

司法正义的区差规律

——基于量刑地区差异的Blinder-Oaxaca分解

江国华 孙琪伟

摘要:司法正义是一个兼具普遍正义和差序正义的复合体。司法正义的普遍性源自法治统一原则的内在稳定性,而司法正义的差序性则属于区域发展差序结构之产物。以刑事司法为例,基于量刑地区差异的Blinder-Oaxaca分解,可以客观地呈现司法正义的区差规律:法定因素是导致量刑地区差异的主要原因,罪行严重程度越高的地区,其量刑结果越重;非法定因素即法官的自由裁量对量刑地区差异的影响是有限的;社会经济发展水平深刻影响法官的个案认知,社会经济发展水平越高的地区,法官对相同犯罪数额危害程度的认定越宽松。为此,有必要建立裁判辅助系统,根据地区经济指标对犯罪数额进行折算,引入量刑地区差异的监测模型,并且强化裁判文书的撰写规范,以确保将正义区差控制在合理的范围内。

关键词:司法正义;差序正义;量刑地区差异;Blinder-Oaxaca分解法;区差规律

DOI: 10.19836/j.cnki.37-1100/c.2024.06.006

一、问题提出

司法裁判地区差异在全球范围内是一种普遍现象,中国也不例外^①。2020年最高人民法院公布的《关于统一法律适用加强类案检索的指导意见(试行)》第四条规定,类案检索的范围限于最高人民法院以及本省行政区划内的法院裁判生效的案件,这意味着最高人民法院也承认中国司法裁判存在地区差异。司法裁判地区差异有侵蚀司法公正的风险。2023年3月8日最高人民法院工作报告强调“维护和促进社会公平。加强裁判尺度统一,切实防止公平正义因地区、城乡、行业差异而打折扣”^②。

党中央高度重视司法公正。习近平总书记指出:“公正司法是维护社会公平正义的最后一道防线。”^③中国宪法第33条规定“公民在法律面前一律平等”。“公平、平等皆以公正为归属,依归于公正。”^④据此,司法公正指的是司法机关平等地适用法律,即同等情况同等对待、不同情况不同对待。正如世界上不存在两片相同的树叶,世界上也不存在绝对意义上相同的案件^⑤。案件之间在具有相似性的同时也具有差异性。所以,司法公正要求法官裁判时要在统一性和差异性之间寻求均衡。最高人民法院为此作出过诸多努力,其公布的一系列文件都阐述了何谓“在统一性和差异性之间寻求均

基金项目:国家社科基金重点项目“中国案例制度体系与实践主义司法哲学研究”(16AZD025)。

作者简介:江国华,武汉大学法学院教授,教育部长江学者特聘教授(武汉 430000; fxyjgh@whu.edu.cn);孙琪伟,武汉大学法学院博士研究生(武汉 420000; 2427057938@qq.com)。

① 荣学磊、余艺:《对司法裁判偏离公正现象的法理分析——基于对民事审判实证调查的视角》,《法学》2012年第2期。

② 周强:《最高人民法院工作报告》,《人民日报》2023年3月18日,第4版。

③ 习近平:《高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告》,《人民日报》2022年10月26日,第1版。

④ 吴忠民:《关于公正、公平、平等的差异之辨析》,《中共中央党校学报》2003年第4期。

⑤ 周少华:《刑事案件的差异化判决及其合理性》,《中国法学》2019年第4期。

衡”。“统一性”是指法官适用法律时注重法律统一适用^①；“差异性”是指法官在维护法律统一适用的前提下，注重差异的包容性^②；“寻求均衡”是指法官对同一时期、同一地区、案情相似的案件作出的裁判应当基本均衡^③。

从司法实践来看，判断司法裁判的地区差异是否符合均衡性原则对深入推进司法公正具有重要意义。区差规律是司法裁判地区差异所反映出来的规律，把握这一规律能够为司法公正的量化评估提供依据，也能够为司法改革的推进以及司法政策的干预提供决策参考。追溯已有的文献，有学者证实中国司法裁判存在地区差异^④。然而，这种差异性的存在是否以坚持法律统一适用为前提？法官在裁判过程中是否在统一性和差异性之间实现均衡？目前，尚未有研究以事实为据回答上述问题。

为探讨这一差异是否合理，以刑事司法为例，通过对2020年我国31个省级行政区受贿罪所有已公开的刑事判决书进行分析，进一步拓展以往研究的深度。首先，参考已有研究的结论，根据31个省级行政区综合经济发展水平的排行，将其划分三个地区，使用OLS法分别量化了三个地区法定因素（数额和情节）与量刑结果之间的关联性，并且使用工具变量法对可能存在的内生性问题进行检验；然后，以回归模型的分析结果为基础，使用Blinder-Oaxaca分解法两两分解地区之间罪行严重程度上的差异和法官在罪行认定（法官自由裁量）上的差异，用以佐证法官的自由裁量如何影响以及在多大程度上影响了量刑地区差异。最后，总结司法公正的区差规律，为进一步维护司法公正提出相应建议。

二、文献综述与研究假设

司法公正是一个内涵丰富的概念。在量化分析之前，首先要基于现有研究对司法公正的含义和分类进行简要回顾，根据学界主流观点，明确欲探讨的司法公正的语境。其次，对现有研究关于司法公正量化的评估方案进行梳理，进而提出合适的评估方案。最后，对探讨量刑地区差异影响因素的相关研究成果进行简要回顾，并提出研究假设。

（一）司法公正的含义和分类

“司法公正”的含义可以分为以下几种学说：适用法律说^⑤、解决纠纷说^⑥、结合说^⑦、个案审理的结果说^⑧、司法权力运作状态说^⑨等。从不同的研究视角出发，“司法公正”主要包括以下几种分类：实体公正和程序公正^⑩、整体公正和个案公正^⑪、社会公正和法律公正^⑫。

借鉴以往研究成果，司法公正可以被认为是司法机关在适用法律时所遵循和体现的公平正义，并且这种公平正义指的是实体公正、整体公正和法律公正。

① 《最高人民法院印发〈关于在审判执行工作中切实规范自由裁量权行使保障法律统一适用的指导意见〉的通知》（法发〔2012〕7号）。

② 《最高人民法院关于切实践行司法为民大力加强公正司法不断提高司法公信力的若干意见》（法发〔2013〕9号）。

③ 《最高人民法院 最高人民检察院印发〈关于常见犯罪的量刑指导意见（试行）〉的通知》（法发〔2021〕21号）。

④ 王剑波：《我国受贿罪量刑地区差异问题实证研究》，《中国法学》2016年第4期；文姬：《醉酒型危险驾驶罪量刑影响因素实证研究》，《法学研究》2016年第1期；陈磊：《贪污受贿犯罪量刑均衡问题实证研究》，《政法论坛》2020年第1期。

⑤ 吕忠梅：《司法公正价值论》，《法制与社会发展》2003年第4期。

⑥ 周友苏：《构建和谐社会中的司法公正问题研究》，《社会科学研究》2006年第3期。

⑦ 王利平：《司法改革的困境与路径》，《中共福建省委党校学报》2011年第9期。

⑧ 秦伟、夏珍：《相对司法公正及其成因》，《山东大学学报（哲学社会科学版）》2008年第5期。

⑨ 梁平：《基层治理的践行困境及法治路径》，《山东社会科学》2016年第10期。

⑩ 董晖：《司法功能与司法公正、司法权威》，《政法论坛》2002年第2期。

⑪ 董玉庭：《司法体制改革不能忽视的四种关系》，《求是学刊》2017年第1期。

⑫ 周帆：《论司法公正的法律标准与社会标准之共生》，《求索》2010年第8期。

(二)司法公正的量化评估

以往对司法公正的量化评估方案以“同案同判”为标准,检验司法裁判结果是否符合“统一性”原则^①。这种评价标准存在一定局限性:其一,这种评价标准的推断不够严谨。从裁判结果来看,虽然不同地区案情相似的案件其裁判结果存在一定的地区差异性,但是不能据此就得出司法不公正的结论。任何一个案件可能都不会仅有一个唯一的处理结果^②,在依法裁判的前提下,同案同判只是法律适用可能出现的结果。因此,从“同案同判”出发,无法反向推断出法官不合理使用自由裁量权的结论。其二,这种评价标准不能直接触及司法公正。同案同判体现了公众基于朴素正义观对司法公正的期盼,属于社会公正。以同案同判为标准对审判差异进行验证,直接证明的是司法权威的问题,只能间接验证司法公正。

由此而言,法官在裁判时需寻求统一性和差异性之间的均衡以实现司法公正。“统一性”表现为罪行越严重的地区其裁判结果越重,应该符合合法性要求,即“罚当其罪”;“差异性”表现为法官自由裁量对量刑地区差异的影响虽然是主观认定的,但并不是武断的、恣意的,应该符合合理性要求,即具有客观依据;“均衡”表现为量刑地区差异主要受法定因素影响,非法定因素对量刑地区差异的影响有限。

(三)量刑地区差异的影响因素

以往研究认为,法定因素(法律明确规定的因素影响裁判结果的因素)和非法定因素(法律没有明确规定的影响裁判结果的因素,如地区、性别、年龄等)都会造成量刑结果的地区差异。其中,法定因素造成量刑结果的地区差异是因为不同地区罪行严重程度不同;非法定因素造成量刑结果的地区差异是因为不同地区的法官对相同罪行的危害性程度认定不同,即不同地区的法官存在自由裁量上的差异^③。由此获得研究假设1:

研究假设1:受贿罪量刑结果受法定因素和非法定因素影响。

据刑法第386条之规定,法定因素包括数额和情节。非法定因素是法定因素以外的因素,最终都由法官自由裁量权的行使而体现出来。值得注意的是,法官自由裁量权大小是指非法定因素造成的差异占总差异的比重。法定因素导致的差异为可解释部分,又称禀赋因素导致的差异,是指由罪行恶性程度差异导致的量刑结果差异,即数额大小和情节严重程度差异导致的量刑结果的差异。非法定因素为不可解释部分,又称环境因素导致的差异,是指由于数额和情节之外的因素导致的差异。

还有研究表明,不同区域间法官的自由裁量差异的主要原因是经济发展水平不同。地区间经济发展水平不同,导致不同地区的法官对相同数额案件的社会危害程度的认定不同^④。由此获得研究假设2:

研究假设2:法官对相同数额案件的危害程度的认定主要受地区经济水平影响。

以经济发展水平为依据划分区域,可以制造地域之间外部环境的经济差异,即可以在模型中控制环境因素中的经济发展水平,因此,环境因素所导致的差异可以近似认为主要由经济发展水平所导致。其中,环境因素中还存在其他因素影响法官的自由裁量,受数据和研究方法所限,不能将它们的影响效果从环境因素的影响效果中分离出来。

① 荣学磊、余艺:《对司法偏离公正现象的法理分析——基于对民商事审判实证调查的视角》,《法学》2012年第2期;陈海嵩:《环境侵权案件中司法公正的量化评价研究》,《法制与社会发展》2018年第6期;武西锋、秦前红:《精神损害赔偿同案同判:实证检验和实现路径》,《学术论坛》2020年第3期;吴雨豪:《量刑自由裁量权的边界:集体经验、个体决策与偏差识别》,《法学研究》2021年第6期。

② 张军:《法官的自由裁量权与司法正义》,《法律科学(西北政法大学学报)》2015年第4期。

③ 胡昌明、马铁丰、赵荀:《量刑确定性的实证研究》,《山东大学学报(哲学社会科学版)》2023年第1期;吴思诗:《美国量刑理论中的关注焦点理论》,《犯罪研究》2021年第1期。

④ 王剑波:《我国受贿罪裁判地区差异问题实证研究》,《中国法学》2016年第4期;陈磊:《贪污受贿犯罪量刑均衡问题实证研究》,《政法论坛》2020年第1期。

三、研究设计

(一) 样本选择与数据来源

以2020年全国已公开的受贿罪判决书为基础,针对受贿罪展开分析以探究量刑存在地区差异的原因。其原因在于相比于民事判决和行政判决,对影响刑事判决的法定因素进行计量分析更为容易。检索时间为2021年3月1日。根据人民法院法律文书公开时限的规定,结案后应该在15日内在中国裁判文书网公开,理论上2021年3月1日检索结果应该已经覆盖全国2020年结案的所有受贿罪案件。样本选择的具体步骤如下所示:第一,通过设置高级检索条件检索到2256份判决书,并将所有裁判文书下载之后进行统一编号;第二,研究人员通过讨论确定实证分析所需的信息,并转化为对应的变量和选项;第三,采用人工阅读裁判文书的形式将相关信息以选项的形式录入Stata16.0软件,为保障数据的准确性,所有数据录入均由2名研究人员单独完成并核对。删除空白或者信息缺失的判决书189份后得到2067份判决书后,最终提取了2201名被告人的判决信息(部分文书中包含多名被告人信息)。

为得到不同的经济分区,需要以经济发展水平为依据对31个省级行政区进行分类。借鉴王丙申的研究结果,以中国经济发展综合水平为依据,将其划分为五类^①。受限于篇幅,为简化分析流程,将五类地区^②整合为三类,分别为:发达地区和较发达地区(一区)、发展较好地区(二区)、发展一般地区和较落后地区(三区)。

(二) 研究方法与模型构建

首先运用普通最小二乘法(ordinary least square method)分区域对影响量刑的相关因素进行回归分析,以一区和二区为例,构建如下模型:

$$Y^E = X^E \beta^E + \epsilon^E \quad (1)$$

$$Y^C = X^C \beta^C + \epsilon^C \quad (2)$$

式(1)和式(2)中的 Y^E 和 Y^C 分别代表一区和二区量刑结果; X^E 和 X^C 分别代表一区和二区受贿罪量刑的影响因素,包括数额和情节(自首、立功、共犯、坦白或如实供述、认罪认罚、认罪悔罪、退赃、索贿和其他较重情节); β^E 和 β^C 分别代表一区和二区受贿罪量刑影响因素的回归系数,即该地区法官对罪行严重程度的认定; ϵ^E 和 ϵ^C 分别代表一区和二区量刑结果的残差。

在回归分析的基础上使用Blinder-Oaxaca分解法测算各个影响因素对量刑地区差异的贡献(率)。结合回归分析的结果,Blinder-Oaxaca分解法的运算公式如下所示:

$$\bar{Y}^E - \bar{Y}^C = \underbrace{(\bar{X}^E - \bar{X}^C) \hat{\beta}^*}_{\text{可解释部分}} + \underbrace{\left[\bar{X}^E (\beta^E - \hat{\beta}^*) + \bar{X}^C (\hat{\beta}^* - \beta^C) \right]}_{\text{不可解释部分}} \quad (3)$$

式(3)中 \bar{Y}^E 和 \bar{Y}^C 分别代表一区和二区受贿罪量刑均值, \bar{X}^E 和 \bar{X}^C 分别代表一区和二区受贿罪量刑影响因素的均值, β^E 和 β^C 分别代表一区和二区受贿罪量刑影响因素的回归系数。

此外,Blinder-Oaxaca分解法的指数基准问题对于分解结果有重要影响,因此,需要结合研究的实际需求对 $\hat{\beta}^*$ 的选择进行讨论。传统的Blinder-Oaxaca分解法直接选择一区(或者二区)群体的回归

① 王丙申、魏艳华、朱琳:《中国经济发展水平的综合评价》,《统计与决策》2021年第9期。

② 依据王丙申等人的研究结论,将我国的省级行政区分为五类,聚类结果如下:第一类为发达地区,包括北京、上海;第二类为较发达地区,包括天津、广东、江苏、浙江;第三类为发展较好地区,包括重庆、吉林、湖北、福建、内蒙古、辽宁、山东;第四类为发展一般地区,包括陕西、河北、宁夏、黑龙江、湖南、海南、河南、新疆、山西、四川、江西、青海、安徽、广西、西藏;第五类为较落后地区,包括云南、贵州、甘肃。

系数作为 $\hat{\beta}^*$,这种方法容易对歧视度的估计造成较大偏差^①。Cotton提出以两个群体的人数占比作为权重分别与两个群体的回归系数进行加权,加权之后的两个回归系数再相加即为 $\hat{\beta}^{*②}$ 。这种方法能够得出理想状态下绝对公正的法官判罚尺度,从而分解出现实情况与理想状态下的差异。然而,司法公正的核心理念并不是盲目排斥法官的自由裁量^③,要求裁判尺度的绝对统一,而是要将地区之间的不合理差异控制在合理范围内。相反,过度追求统一会造成刑罚适用的机械和僵化,同样有损司法公正^④。与司法公正的核心理念相结合来考虑,不难发现,这种方法也不是理想选择。Neumark提出以选择全样本系数作为 $\hat{\beta}^{*⑤}$ 。这种方法能涵盖更多的信息,其分解结果能够直接体现法官判罚尺度的地域差异,比较契合司法公正的核心理念。综上所述,可选择全样本系数作为 $\hat{\beta}^*$,即式(3)中, $\hat{\beta}^*$ 代表一区和二区的全样本回归系数。

(三)描述性统计

由于主刑量刑结果既包括年份,也包括月份,为方便运算,运用进一法全部换算为年,如:判处有期徒刑六个月赋值为1,判处有期徒刑一年零九个月赋值为2,其余照此设定。最终的样本量共计2123人,一区、二区和三区分别有533人、381人和1209人^⑥。

如表1所示,数额和情节这两个影响受贿罪量刑的法定因素为自变量。为了使数据满足正态分布,便于统计分析,受贿数额取自然对数值。从受贿数额来看,2020年中国不同经济区域的平均受贿数额由高到低的排序为:三区>一区>二区。选取的情节涵盖了所有法定量刑情节、判决书中出现的频数较高的酌定情节以及《关于办理贪污贿赂刑事案件适用法律若干问题的解释》第1条第2款和第3款中规定的9种其他较重情节。

表1 2020年全国受贿罪数额和情节的分布情况

变量名称	变量赋值	全部样本		一区		二区		三区	
		Mean (N)	S.D. (%)						
受贿数额 (万元)	Ln(受贿数额+1)	4.04	1.40	3.94	1.422	3.88	1.39	4.13	1.39
自首	是=1	655	30.85	173	32.46	129	33.86	353	29.20
	否=0	1468	69.15	360	67.54	252	66.14	856	70.80
立功	是=1	124	5.84	32	6.00	34	8.92	58	4.80
	否=0	1999	94.16	501	94.00	347	91.08	1151	95.20
共犯	以非共犯为对照组	1820	85.73	446	83.68	328	86.09	1046	86.52
	不区分主从犯=1;其他=0	145	6.83	41	7.69	32	8.40	72	5.96
	主犯=1;其他=0	82	3.86	13	2.44	11	2.89	58	4.80
	从犯=1;其他=0	76	3.58	33	6.19	10	2.62	33	2.73

① 陈杰、郭晓欣:《城市外来劳动力市场上的农业户籍歧视程度研究》,《华东师范大学学报(哲学社会科学版)》2019年第5期。

② Cotton J., "On the Decomposition of Wage Differentials", *The Review of Economics and Statistics*, 1988, 70(2), pp. 236-243.

③ 刘海渤:《论法治社会中的司法自由裁量权》,《求是学刊》2004年第2期。

④ 王牧:《论刑罚概念:从“本质”到“意义”》,《当代法学》2018年第2期。

⑤ Neumark D., "Employers Discriminatory Behavior and the Estimation of Wage Discrimination", *The Journal of Human Resources*, 1988, 23(3), pp. 279-295.

⑥ 原始数据中还存在主刑被判处免于刑事处罚、拘役、无期徒刑和死刑的个体。为满足定量研究的需要,该部分个体不纳入研究,考虑到这部分个体人数占总样本人数的比重只有3.4%,故不会影响模型估算结果。

续表 1

变量名称	变量赋值	全部样本		一区		二区		三区	
		Mean (N)	S.D. (%)	Mean (N)	S.D. (%)	Mean (N)	S.D. (%)	Mean (N)	S.D. (%)
坦白或如实供述	是=1	1924	90.63	477	89.49	350	91.86	1097	90.74
	否=0	199	9.37	56	10.51	31	8.14	112	9.26
认罪认罚	是=1	1555	73.25	411	77.11	285	74.80	859	70.80
	否=0	568	26.75	122	22.89	96	25.20	350	29.20
认罪悔罪	是=1	557	26.24	115	21.58	89	23.36	353	29.21
	否=0	1566	73.76	418	78.42	292	76.64	856	70.79
退赃	是=1	1774	83.56	455	85.37	321	84.25	998	82.55
	否=0	349	16.44	78	14.63	60	15.75	211	17.45
索贿	是=1	353	16.63	76	14.26	63	16.54	214	17.70
	否=0	1770	83.37	457	85.74	318	83.46	995	82.30
其他较重情节	是=1	124	5.84	31	5.82	19	4.99	74	6.12
	否=0	1999	94.16	502	94.18	362	95.01	1138	93.88

四、实证分析

(一)各地区量刑结果的主要影响因素分析

为验证估算结果的稳健性,分次纳入影响因素检验法定因素与量刑结果的关联性。模型一仅纳入受贿数额进行单因素分析;模型二将受贿数额、自首、立功、共犯、坦白或如实供述、认罪认罚、认罪悔罪、退赃、索贿和其他较重情节都纳入回归模型进行多因素分析。

如表 2 所示,在同一区域内,模型一和模型二的估算结果较为稳健;在不同区域内,模型一和模型二估算结果的对比分析发现,2020 年中国受贿罪主刑量刑存在地区差异,体现在数额与主刑量刑结果之间的回归系数存在明显差异。

一区、二区和三区的 OLS 分析结果显示,受贿数额与主刑量刑结果之间均具有显著性关联。模型一在仅纳入受贿数额的情况下,受贿数额每增加 1%,一区、二区和三区的主刑量刑结果分别增加 0.01448、0.01567 和 0.01638 年。模型二在纳入受贿数额和情节因素之后,发现其回归系数均在 1% 的水平上显著,受贿数额每增加 1%,一区、二区和三区的主刑量刑结果分别增加 0.01487、0.01505 和 0.01641 年,模型估算结果比较稳健。通过对比发现,法官对单位受贿数额判罚结果由轻到重的排序为:一区<二区<三区,说明经济越发达的地方的法官对受贿数额的敏感度更低、容忍度更高。

表 2 各区域主刑量刑结果的 OLS 分析

变量名称	一区		二区		三区	
	模型一	模型二	模型一	模型二	模型一	模型二
受贿数额	1.448*** (0.061)	1.487*** (0.057)	1.567*** (0.082)	1.505*** (0.080)	1.638*** (0.040)	1.641*** (0.036)
其他变量	—	控制	—	控制	—	控制

续表 2

变量名称	一区		二区		三区	
	模型一	模型二	模型一	模型二	模型一	模型二
常数项	-1.933*** (0.202)	-1.313*** (0.316)	-2.477*** (0.277)	-1.410*** (0.457)	-2.726*** (0.147)	-1.713*** (0.207)
R ²	0.686	0.765	0.666	0.708	0.713	0.750
样本量	533	533	381	381	1209	1209

注:受篇幅所限,其他变量的分解结果未展示;*,**、***分别表示10%、5%、1%的水平上显著;括号内为稳健标准误。

(二)可能的内生性问题

内生性问题可能来源于遗漏变量所带来的偏差。虽然模型已经尽可能地纳入了受贿罪量刑所涉及的法定影响因素,但是依然存在不可观测的变量,例如,司法公开就可能同时对“受贿罪主刑量刑结果”和“受贿罪嫌疑人坦白或如实供述”产生影响。一方面,司法公开具有防止司法腐败的作用^①。司法透明度是司法公开的表现^②。吴雨豪基于刑罚社会学理论提出,在中国现有审判管理模式之下,司法透明度可能会影响法官量刑结果,司法透明意味着量刑决策被关注的可能性增大,有利于促使法官审慎裁判^③。另一方面,司法公开具有法治宣传教育功能^④。实务界认为职务犯罪嫌疑人具有特殊性,具有较高的文化知识水平,个人素质和能力较强,看待问题的方式较为理性,具备教育感化的基础^⑤。因此,司法公开有利于教育感化职务犯罪嫌疑人,从而顺利实现如实供述的目的。

为解决内生性问题,拟采用2020年31个省级纪委监委开展监督执纪“四种形态”^⑥的次数作为工具变量。第一,纪委监委开展“四种形态”的次数与职务犯罪嫌疑人坦白或如实供述具有强关联性。监督执纪“四种形态”是党的十八大以来反腐败斗争理论与实践创新及经验的集大成者,是党和国家的“反腐败基本政策”^⑦。习近平总书记强调:“精准运用‘四种形态’,让党员、干部切身感受到党的严管和厚爱”^⑧。实务界认为职务犯罪嫌疑人长期接受党的教育和熏陶,比较容易对讯问人员的教育感化形成认同,从而唤醒犯罪嫌疑人的正确价值观念,顺利实现犯罪嫌疑人如实自愿供述的目的^⑨。弱工具变量检验的结果也显示,“四种形态”的次数确实与嫌疑人是否选择坦白或如实供述存在显著关联性。第二,纪委监委监督执纪“四种形态”具有外生性。习近平总书记指出:“要有效防止腐败滋长,把反腐败防线前移,加强日常管理监督,精准运用‘四种形态’,抓早抓小、防微杜渐、层层设防。”^⑩在腐败治理方面,党的纪律工作有别于司法机关的审判工作,故有理由相信监督执纪“四种形态”外生于受贿罪主刑量刑结果。过度识别检验的分析结果也不拒绝“所有工具变量均外生”的原假设,故有理由相信所选择的工具变量满足外生性的前提条件。第三,运用工具变量法进行回归分析的前提条件是将职务犯罪嫌疑人坦白或如实供述作为内生选择变量。将各省级纪委监委开展监督执纪“四种形态”的次数作为工具变量分经济区域纳入对应的回归模型,Durbin-Wu-Hausman(DWH)检验的P值

① 陈柏峰:《习近平法治思想中的加强权利司法保护理论》,《法学家》2022年第5期。

② 郑智航、曹永海:《大数据在司法质量评估中的运用》,《吉林大学社会科学学报》2023年第2期。

③ 吴雨豪:《论集体量刑倾向对自由裁量权的塑造》,《法制与社会发展》2024年第1期。

④ 刘敏:《论司法公开的扩张与限制》,《法学评论》2001年第5期。

⑤ 王雄飞:《职务犯罪嫌疑人供述心理与讯问对策》,《人民检察》2016年第Z1期。

⑥ 《中国共产党纪律处分条例》第5条规定了“四种形态”的内涵。

⑦ 王军仁:《“风腐一体”视野下监察机关性质的再审视》,《中国法律评论》2023年第5期。

⑧ 习近平:《健全全面从严治党体系 推动新时代党的建设新的伟大工程向纵深发展》,《求是》2023年第12期。

⑨ 王雄飞:《职务犯罪嫌疑人供述心理与讯问对策》,《人民检察》2016年第Z1期。

⑩ 《习近平在中共中央政治局第四十次集体学习时强调 提高一体推进“三不腐”能力和水平 全面打赢反腐败斗争攻坚战持久战》, <https://china.huanqiu.com/article/48TOMR4v3it>, 访问日期:2024年3月11日。

均明显大于0.1,故不拒绝职务犯罪嫌疑人坦白或如实供述是外生变量的原假设(详见表3)。有理由相信,表3所显示的回归分析结果受到内生性问题的影响导致结果失真的可能性不大。

表3 工具变量的检验结果

经济区域	过度识别检验		弱工具变量检验		DWH检验	
	统计量	P值	统计量	P值	统计量	P值
一区	2.708	0.100	3.555	0.029	0.018	0.892
二区	0.186	0.666	3.186	0.043	2.261	0.134
三区	1.584	0.208	2.482	0.084	0.745	0.388

(三)Blinder-Oaxaca法分解我国受贿罪量刑地区差异的影响因素

如表4所示,一区比二区总体上多0.164个单位,说明一区比二区的量刑结果更重;可解释部分的贡献为0.030,说明一区比二区的罪行更为严重;不可解释部分的贡献为0.134,说明一区比二区的法官量刑更为严格。二区比三区总体上少0.423个单位,说明二区的量刑结果比三区轻;可解释部分的贡献为-0.462,说明二区比三区的受贿罪罪行更轻;不可解释部分的贡献为0.039,说明二区比三区的法官量刑更为严格。一区比三区总体上少0.259个单位,说明一区比三区量刑结果更轻;可解释部分的贡献为-0.411,说明一区比三区的罪行更轻;不可解释部分的贡献为0.152,说明一区比三区的法官量刑更为严格。

表4 分区之间主刑量刑总体差异比较

项目	一区-二区		二区-三区		一区-三区	
	贡献	贡献率	贡献	贡献率	贡献	贡献率
总体差异	0.164	100.00%	-0.423	100.00%	-0.259	100.00%
可解释部分	0.030	18.29%	-0.462	109.22%	-0.411	158.69%
不可解释部分	0.134	81.71%	0.039	-9.22%	0.152	-58.69%

如表5所示,在一区与二区之间,受贿数额在可解释部分和不可解释部分的贡献分别为0.087和-0.072,在5%的检验水准之下均不具有统计学意义;在二区与三区之间,受贿数额在不可解释部分的贡献为-0.534,在10%的检验水准之下具有统计学意义,表明二区与三区的法官就受贿数额的认定存在差异,致使两地的量刑结果扩大了1369.23%;在一区与三区之间,受贿数额在不可解释部分的贡献为-0.617,在5%的检验水准之下具有统计学意义,表明一区与三区的法官就受贿数额的认定存在差异,致使两地的量刑结果扩大405.92%。

表5 受贿数额对分区之间主刑量刑差异的贡献

分区对比	可解释部分		不可解释部分	
	贡献	贡献率	贡献	贡献率
一区-二区	0.087	290.00%	-0.072	-53.73%
二区-三区	-0.391***	84.63%	-0.534*	-1369.23%
一区-三区	-0.293**	71.29%	-0.617**	-405.92%

注:受篇幅所限,其他变量的分解结果未展示;*,**、***分别表示10%、5%、1%的水平上显著。

五、区差规律的提炼

通过对比分析2020年全国受贿罪公开判决书的相关资料,在回归分析的基础上使用Blinder-

Oaxaca分解法量化了法定因素(数额和情节)以及非法定因素(法官自由裁量)对量刑地区差异的贡献,从而验证量刑地区差异是否符合司法正义的要求。对比三组地区的分解结果,可以归纳出中国受贿罪量刑的区差规律:

结论一:各区域的主刑量刑结果均与其罪行的严重程度相对应。对比表4的研究结果发现:受贿罪罪行严重程度从重到轻排序为:三区>一区>二区;受贿罪主刑量刑结果由重到轻的排序为:三区>一区>二区。主刑量刑结果轻重程度的排序与罪行严重程度的排序相同,说明罪行越严重的地区主刑量刑结果也越重。

结论二:经济发展水平主要影响法官对受贿罪数额危害程度的认定。依据经济发展水平将31个省级行政区分为三个地区,由高到低的排序为:一区>二区>三区。对法官自由裁量所产生的贡献(不可解释部分的贡献)进行分析可以发现,对地区差异的影响显著且整体贡献最高的因素是受贿数额,说明法官对数额危害程度认定的差异是导致地区间自由裁量差异的主要原因。一区与二区之间,受贿数额所导致的不可解释部分的贡献为-0.072,表明一区法官对相同数额危害程度的认定比二区法官更为宽松。法官对数额危害程度认定的差异不能显著影响一区与二区之间主刑量刑结果的差异,可能是由于一区与二区之间的经济发展差异不大所致。二区与三区之间,受贿数额所导致的不可解释部分的贡献为-0.534,表明二区法官对相同数额危害程度的认定比三区法官更为宽松。一区与三区之间,受贿数额所导致的不可解释部分的贡献为-0.617,表明一区法官对相同数额危害程度的认定比三区法官更为宽松。据此可知,法官对犯罪数额危害程度的认定由宽到严的排序为:一区>二区>三区,与三个区域经济发展水平的排序相对比不难发现,经济越发达地区的法官对相同犯罪数额危害程度的认定越低。

结论三:法官自由裁量对受贿罪主刑量刑地区差异的影响是有限的。一区与二区主刑量刑总体差异为0.164,其中18.29%的差异是由数额和情节所导致的,81.71%的差异是由法官自由裁量所导致的,说明一区与二区的主刑量刑地区差异主要由法官自由裁量所导致的。尽管非法定因素对两地区主刑量刑差异的贡献率高达81.71%,但是一区与二区之间主刑量刑总体差异较小,可以认为法官自由裁量所导致的差异依然是有限的。二区与三区主刑量刑总体差异为-0.423,其中109.22%的差异是由数额和情节导致的,-9.22%的差异是由法官自由裁量所导致的,说明二区与三区的主刑量刑地区差异主要是由数额和情节所导致的。一区与三区主刑量刑总体差异为-0.259,其中158.69%的差异是由数额和情节导致的,-58.69%的差异是由法官自由裁量所导致的,说明一区与三区的主刑量刑地区差异主要由数额和情节导致的。

通过对比同一时期、不同地区的受贿罪法官自由裁量权与裁判结果的差异可以发现,从整体上看,中国司法裁判是公正的。结论一证实了罪行越严重的地区量刑结果越重,说明裁判结果符合“罚当其罪”的原则。这与“统一性”的宗旨相吻合,司法裁判做到了同等情况同等对待,即同罪同罚。结论二证实了从地区这一非法定因素与裁判结果的关系来看,法官的裁判符合地区经济发展水平这一客观实际。这与“差异性”的宗旨相吻合,法官自由裁量造成的量刑地区差异是符合地区经济发展水平的,具有合理性。法官自由裁量导致的量刑地区差异虽然是由法官主观导致的,但是有客观依据。结论三证实了除一区与二区这组对比之外,罪行严重程度造成的量刑地区差异高于法官自由裁量造成的量刑地区差异。这与“均衡性”的宗旨相吻合,即量刑地区差异主要是由地区罪行严重程度的差异所导致的,法官自由裁量对量刑地区差异的影响是有限的,也就是说司法裁判做到了在遵循统一性的前提下兼顾差异性。

司法公正区差规律背后的原因在于司法正义的差序性。司法正义的差序性是指司法正义在实现的过程中,其内含的各环节、各层次、各部分在时间上和空间上的优先性方面所体现出来的不平衡性。它要求在空间方面,应当从实际出发,考虑到不同地区社会经济的非同质性和不平衡性的特点,在当前不同地区的裁判尺度无法完全均等化的情况下,保持不同地区的案件在实现正义的要求和层次上

适度不同的空间秩序^①。司法裁判作为司法运行结果,既是法官适用法律的产物,也是司法与社会交互作用的结果。因此,司法裁判既具有绝对性和客观性又具有“时宜性”、相对性或主观性^②。差序公正和公平正义、平等原则并不矛盾,它不是要摒弃公平正义和平等原则,而是社会主义社会实现公平正义和社会平等的战略选择^③。中国幅员辽阔、地区之间经济社会发展不平衡,不同地区的案件之间既具有相似性又具有差异性。裁判对象的复杂性决定了司法裁判中普遍公正和差序公正不可偏废^④。因此,需要法官在维护裁判尺度统一性的同时包容案件之间的差异性,寻求一种有差序的公正。

六、相关建议

为防止司法公正因地区差异而受侵蚀,依据上述规律,提出几点建议。

第一,建立裁判辅助系统,为法官合理把握自由裁量权的使用幅度提供帮助。对国内不同经济区域的受贿罪主刑量刑结果进行分析之后发现,法官对于法律的适用基本符合“统一性”的原则。为巩固现有成果,增强地区之间裁判尺度的统一性,需要进一步加快推进人工智能技术与司法工作的深度融合。

经过十多年的探索,中国地方层面的司法机关在裁判辅助系统建设方面积累了一些经验:2006年山东法院系统研发了“刑法常用百种罪名电脑辅助量刑系统”;2011年福建检察院系统研发了“量刑建议管理系统”;2016年贵州法院系统研发了“大数据量刑系统”;2017年上海法院研发了“大数据审判辅助体系”。最高人民法院在2022年公布的《关于规范和加强人工智能司法应用的意见》中指出“到2025年,基本建成较为完备的司法人工智能技术应用体系,为司法为民、公正司法提供全方位智能辅助支持”。不同于在地方层面建立的裁判辅助系统,如果要搭建全国裁判辅助系统,需要考虑地区之间裁判尺度的差异性,因此,系统的搭建难度是可以预见的。在裁判辅助系统中运用人工智能进行量刑估算时,不仅需要考虑法官的集体经验,还需要充分考虑地区的特异性。

第二,按经济发展水平对犯罪数额进行折算,消减法官对数额认定的差异。自由裁量是法官心证的过程,裁量的过程不具有直观性。案情类似案件的裁判结果往往具有差异性,故而其裁判结果容易受到质疑。2012年最高人民法院在《关于在审判执行工作中切实规范自由裁量权行使保障法律统一适用的指导意见》的第三条中提出,“努力增强行使自由裁量权的确定性和可预测性”。自由裁量权的确定性和可预测性越强,法官招致的不合理怀疑会越少。结论二已经证实,法官受经济发展水平影响对相同数额案件的危害性程度认定不同,这会直接导致量刑地区差异。考虑到经济发展水平的相关指标是已知的和可量化的,可以对犯罪数额进行消减换算。

从消减的具体办法来看,以往研究给出了两种方式:王剑波认为由省级人民法院、省级人民检察院根据本地区经济发展状况和当地犯罪状况,在司法解释规定的数额幅度内,确定本地区执行的具体数额标准,报最高人民法院、最高人民检察院批准^⑤。陈磊认为可以借鉴国家赔偿法的立法方式,以省级行政区划为单位,采取“各省同一经济指标(城镇居民人均可支配收入或职工年平均工资)”乘以倍数的方式,每五年启动一次立法评估,在调研的基础上确定是否需要调整以及如何调整^⑥。从便捷性、明确性和可操作性的角度来考虑,陈磊的折算标准具有可能性。从操作流程来看,有两种选择:一

① 吴强玲:《社会正义及其在公共行政中的差序性实现》,《上海经济研究》2007年第3期。

② 江国华:《司法规律层次论》,《中国法学》2016年第1期。

③ 凌文豪:《差序平等:中国社会保障新理念》,《道德与文明》2012年第1期。

④ 罗诗钿:《论生态差序正义困境及其对空间生产正义的反思》,《学术界》2021年第7期。

⑤ 王剑波:《我国受贿罪裁判地区差异问题实证研究》,《中国法学》2016年第4期。

⑥ 陈磊:《贪污受贿犯罪量刑均衡问题实证研究》,《政法论坛》2020年第1期。

种是参照国家赔偿法的方式定期公布折算标准,法官在衡量犯罪数额的危害程度时,不使用货币单位而是时间单位;另一种是将已生效判决的量刑信息录入裁判辅助系统,由系统进行折算,提供折算后的量刑幅度参考。无论采取哪种方式,将犯罪数额按经济发展水平进行换算均能消减不同区域法官对犯罪数额危害程度的认定差异,从而提升法官自由裁量权的确定性和可预测性。

第三,建立数理模型监测法官自由裁量权的使用幅度。运用Blinder-Oaxaca分解法量化了法定因素和非法定因素对受贿罪量刑地区差异的贡献,证实了中国受贿罪量刑在遵循“统一性”的前提下兼顾了“差异性”。但从方法的运用上来看,还存在社会经济发展水平之外的其他环境因素影响法官自由裁量,现有模型没能将它们的影响效果从整体环境因素的影响效果中分离出来。为了便于监测,将来的研究有必要进一步探讨上述问题,逐步优化监测模型。

从研究结果来看,构建数理模型来监测司法公正具体可以分为四个步骤:(1)监测员在年末将国内全年所有的案件按照罪名分类,并对每一类案件逐一进行分析;(2)以受贿罪为例,将某一个省级行政区作为一个整体,其余所有省级行政区作为另一个整体,利用数理模型,得出量刑结果总差异、法定因素贡献和非法定因素贡献,每个省级行政区和每类案件均重复一样的操作;(3)对于量刑结果总差异较大,且非法定因素贡献过高的省级行政区进行调研,具体分析法官自由裁量所导致的量刑地区差异是否具有客观依据;(4)如果没有足够的依据,且非法定因素所产生的贡献过高,需要对法官自由裁量权的使用幅度进行相应的控制。

使用Blinder-Oaxaca分解法构建监测模型可以了解差异性的存在是否符合司法公正的要求,以及差异所呈现的趋势是否合理。如果Blinder-Oaxaca分解结果显示可解释部分的贡献是100%,则说明完全排除了法官的自由裁量,实现了地区间裁判结果的“同案同判”。然而,正如韩士彦和田宏杰所说:“很难想象,可以完全摒弃自由裁量而能够实现法律的正义性要求”^①。运用数理模型对量刑结果进行监测旨在将法官自由裁量权控制在合理范围内,而不是追求裁判结果的高度统一,这是推进司法公正需要遵循的重要准则。

第四,强化裁判文书书写规范,尽可能全面地展示法官自由裁量的依据。无论是运用裁判辅助系统来进行量刑参考,还是构建数理模型来监测法官量刑结果,都需要将裁判文书中的信息从文字形式转换为数字形式,受贿罪判决书撰写存在一些不规范之处导致法官裁判依据的提取难度较大,给监测法官自由裁量权的使用情况带来障碍,具体而言:(1)部分文书对受贿情节的表述不一致。例如,“初犯”一词,有的文书表述为“没有犯罪前科”或“未曾故意犯罪”;“认罪悔罪”一词,有的文书表述为“认罪态度诚恳、确有悔改之意”;“退赃”一词,有的文书表述为“退回违法所得”。(2)部分文书“本院认为”部分所引用的法条与“案情描述”部分所提及的情节并不完全对应。通过访谈一线办案法官得知,判决书中“本院认为”部分所引用的法条可能既包括对被告主张的回答,也可能包括对检方主张的回答;判决书中所引用的法条可能既包括法官认定某种情节成立的依据,也可能包括不成立的依据。(3)文书说理不足。量刑说理并非简单地重复事实和法律,而是具有很强的理论性与逻辑性,对维护量刑均衡与一致、实现量刑公正具有重要的现实意义^②。强化文书说理可以防止自由裁量权异化为司法专横,提高司法裁判的公信力和权威性^③。

① 韩士彦、田宏杰:《自由裁量权探析》,《社会科学研究》1996年第5期。

② 彭文华:《盗窃罪量刑规范化问题实证研究》,《华东政法大学学报》2021年第2期。

③ 张军:《法官的自由裁量权与司法正义》,《法律科学(西北政法大学学报)》2015年第4期。

Regional Disparities in Judicial Justice —A Blinder-Oaxaca Decomposition of Sentencing Differences

Jiang Guohua Sun Qiwei

(School of Law, Wuhan University, Wuhan 430000, P.R.China)

Abstract: Regional disparities in judicial decisions pose a risk of undermining judicial justice and have therefore always received significant attention from the Chinese judiciary. Taking criminal justice as an example, the Supreme People's Court of China has been implementing sentencing standardization reforms since 2008, emphasizing that this is a long-term endeavor. On March 8, 2023, the Supreme People's Court stated in its work report: "Maintain and promote social fairness; strengthen the unification of adjudication standards, and effectively prevent fairness and justice from being compromised due to regional, urban-rural, and industry differences." Despite years of effort and significant progress in sentencing standardization, there remains a risk of sentencing imbalances in China's judiciary. Existing studies have shown that regional sentencing disparities objectively exist in China. However, due to limitations in research methods, there has been no fact-based response to whether these disparities are within reasonable limits and what factors primarily cause them. This study uses publicly available verdicts of bribery cases nationwide from 2020 as samples and employs the Blinder-Oaxaca decomposition method to quantify the extent of sentencing disparities between regions with different levels of economic development, as well as the contribution of relevant influencing factors to these disparities. Regarding overall differences: the ranking of the severity of bribery offenses from most to least serious is Region 3 > Region 1 > Region 2; the ranking of sentencing outcomes from harshest to lightest is Region 3 > Region 1 > Region 2; the ranking of strictness in judges' application of the law from strictest to most lenient is Region 3 > Region 2 > Region 1. It shows that although judges in different regions do not have uniform standards of discretion, they overall adhere to the principle of proportionality between crime and punishment. In terms of group differences, the ranking of judges' discretionary power in determining the amount involved in the crime, from strict to lenient, is Region 3 > Region 2 and Region 3 > Region 1, consistent with the levels of regional economic development. This suggests that judges in more economically developed regions assess the social harmfulness of the crime amount as lower. Therefore, the following suggestions are put forward. Firstly, It is necessary to establish a judicial assistance system to provide assistance for judges to reasonably grasp the extent of their discretionary power. Secondly, the amount of crime should be converted based on the level of economic development to reduce the differences in judges' determination of the amount. Thirdly, It is imperative establish a mathematical model to monitor the extent to which judges exercise their discretion. Fourthly, It is imperative strengthen the writing standards of judicial documents and comprehensively demonstrate the basis for judges' discretion as much as possible.

Keywords: Judicial justice; Differential justice; Regional sentencing disparities; Blinder-Oaxaca decomposition; Regional disparity patterns

[责任编辑:王苏苏]