,下文将对最高院此次量刑指南实践对于司法体系所产生之实效加以评估。^[69]

(1)评估样本概要。在相关司法机关的大力协助下,笔者从全国21个省级行政区^[70]共77个法院中收集了包括盗窃罪、故意伤害罪、交通肇事罪在内的共22450份判决书,在这些判决书中,共涉及被告人33512人,因此,相应形成了共33512份罪刑关系,该33512份罪刑关系的基本情况如下表所述:

表四

罪名	判决年份/比重					法院级别/比重	
	2006年	2007年	2008年	2009年	缺失	基层	中级
盗窃罪	4684	5889	6680	4335	490	17259	4819
22078	21.2%	26.7%	30.3%	19.6%	2.2%	78.2%	21.8%
故意伤害罪	2074	2195	2738	2032	444	6205	3278
9483	21.9%	23.1%	28.9%	21.4%	4.7%	65.4%	34.6%
交通肇事罪	406	482	656	389	18	1725	226
1951	20.8%	24.7%	33.6%	19.9%	0.9%	88.4%	11.6%
罪名	刑罚种类/比重					是否试点院/比重	
	管制	拘役	有期徒刑	无期徒刑	其他	是	否
盗窃罪	44	4223	17357	179	275	17268	4810
22078	0.002%	19.1%	78.6%	0.8%	0.012%	78.2%	21.8%
故意伤害罪	85	1042	8048	186	122	8382	1101
9483	0.9%	11%	84.9%	2%	1.3%	88.4%	11.6%
交通肇事罪	1	52	1898	0	18	1530	421
1951	0.1%	2.6%	97.3%	0	0.9%	78.4%	21.6%

根据前述表二可知,在本次实证评估中,首先,该三个犯罪的案件从2006年开始到2009年结束,并且每年罪刑关系样本所占比重数大体相近,在20%左右;其次,由于最高院2009年6月1日开始的规范化试点法院不包括高级法院及最高院本身,因此,本次实证评估仅收集了基层法院以及中级法院的判决书,并且,在数量上,三个犯罪均呈现出基层法院判决书所占比

[69] 为了确定量刑规范化改革之实效,以决定是否开展下一阶段的量刑改革,2009年6月1日之后,在规范化改革进入第二个阶段之后,最高院委托了北京大学实证法务研究所对其120余个法院的量刑规范化试点工作进行实证评估,在研究所所长、北京大学法学院白建军教授的推荐与指导下,笔者全程参与并完成了此次实证评估工作。

[70] 其分别为:安徽、北京、福建、甘肃、广西、海南、河南、黑龙江、湖南、吉林、江苏、宁夏、山东、山西、上海、四川、浙江、西藏、重庆、浙江、天津。

重明显高于中级法院的趋势;^{〔71〕}在判处的刑罚种类上,有期徒刑是我国司法实践中适用最多的刑种,因此,在该三罪数据库中非常明显的趋势为判处有期徒刑的罪刑关系样本占据了极大的比重,其他刑种比重非常小;最后,为了分析试点之后的量刑实践,也引入了非试点法院的案件,但在本数据库中,来自试点法院的案件数量明显高于非试点法院。

(2)评估过程展示。“指导意见”出自中国最高司法机关之手,可以被假定为是公正合理的,而遵循“指导意见”的量刑实践自然实现了量刑的平衡,反之亦然。由此,通过将试点法院试点后量刑符合“指导意见”的比重(下文简称“合规率”)与试点法院试点前以及非试点法院量刑的合规率之间加以对照,就可以得出“指导意见”的实施是否能够带来量刑失衡程度降低的结论。

计算合规率的具体方法为:首先,通过对相关犯罪每个具体罪刑关系样本的浏览,得出该样本所属的犯罪构成要件等级,以及法官在量刑中所考虑的量刑情节;其次,根据“指导意见”的规定,得出前述犯罪构成要件等级所对应的起点刑。根据“指导意见”的规定,对于盗窃罪的各构成要件等级,其起点刑为固定值,而对于故意伤害罪、交通肇事罪各构成要件等级,则均规定了幅度式的起点刑;再次,根据“指导意见”的规定,基于已经得出的起点刑点或幅度,得出前述各量刑情节所对应的调节刑罚量幅度。根据“指导意见”的规定,所有量刑情节所对应的量刑调节比例均为直接的刑罚幅度或者以起点刑为基点的百分比幅度;最终,将起点刑与调节刑罚量进行相加或者相减的运算,得出根据“指导意见”所应当量处刑罚幅度区间。^{〔72〕}通过对数据库中的罪刑关系样本进行搜索,如果同样具有类似基本构成要件等级以及量刑情节的犯罪,当实际量刑处于该区间以内时,则属于合规率的范围之中,如果量刑处于该区间之外时,则属于非合规率的范围之中,由此,即可以得出罪刑关系样本的整体合规率情况。

表五

法院类型/样本数			合规样本数/合规率		
			盗窃罪	故意伤害罪	交通肇事罪
试点法院试点前			3315	2651	277
盗窃 11354	故意伤害 5458	交通肇事 392	29.2%	48.6%	70.7%
试点法院试点后			294	304	42
盗窃 1129	故意伤害 528	交通肇事 50	26.1%	57.6%	84%

〔71〕 由于对 77 个法院采取的是全样本收集法,在判决书数量上基层高于中级法院是正常的。

〔72〕 比如,根据“指导意见”的规定,假设某盗窃罪满足了基本构成要件等级,起点刑为 12 个月有期徒刑,被告人又只具有“已满十六周岁不满十八周岁”这一量刑情节,因此,该一量刑情节的调节刑罚量幅度在 2.4 个月到 6 个月的区间内,最终,根据“指导意见”的量刑区间为 6 个月(12 月减去 6 月)到 9.6 个月(12 月减去 2.4 月)有期徒刑的区间。

非试点法院			757	399	105
盗窃 3944	故意伤害 665	交通肇事 131	19.2%	60%	80.2%
总计			4369	3354	424
盗窃 16427	故意伤害 6651	交通肇事 573	26.6%	50.4%	74%
缺失					
盗窃 5651	故意伤害 2832	交通肇事 1378			

根据上表,盗窃罪的合规率在试点法院试点后竟然有下降趋势,这显然是难以接受的。通过对数据库构成的进一步分析,笔者发现了其中的原委。在盗窃罪数据库共 22078 个罪刑关系样本中,判处有期徒刑实刑的样本共有 17357 个,而在所有这 17357 个样本中,属于盗窃罪裸罪^{〔73〕}的样本有 9314 个,占有有效样本的 54%,^{〔74〕}由于裸罪不存在任何能够被法官认可的起调节起点刑轻重的量刑情节,因此对于裸罪案件不需要考虑量刑情节的调节刑罚量幅度,换句话说,根据“指导意见”,对裸罪的量刑就等于其相应构成要件等级的起点刑。但是,由于“指导意见”将盗窃罪各构成要件等级所对应的起点刑规定为一个固定的刑期(盗窃罪“数额较大”、“数额巨大”、“数额特别巨大”时起点刑分别为 12、36、120 个月),因此,数据库中这 9314 个样本的量刑如果要符合“指导意见”的规定,只有恰好等于该三个固定数值,而这显然是不合理的。再进一步统计,在该 9314 个裸罪样本中,仅有 11.7% 的样本恰好等于 12 或 36 或 120 月,属于合规率之内;相反,在 8043 个非裸罪样本中,却有 32.8% 的样本属于合规率之内,而后者合规率增大的原因,并非因为法官在存在法定量刑情节以及部分酌定量刑情节的犯罪中更容易做出合规的量刑,而是因为根据“指导意见”的规定,对于法定量刑情节以及部分酌定量刑情节对量刑所起的作用均规定了一个调节比例的幅度,因而,增大了实践样本合规的几率。总而言之,根据统计结论,盗窃罪合规率在非试点法院、试点法院试点前、试点法院试点后均保持较低的合规率,甚至在试点法院试点后较试点法院试点前在合规率上降低的“事实”,并不意味着司法实践中在对盗窃罪量刑时存在问题,而是最高院“指导意见”在盗窃罪量刑规范上存在问题。在“指导意见”将盗窃罪构成要件等级起点刑规定一个固定刑期的前提下,上表中对盗窃罪试点法院试点前后的合规率统计,不具有参考价值。

〔73〕 所谓裸罪,是指在司法实践中,刚达到犯罪某一具体构成要件等级,并未有任何量刑情节获得法官认可的犯罪状态。

〔74〕 基于数据处理的便利,此处的共 9314 个盗窃罪裸罪样本,在盗窃数额这一罪量上,并非绝对等于“数额较大”、“数额巨大”、“数额特别巨大”的起点数额,因此,该 9314 个盗窃罪裸罪样本,实际上是指:与该三个盗窃数额罪量起点相近,同时法官在刑事裁判文书中并未认可任何量刑情节的样本。

与盗窃罪的规定不同,“指导意见”在故意伤害罪、交通肇事罪各构成要件等级起点刑的问题上,代之以起点刑幅度,而非固定刑期。这样,上表中对故意伤害罪以及交通肇事罪合规率的统计就不会存在盗窃罪统计中出现的问题,根据上表可以明显看出,在故意伤害罪合规率上,试点法院在试点前,其量刑合规率仅为48.6%,而在试点之后,该合规率上升至57.6%,上升了9个百分点;而在交通肇事罪合规率上,试点法院则从试点前的70.7%上升至试点后的84%,上升了约15个百分点。然而,在非试点法院中,故意伤害罪的合规率为60%,交通肇事罪合规率为80.2%,后者较试点法院试点后合规率仅高约3.8个百分点,而前者甚至较试点法院试点后的57.6%还要低2.4个百分点。为何存在“指导意见”的试点法院在量刑合规率的比重上较不存在“指导意见”的非试点法院并不具太大优势,甚至存在劣势呢?在假定“指导意见”是从量刑实践中总结升华的在法院系统内广为推崇之规则的前提下,对此的一种估计为:非试点法院相较于试点法院而言,其量刑公正合理性的程度一贯更高,在解决量刑失衡的问题上本来就更出色。假设将“指导意见”的试点工作交由本次实证评估抽点的非试点法院去试点,这些“试点法院”的预期合规率将比真正试点法院的合规率更高。

(3)评估结论汇总。根据前文对盗窃、故意伤害、交通肇事三罪量刑合规性的实证评估,可以得出如下两个结论:

首先,盗窃罪实证评估的失败,说明“指导意见”对犯罪各构成要件等级的起点刑应当规定一个幅度,而非一个固定刑期,否则,将与司法实践中法官群体的量刑经验相悖,导致“指导意见”的规定不被遵循。^[75]

其次,也是最重要的结论,“指导意见”在试点法院试点后合规率上升的现象证明了指南型量刑失衡克服模式在中国的可行性。

(三)中国化量刑指南的经验根基

从前文的实证分析可以看出,量刑指南在中国具有实践可行性。就量刑指南中国化的具体操作而言,通过参考国外量刑指南的制定历程可知:量刑指南的规定无一不是司法实践经验升华的结果,^[76]我国最高院亦宣称自己的“量刑指导意见”来源于审判实践,^[77]由此可见,是否具有经验根基是量刑指南绕不开的话题。“量刑指导意见”提出了规范化量刑的三个步骤,分别是:根据基本犯罪构成事实在相应的法定刑幅度内确定量刑起点;根据其他影响犯罪构成的犯罪数额、犯罪次数、犯罪后果等犯罪事实,在量刑起点的基础上增加刑罚量确定基准刑;根据量刑情节调节基准刑,并综合考虑全案情况,依法确定宣告刑。而遵循该三步量刑法的核心在于确定量刑起点、增加刑罚量以及量刑情节调节基准刑。“量刑指导意见”对于量刑起点以

[75] 实际上,在2010年10月1日的《人民法院量刑指导意见》中,最高院即对此做出了修改,给盗窃罪各构成要件等级起点刑规定了一个幅度。

[76] 美国联邦量刑指南得源于对10000份判决前调查报告以及美国假释委员会与假释相关的统计数据实证统计之结果;韩国量刑委员会在制定其量刑指南前,即建立了一个从2004年3月1日起到2007年2月28日为止的包括43000个案件的数据库,希冀从中获知实践量刑经验以体现在其量刑指南中。

[77] “量刑规范化改革所确定的基本思路和主要内容都来源于审判实践,是对长期以来各地所积累的审判实践经验的概括和提炼。”参见高憬宏:“量刑改革应澄清的认识误区”,《法制资讯》2010年第22期。

及量刑情节调节基准刑均有明确的指导性取值幅度,而对于增加刑罚量,则规定得较为模糊。^[78]撇开尚未做明确规定的增加刑罚量不谈,就量刑起点及量刑情节调节基准刑而言,“量刑指导意见”对二者的规定是否存在经验性根基?又该如何找寻其经验性根基?在下文中,笔者将尝试对其展开研究。

1. 量刑起点的经验根基

在中国学界,对于量刑起点如何确定,有各种观点,比如中线论、上限论、下限论、分格论、重心论、形势论、均值论、众值论等。^[79]笔者认为,上述各种方法可以从本质区别成两类,一为演绎型确定法,即通过逻辑推演来确定裸罪所对应的起点刑,上述中线论、上限论、下限论、分格论、重心论、形势论,均属于此类;一为经验型确定法,即通过对司法实践中法官量刑经验的归纳来确定裸罪所对应的起点刑,上述均值论、众值论均属于此类。前者凸显着量刑起点所具有的规范意义,而后者更看重与实际问题联系在一起的规则以及尊重法官群体的量刑经验。在此,笔者并不排斥由权力部门基于一定刑事政策的目的,通过逻辑演绎最终确定量刑起点的做法,但是,这种演绎的起点应当来源于司法经验,与司法经验差距甚远的逻辑演绎,必将因为缺乏实践支撑而难以为继。因此,“量刑指导意见”中量刑起点的规定应当与实践经验相当,从而获得经验根基。

为了验证“量刑指导意见”中量刑起点是否具有经验根基,下文将对盗窃罪、故意伤害罪以及交通肇事罪三罪裸罪各构成要件等级实践量刑^[80]之均值与“量刑指导意见”的具体规定加以对比,情况如下表所示:

表六

裸罪犯罪构成要件等级		数据库量刑均值	量刑指导意见	对比情况
盗窃罪	基本构成要件	9.7	3 * ~6	高于上限(不符合)
	加重构成要件	43.8	36~48	基本居中(符合)
	再加重构成要件	122.7	120~144	接近下限(基本符合)

[78] 以盗窃罪为例,在最高法院 2010 年 10 月 1 日公布的《人民法院量刑指导意见》中,其基本构成要件等级的量刑起点规定为:“达到数额较大起点的,或者一年内入户盗窃或者在公共场所扒窃三次的,可以在三个月拘役至六个月有期徒刑幅度内确定量刑起点”;而对于增加刑罚量,则是这样规定的:“在量刑起点的基础上,可以根据盗窃数额、次数、手段等其他影响犯罪构成的犯罪事实增加刑罚量,确定基准刑”。对于量刑情节调节基准刑规定得较为明确,如就未成年人犯罪而言:“已满十四周岁不满十六周岁的未成年人犯罪,可以减少基准刑的 30%—60%”。

[79] 参见苏惠渔等,见前注[1],页 102—105。

[80] 提取该三罪裸罪量刑起点均值的样本仍然来自前文中从全国 21 个省级行政区共 77 个法院中收集的共 22450 份判决书所组成的共 33512 份罪刑关系数据库中。至于该三罪基本、加重、再加重构成要件等级的确定标准,则与“量刑指导意见”的相关规定保持基本一致。

故意伤害罪	基本构成要件	12.1	6~18	基本居中(符合)
	加重构成要件	54.9	36~48	高于上限(不符合)
	再加重构成要件	138.1	120~180	接近下限(基本符合)
交通肇事罪	基本构成要件	17.3	6~24	基本居中(符合)
	加重构成要件	43.5	36~48	基本居中(符合)
	再加重构成要件	94.3	84~96	接近上限(基本符合)

* 根据最高院“量刑指导意见”的规定,盗窃罪基本构成要件起点刑下限为3个月拘役。除此之外,上表中起点刑数值均表示有期徒刑月数。

根据“量刑指导意见”的规定,上述三罪量刑起点并不是一个固定值,而是一个幅度。一般来说,这个幅度对应着司法实践从最轻微到最严重之两极,而数据库中对三罪裸罪各构成要件等级量刑之均值,则体现着最通常的情况,其至少应位于“量刑指导意见”规定幅度之内,而最佳的情况则是量刑均值应位于“量刑指导意见”规定幅度之中心位置。对比上表来看,其中:盗窃罪基本构成要件以及故意伤害罪加重构成要件量刑均值均已超过“量刑指导意见”幅度之上限,因此,“量刑指导意见”相应的量刑起点之规定从量刑实践检验的结果来看,是有待商榷的。而在其他情况下的量刑均值则基本位于“量刑指导意见”量刑起点幅度的中心位置或者处于量刑起点的幅度之内,意味着“量刑指导意见”量刑起点幅度在这些情况下的规定反映了司法实践的现实情况或者至少没有与实践相去太远。

总之,从上述三罪数据库的实证分析来看,假设本数据库的案件组合代表了司法实践中盗窃罪、故意伤害罪以及交通肇事罪量刑现状,那么,“量刑指导意见”中对该三罪量刑起点的规定则并非全具有经验根基,有再完善的必要。退一步说,即使认为笔者的数据库不能代表司法实践该三罪量刑现状,此处的研究亦并非全无意义,我们完全不必过分纠结于前文所得出的对“量刑指导意见”经验根基之检验结论,所关注的应是求得量刑起点经验根基的方法,在各方达成一致的前提下建立一个最贴近司法实践现状的数据库,通过前述方法归纳出各罪各构成要件等级裸罪量刑均值,再将其在理性评价的基础上升华为量刑指南上所规定的量刑起点,从而达到理性与实践的协调,造就起中国量刑指南在量刑起点上的经验根基。

2. 量刑情节调节基准刑的经验根基

(1) 量刑情节设定范围的经验性。“量刑指导意见”在罪名范围上只针对司法实践中最常见的15类犯罪,根据最高院的统计,该15个犯罪位列全国刑事案件比重的前20位以内,而前20位的犯罪占全国案件比重的90%以上,因此,规范好了“量刑指导意见”中规定的这15个犯罪,同时在将来进一步规范好剩下的5个犯罪,就将“很大程度上解决量刑不规范的问题”。^{〔81〕}在罪名范围上的这种“不理睬琐细”的指导方针,当然还可以体现在量刑情节的范围

〔81〕 参见陈秀军:“严格程序 规范量刑 确保公正——最高人民法院刑三庭负责人答记者问”,载《人民法院报》,2009年6月1日,第2版。

上,“量刑指导意见”在量刑情节的设定范围上,应仅针对司法实践中最常见的量刑情节。这一方面是基于从经验主义的路线出发,只有那些经常出现在法官量刑时的量刑情节,才应当被规定在“量刑指导意见”中,而那些备而不用量刑情节,由于其发生概率过小,而不具有规范该类量刑情节在量刑中所起作用程度的紧迫性;另一方面,也是由于“量刑指导意见”上量刑情节调节基准刑的规定应具有经验根基,而经验根基的求得则依赖大规模量刑样本的存在,至于司法实践中发生率不高的量刑情节,则可以暂时不纳入最初版本的“量刑指导意见”中,留待之后在司法实践积累了成规模样本的前提下,再归纳出其中的实践经验,最终上升至细化了的后续版本中。

总之,笔者认为在探讨量刑情节调节基准刑的经验根基之前,应首先从发生概率的基础上,确定“量刑指导意见”中量刑情节的设定范围。下文将对 71653 个罪刑关系样本进行实证分析,^{〔82〕}以统计样本数据库中各量刑情节的发生频率,并将之与“量刑指导意见”中量刑情节的设定范围加以对照。

表七

量刑情节	发生频率	发生比重	“量刑指导意见”规定与否
自首	8057	11.2%	有
未成年人犯罪	7230	10.1%	有
累犯	6448	9%	有
未遂犯	5884	8.2%	有
从犯	5850	8.1%	有
立功	3856	5.4%	有
预备犯	784	1.1%	无
聋哑盲人犯罪	554	0.8%	无
中止犯	195	0.3%	无
正当防卫过当	55	0.08%	无
教唆犯	17	0.02%	无
胁从犯	8	0.01%	无
缺省值	32715	45.69%	

〔82〕 本次分析的数据库情况介绍:在前述三罪共 33512 份罪刑关系所组成数据库的基础上,笔者又加入了同样时间段内相同法院判处的强奸罪,非法拘禁罪,抢劫罪,诈骗罪,抢夺罪,职务侵占罪,敲诈勒索罪,妨害公务罪,聚众斗殴罪,寻衅滋事罪,走私、贩卖、运输、制造毒品罪,掩饰、隐瞒犯罪所得、犯罪所得收益罪的所有判决书,新数据库共有样本 38141 份,前后数据库合并,组合成 71653 个罪刑关系样本数据库。

根据上表所示,自首、未成年人犯罪、累犯、未遂犯、从犯、立功这几个在数据库中发生频率在5%以上的量刑情节,“量刑指导意见”均作出了明确的规定;而对于预备犯、聋哑盲人犯罪、中止犯、正当防卫过当、教唆犯、胁从犯这几个在数据库中发生频率在5%以下的量刑情节,“量刑指导意见”则未作规定。虽然预备犯、聋哑盲人犯罪、中止犯这三个量刑情节在数据库中的发生频率绝对值分别为784、554、195,均过百数,但是,由于其所占整体比重不高,笔者亦认为在第一版的“量刑指导意见”中不宜加以规定,有待以后再细化。总之,如果以发生比重是否超过5%作为衡量某一量刑情节发生多寡的标准,那么,“量刑指导意见”在量刑情节设定范围上的经验根基不言而喻。

(2)量刑情节调节基准刑的经验性。在刑事立法中,法定刑幅度的设定耐人寻味,比如,在盗窃罪的规定上,“盗窃公私财物,数额较大的,或者多次盗窃、入户盗窃、携带凶器盗窃、扒窃的,处三年以下有期徒刑、拘役或者管制,并处或者单处罚金”,对此规定,一旦深究,难免会产生为何盗窃公私财物,数额较大的,或者多次盗窃、入户盗窃、携带凶器盗窃、扒窃的行为法定刑幅度是“三年以下有期徒刑、拘役或者管制,并处或者单处罚金”而不是其他的疑问。如果存在着对上述行为法定刑幅度之“定在”,那么,立法者所要做的就是发现这一“定在”,并将之上升为刑法规定。然而,现实中某一罪行所应配置的法定刑幅度,因时、因地、因人而不同,根本没有“定在”可寻,绝对精确的罪刑阶梯是不存在的,只能寻求建立相对精确的罪刑阶梯,即:在社会发展状况拟制出某一罪行之法定刑幅度后,再比照其他罪行的危害程度,按照相当的比例来确定全部罪行的法定刑幅度,最终达到“重罪法定刑更重,轻罪法定刑更轻”的相对罪刑阶梯状态。相对罪刑阶梯同样可以适用在量刑情节调节基准刑的设定上,“量刑指导意见”中对一般立功的规定为“可以减少基准刑的20%以下”,对此,在相对罪刑阶梯状态下,追问为何是“20%以下”而不是“40%以下”意义不大,值得考究的应是:如果一般立功能带来减小“20%以下”的量刑效果,重大立功理应更甚。“量刑指导意见”规定:“重大立功的,可以减少基准刑的20%~50%”,据此可知,在量刑情节调节基准刑比例的设定上,重大立功对量刑的影响效果大于一般立功,而这样的规定不仅于法有据,而且于情有理,两者之间遵循了相对罪刑阶梯原则。

对于刑法有明确规定或者根据刑法内在解释能够推论出来的各量刑情节在量刑影响效果上的序位关系,“量刑指导意见”理应加以采纳,但是,并非所有量刑情节之间的序位关系均能通过逻辑推演得出,比如,一般立功和一般情况下的自首这两个法定量刑情节所带来的量刑效果均为“可以从轻或者减轻处罚”,那么,两者在量刑影响效果上的序位关系应如何规定?在法定量刑情节之外,一般立功和积极赔偿被害人损失这样一个明确化的酌定量刑情节,两者的序位关系又应如何规定?对此,一种方案为根据量刑情节的法律效力层级来确定相互间的影响序位,而对于处于同一法律级别的量刑情节,在没有明确规定的情况下,则具有相同的量刑影响程度。我国学者赵廷光教授的“辅助量刑系

统”即为其中的佼佼者。赵教授根据量刑情节的出处,比如,是属于刑法规定,还是司法解释规定,还是酌定量刑情节,人为地确定为五个“重要性等级评价”,即:“特别次要情节”(积0分)、“次要情节”(积5分)、“一般情节”(积10分)、“重要情节”(积15分)、“特别重要情节”(积20分)。^{〔83〕}根据量刑情节的来源确定量刑情节的量刑影响程度,虽然在逻辑上是有道理的,但是,其忽视了一个最根本的问题,即:不同量刑情节虽然来源于同一个法律层级,但其对量刑的影响程度是有差异的,比如,根据赵教授的“辅助量刑系统”,自首与累犯均被归入“重要情节”的范畴,具有相同的量刑影响程度,而这显然是有悖于法律规定的。^{〔84〕}再比如,完全从规范的角度,所有司法解释规定的量刑情节对量刑的影响程度都小于任何法定量刑情节,但在司法实践中往往并非如此,在诸如交通肇事、过失伤害等犯罪中,被告人足额赔偿被害人经济损失这一司法解释上的量刑情节对量刑产生的作用,甚至比自首这一刑法明文规定的量刑情节还要大。^{〔85〕}由此可见,根据量刑情节的法律效力层级来确定其对量刑的影响程度的做法是值得商榷的,即使其最终能够在形式上“符合”法理,其在司法实践中也将难以为法官群体所认可。

解决量刑情节序位关系的另一种方案则不拘泥于形式地遵循法律效力层级,对于刑法有明确规定以及依据刑法精神可以明确推导出的各量刑情节对量刑影响程度的排序情况,该方案予以全盘采纳;但对于刑法并无明确规定、亦无法从刑法精神中明确推导的法定量刑情节以及非法定量刑情节的序位情况,则并不形式地依照法律效力层级,而更多地依赖司法实践中法官的量刑经验,在量刑经验的基础之上获得各量刑情节对量刑影响的排序,在此基础上根据立法精神、刑事政策进行理性再评价,最终得出各量刑情节的排序情况。下表为笔者对与前文相同的,涉及到“量刑指导意见”所规定的15类犯罪的共71653个判处有期徒刑实刑的罪刑关系样本的回归分析结论,通过该分析,得出数据库中各法定量刑情节的非标准回归系数以及其排序情况,并将该实践排序情况与最高院“量刑指导意见”规定的排序情况进行对比,以此来检验“量刑指导意见”的经验根基。

〔83〕 当然,在现实案件中,某一个具体的量刑情节,由于其在具体案件中的表现形式不同,因此,赵教授再次对具体量刑情节分出五个“具体表现等级评价”,即:“危害危险轻微的”,积5分;危害危险较轻的,积10分;危害危险一般的,积15分;危害危险较重的,积20分;危害危险严重的,积25分。最后,根据“重要性等级评价”以及“具体表现等级评价”的综合得分,得出现实案件中某量刑情节的最终积分,该积分即对应着一定的量刑结论。

〔84〕 根据刑法规定,一般自首情节“可以从轻或减轻处罚”,而一般累犯情节却是“应当从重处罚”,两者在对量刑的影响程度上很难说是相同的。

〔85〕 比如,在笔者对261个交通肇事罪案例进行实证研究的过程中,当以足额赔偿被害人经济损失与自首作为自变量,以最终的量刑结果作为因变量进行回归分析时,发现足额赔偿被害人经济损失这一自变量的非标准化回归系数为-11.024,而自首自变量的非标准化回归系数为-8.865,两者显著性值均为0.000,说明该两个自变量对量刑结果具有显著的影响关系,同时在程度上,足额赔偿被害人经济损失对最终量刑结果的影响程度要大于自首。参见蔡曦蕾,见前注〔67〕。

表八

量刑情节	“指导意见”规定		实践数据库情况	
	调节幅度	排序	非标准回归系数	排序
重大立功	20%~100%	1	-35.233	1
从犯	20%~100%	1	-3.840	5
自首*	0%~100%	2		
未成年人犯罪	10%~60%	3	-12.419	2
未遂犯	0%~50%	4	-6.892	4
累犯	10%~40%	5	3.784	6
一般立功	0%~20%	6	8.753	3

* 除自首这一量刑情节由于对最终量刑结果不具有显著性而被系统自动排除外,其作为自变量的量刑情节对作为因变量的量刑结果的显著性均为 0.000,显著性程度非常高。

根据上表,“量刑指导意见”中各法定量刑情节对量刑结论的影响程度排序与实践数据库中各法定量刑情节的排序相比,在重大立功、未成年人犯罪、未遂犯、累犯这几个情节上,排序基本一致,但是,在从犯与一般立功这两个情节上,在“量刑指导意见”中,从犯排序为第 1,就也就是说,最高法院认为从犯这一法定量刑情节在如上 7 个法定量刑情节中,对量刑结论的影响程度最大,但在实践数据库中,从犯排序仅为第 5,对量刑结论的影响相对较小;同样,在“量刑指导意见”中,一般立功排序为第 6,在最高法院看来,一般立功对量刑结论的影响程度相对较小,但是,在实践数据库中,一般立功排序却为第 3,对量刑结论的影响相对较大。总之,在量刑情节调节基准刑的经验根基问题上,“量刑指导意见”在多个法定量刑情节的排序上与实践经验相符,具有经验性根基;但是,在从犯与一般立功这两个量刑情节上,则与司法实践中的普遍经验存在较大背离。

文末,再次重申的是,虽然前文的实证分析指出“量刑指导意见”中部分量刑情节调节基准刑在相对序位上缺乏经验根基,但重点并不在于结论,而在于提供一个如何确定量刑情节调节基准刑排序的研究范例,笔者将其总结为以下流程:首先,根据法律明文规定确定各量刑情节调节基准刑序位;其次,当法律并无明文规定时,以司法实践序位为基础,形成各量刑情节调节基准刑的实践序位;最后,已经形成的序位并非必须原样“复制”到成文的“量刑指导意见”中,而是在立法精神、刑事政策的指导下进行理性再评价,对不合理的实践经验加以微调,最终形成理性与经验并存的量刑情节调节基准刑。

(责任编辑:江 溯)