

量刑建议“一般应当采纳”条款之省思

刘磊*

目次

- 一、量刑建议的效力延展及高采纳率对程序公正的影响
二、“一般应当采纳”条款与罪刑相当原则的待量刑建议的效力
三、回归审检制衡与量刑公正：辩证理性对

摘要 依据现行《刑事诉讼法》第201条的规定，量刑建议的效力延展至审判阶段，约束庭审法官的量刑裁量权。由于量刑建议“一般应当采纳”条款的扩张适用，其对程序公正与量刑公正均产生了连锁影响。由于审查起诉阶段牺牲了被告人量刑听审与量刑辩论权利，加之检察官担当审判之前的量刑法官产生的各种反作用力，量刑建议“一般应当采纳”条款对刑事程序公正有一定的反向阻力。同时，由于检察官在审查起诉阶段难以尽职尽责进行量刑调查，更难以全面考量兼顾所有的量刑因素，确定刑的量刑建议很难完全体现刑罚公正。法院在个案中应当启动必要的庭审公诉审查，庭审阶段也应赋予法官充分的变更起诉罪名与量刑权限。设定科学的量刑方法、量刑情节的合理赋值等改革举措，有利于回归诉审制约与实现量刑公正。

关键词 量刑建议 法官保留原则 点的理论 幅的理论 量刑公正

根据我国现行《刑事诉讼法》第201条之规定，“对于认罪认罚从宽案件，人民法院依法作出判决时，一般应当采纳人民检察院的罪名和量刑建议。”由检察官司职追诉、法官司职审判的司法权力分立结构，既是当代刑事司法权力配置与运行的基本特征，也是有效保护人权的司法智慧。^{〔1〕}然而，如果由检察机关主导量刑结果，将量刑建议的高采纳率作为司法绩效的指征，量刑裁判的场所可能从正式的法庭庭审前移至检察官的密室。^{〔2〕}我国《刑事诉讼法》第201条的“一般应当采纳”规定有重新审视之必要，检察机关为实现司法绩效考核片面追求量刑建议的“高采纳率”可能衍生诸多的负面效应。如果过度提高量刑建议的采纳率，量刑建议的效力会自然延展至审判阶

* 同济大学法学院教授、法学博士。本文系中央高校基本科研业务费项目“司法供给侧改革背景下的认罪速裁程序”(项目编号: 07012150028)的阶段研究成果。

〔1〕 参见王皇玉:《认罪协商程序之法社会学考察》,载《台大法学论丛》2008年第4期,第81页。

〔2〕 Albert W. Alschuler, *The Trial Judge's Role in Plea Bargaining*, 76 Columbia Law Review 1059, 1063 (1976).

段。量刑建议的效力延展不仅仅会侵蚀法官的量刑裁量权,亦会影响刑罚公正的实现,甚至个案中有违罪刑均衡原则。认罪速裁程序背后的诉讼效率优先导向固然值得肯定,但其仍然应遵循正当程序原则与恪守罪刑相当原则,不应当以严重牺牲程序正义及刑罚公正为代价。

一、量刑建议的效力延展及高采纳率对程序公正的影响

随着2016年以来认罪认罚从宽制度改革的实践展开,认罪从宽程序的高适用率与量刑建议的高采纳率已成为我国当前刑事司法中的新常态。在诉讼效率导向下认罪速裁程序的扩张适用背景下,我国法院极少直接变更公诉机关的起诉罪名,对于公诉机关的量刑建议的采纳率甚至高达99%。^{〔3〕} 量刑建议的高采纳率,亦被检察机关作为重要的绩效考核指标。但是,过度强调诉讼效率与检察主导量刑,却易侵蚀、限缩法官在个案中量刑的自由裁量权,认罪认罚从宽的改革之路有误入歧途的风险。^{〔4〕}

(一) 对被告人量刑听审、量刑辩论权利的影响

与审判阶段集中审理、言词审理的听审模式不同,审查起诉阶段通常既无量刑听证程序,亦缺少充足的量刑情节调查过程,员额检察官在立法限定的办案期限内难以兼顾量刑公正。在审查起诉程序阶段,除讯问已被羁押的被告人外,检察官主要通过侦查机关移送的卷宗书面审查犯罪行为,目前尚无独立的量刑听证程序或庭前量刑尽职调查程序。^{〔5〕} 尤其是在复杂疑难的刑事案件中,量刑因子往往呈现复杂性与多样化,加之量刑背后的刑罚理念与刑事政策的易变性,仅仅通过书面化的卷宗审查,检察官难以系统、全面地权衡案件中的量刑情节与确立合理的预防刑期。而且,侦查卷宗内所记载的证据与犯罪情节内容,往往不利于犯罪嫌疑人、被告人,检察官也难以“兼听”侦查机关之外其他证人、当事人的合理诉求与客观评估各项酌定量刑情节的刑度赋值比例。刑事卷宗制度在实践中展现为诉讼卷与证据卷两类,卷宗记载的内容由侦查、监察机关所取得,卷宗是侦查终结后对犯罪事实、证据、程序的记载,量刑问题往往并非侦查、起诉阶段的聚焦点与重心,依赖卷宗确定的量刑建议难以全面反映案件的综合情况与量刑情节,办案检察官难以充分权衡诸多量刑情节并运用科学的量刑方法给出公正的量刑建议。而且,在审查起诉阶段,由于不存在集中的、以言词辩论审理为基本特征的庭前“听审”程序,辩护律师的量刑诉求难有实质性的影响力,检察官的办案重心与聚焦点往往集中在案件定性、犯罪事实、证据审查问题上,量刑建议是否达到罪刑相当、是否能够兼顾责任刑与预防刑的问题,往往只作为从属或辅助问题。审查起诉阶段缺乏公开听证与控辩平等协商程序的背景下,检察机关“精准量刑建议为主、幅度刑为辅”的

〔3〕 以2023年上半年为例,“已办理的审查起诉案件中,适用认罪认罚从宽制度审结人数占同期审结人数的90%以上;检察机关提出确定刑量刑建议已占量刑建议提出数的95%以上;对检察机关提出的量刑建议,法院采纳人数占同期提出量刑建议数的99%。”参见《2023年1至6月全国检察机关主要办案数据》,载最高人民法院网站2023年7月26日, https://www.spp.gov.cn/spp/xwfbh/wsfbt/202307/t20230726_622572.shtml#1。高达99%的量刑采纳率,印证了法院对量刑结果的实质影响力在日渐式微。

〔4〕 “我们在抱持很大的理想来做司法改革时,如果忽略社会层面、国家治理层面、制度来源以及文化背景时,尽管抱持很大理想做自认很好的改革,会发现成本过高而效果不如期待。”苏永钦等:《检察制度的新发展——以欧陆法系为中心》,载《检察新论》2010年第8期,第120页。

〔5〕 据实证调研,在审查起诉阶段,并无一个实质化的量刑平等协商机制。无论是确定刑抑或幅度刑,目前的量刑建议完全由检察机关主导,承担部分辩护职能的“值班律师”目前更似“见证人”的角色,很难发挥实质性的辩护职能,认罪协商的时间通常也仅有十分钟左右,被告人往往只能接受检察机关的量刑建议而难以协商,犯罪嫌疑人签署认罪具结书的过程难以保证程序公正。参见孙长永、田文军:《认罪认罚案件量刑建议机制实证研究——以A市两级法院适用认罪认罚从宽制度审结的案件为样本》,载《西南政法大学学报》2021年第5期,第5页。

改革尝试,既缺乏配套制度的支撑,亦难以全面权衡各种量刑因子而易导致刑罚失当。

由于第201条的“一般应当采纳”条款的扩张适用,认罪认罚从宽案件的量刑决定过程事实上被前移至审查起诉阶段,量刑辩护的重心亦因此随之转移至审查起诉阶段。量刑辩护的前移可能造成消极影响:(1)正式庭审中的量刑辩护效果遭受减损。如果辩护人在审查起诉阶段不能影响检察官的量刑拟定,在审判阶段再提出量刑辩护意见,除非庭审阶段出现新的足以影响判决结果的量刑事实,量刑辩护意见对一审量刑结果的影响力几近为零。(2)审查起诉阶段,辩护人量刑的协商空间受限。相较美国等国检察官,我国员额检察官的人均办案负荷较低,检察阶段“案多人少”的矛盾并不明显。当人均承办案件量较低时,通常检察官意愿与辩护人进行量刑协商的驱动力不足,辩护人向检察官提请量刑协商的议价空间有限,量刑结果往往由承办案件的检察官主导决定。(3)检察环节量刑“听证”程序的缺失或形式化。辩护律师在公诉阶段提交的书面量刑辩护意见,只是作为检察机关拟定量刑建议的参考,对量刑建议难有实质性的影响。审查起诉阶段,即使允许辩护人到场参加量刑听证,由于被告人及辩护人诉讼权利的不足,能够与检察官进行量刑协商的空间有限,检察阶段进行量刑辩护的实效不足。^{〔6〕}

(二) 量刑建议效力的延展与法官保留原则的冲突

“法官保留”原则,是指“将特定的公法上事项保留由法官行使,并且也仅法官始能行使的原则”。^{〔7〕}依据当代公法上的权力分立原理,对于公民生命、人身自由、财产、隐私等基本权利的剥夺与克减,应当遵守法律保留原则与法官保留原则。^{〔8〕}从二战后刑事司法发展的国际趋势观察,采用控诉原则与诉审分离原则是当代法治国家处理诉审关系的基石,量刑权则专属于法院,公诉机关只有量刑建议的权力而无量刑决定权。由于诉审机关在量刑权限上客观上存在互相消长的关系,扩张检察机关量刑裁量权往往会抑制法院量刑审理阶段的积极性,亦会抑制法官庭审中的量刑裁量权的积极行使。

1. 法院不采纳量刑建议需要承担额外的说理义务

我国现行《刑事诉讼法》第201条第1款规定人民法院有“一般应当采纳”量刑建议的义务,如果不予采纳,审判法官有义务释明其不采纳的法定事由。一方面,只有在案件定性错误、认罪意愿不真实、被告人否认犯罪事实、适用罪名分歧等五种情形下,人民法院才有权例外地不采纳量刑建议。^{〔9〕}另一方面,如果不属于立法明确列举的五种法定事由,人民法院审理后即使认定量刑建议畸重或畸轻,如果要变更检察机关的量刑建议,则必须遵循法定的变更程序,即先要通知检察机关对之前的量刑建议进行调整修正。只有在检察机关拒绝调整或调整后人民法院仍然认为“明显不当”时,人民法院才能变更量刑建议。我国现行《刑事诉讼法》第201条的立法目的为:对于认罪认

〔6〕 无论是从律师聘请率、有效辩护的实效、辩护对裁判的影响力观察,还是从律师的诉讼权利衡量,我国刑事辩护的实效与立法预期有一定的差距,在有效的实质性辩护不足的情形下,犯罪嫌疑人、辩护律师与检察机关的量刑协商空间有限,检察机关“从宽”宽宥的刑罚幅度至多只是参照刑法中的自首来处理,量刑从宽的刑度赋值不足,辩护律师与检察机关平等协商量刑幅度的制度保障不足。

〔7〕 林钰雄:《刑事诉讼法》(上),新学林出版公司2023年版,第98页。

〔8〕 我国《宪法》第128条规定:“中华人民共和国人民法院是国家的审判机关。”不论是被告人认罪的简化审理程序,还是被告人不认罪的普通审理程序,案件事实与证据的认定、罪名的适用、量刑等均应当由人民法院决定,量刑决定权由人民法院行使,检察机关的量刑建议性质上属于“求刑权”而非裁判权。检察机关在我国宪法上的定位是法律监督机关,而非裁判机关。

〔9〕 根据《刑事诉讼法》第201条第1款规定,在五种例外情形下,人民法院有权不采纳量刑建议,具体包括:(1)被告人的行为不构成犯罪或者不应当追究刑事责任的;(2)被告人违背意愿认罪认罚的;(3)被告人否认指控的犯罪事实的;(4)起诉指控的罪名与审理认定的罪名不一致的;(5)其他可能影响公正审判的情形。

罚从宽的案件,出于诉讼效率与认罪激励机制的考虑,立法机关对法院变更量刑建议的权力做出一定的立法限制。对于认罪认罚从宽的案件,法院的终局量刑决定权受到检察机关的制约。如果法官要变更量刑建议,需要承担举证责任并在判决书中详尽说理。

依据案件的轻罪与重罪类型,量刑建议可区分为确定刑与幅度刑。对于三年以下有期徒刑的认罪认罚从宽案件,检察机关通常倾向于提出确定刑类型的量刑建议。检察机关确定刑量刑建议的采纳率越高,法院量刑的自由裁量权则越低。认罪简化审理的案件中,法院如果变更量刑,需要详尽阐释量刑建议“明显不当”的判由。对于认罪简化审理的轻罪案件,一审法院即使发现检察机关的量刑建议偏轻或偏重,法院因受《刑事诉讼法》第201条的立法掣肘而难以变更量刑建议。

由此可见,随着认罪认罚从宽制度的推行与实践,检察机关量刑建议的刚性效力已延展至审判阶段,具有了一定的“准终局性”的刚性效力,^[10]法院的量刑裁量权受到一定的立法限制。^[11]

2. 量刑建议的效力延展对法官量刑裁量权的限缩

检察机关的求刑权效力扩张与法院行使量刑裁量权,系此消彼长的关系。目前,我国认罪认罚从宽案件已高达90%以上,量刑建议的采纳率事实上已高达90%以上,这意味着我国检察机关现实中可能主导了80%以上的刑事案件的最终量刑结果。^[12]如前所述,除非符合五种特定的情形,人民法院审理案件后,认定量刑建议无论偏重还是偏轻,均不得随意变更检察机关的确定刑抑或幅度刑范围。只有当认定量刑建议属于“明显不当”的畸重畸轻情形,而且还要遵守严格的变更程序通知检察机关调整修正,才能予以变更。固然,限制法院的量刑变更权,有利于检察机关求刑权的效力提升及认罪激励机制的实现,但其产生的负面效应不可忽视。诉审之间即使存在量刑分歧,只要量刑建议未达到“明显不当”,出于诉讼效率与认罪激励机制的考量,一审法院仍然应当采纳量刑建议。^[13]除非出现适用主刑错误、法定刑适用的幅度严重失当、主要量刑情节不属实、立功等新的对量刑有重大影响的事实、基准刑计算失当等严重的量刑错误,法院才能依照第201条第2款之规定,通知检察机关变更之前的量刑建议。检察机关调整修正之前的量刑建议后,即使仍然与法院设定的量刑结果有差

[10] 参见胡铭:《认罪认罚案件中的量刑协商和量刑建议》,载《当代法学》2022年第2期。

[11] 毋庸讳言,目前认罪认罚从宽程序实践中呈现一定的检察官司法色彩、检察官成为“法官之前的法官”,量刑程序从之前的法官主导模式演变为诉审协同模式。参见孙长永:《中国检察官司法的特点和风险——基于认罪认罚从宽制度的观察与思考》,载《法学评论》2022年第4期,第73页。

[12] 我国2018年正式推行认罪认罚制度之后,虽然存在一定的地域差异,总体上各地法院量刑建议的采纳率高达90%以上。例如,有学者统计了112万余件刑事裁判文书,法院明确不采纳量刑建议的案件仅有552件。参见周长军:《认罪认罚案件中法院变更量刑建议的法理分析》,载《云南社会科学》2022年第2期。以某市检察机关2022年上半年数据为例,认罪认罚从宽案件的适用比例高达91.5%,约80%的案件提起公诉后,量刑建议的采纳率高达97%。2022年上半年,该市检察机关刑事犯罪案件适用认罪认罚从宽制度提出量刑建议合计1054人,其中,提出幅度型量刑建议47人,占比4.46%;提出确定型量刑建议1007人,占比95.54%。法院采纳检察机关适用认罪认罚从宽制度量刑建议合计1033人,量刑建议采纳率为98.01%,同比下降了0.12%。其中,采纳确定型量刑建议986人,确定型量刑建议采纳率为97.91%。参见《关于2022年上半年全市认罪认罚工作情况的通报》,载安庆市人民检察院网站2022年7月11日,http://www.anqing.jcy.gov.cn/jwggk/gzbg/202207/t20220711_3745464.shtml。

[13] 有学者主张,不仅审前阶段要检察主导,审判阶段亦要“控审交融”,一是检察机关借助量刑辅助软件系统科学量刑,二是审检在开庭前进行量刑沟通,三是由省级检察机关统一全省量刑标准,量刑公正与量刑科学均能实现。即在认罪认罚从宽程序中,主张诉审机关从现行的分工制约模式转向合作、协作量刑模式。该论点不仅与我国《刑事诉讼法》第7条“分工制约”的相关条款抵触,对审判机关在庭前参与检察机关量刑建议确定的司法风险也未进行必要的评估与实证研判。审查起诉阶段与庭前会议阶段即提前确立宣告刑,开庭前诉审机关的“协同量刑”将使正式开庭中量刑程序流于形式。参见闵丰锦:《检察主导抑或审判中心:认罪认罚从宽制度中的权力冲突与交融》,载《法学家》2020年第5期,第161—162页。

异,只要量刑分歧的刑度不超过30%,不属于量刑的“明显不当”,法院仍然应当依据第201条第1款规定予以采纳。易言之,无论是提起公诉时的量刑建议抑或开庭后调整修正后的量刑建议,只要诉审机关的量刑分歧不超过一定的幅度,法院必须采纳。正是由于立法机关严格限制了认罪认罚案件中法官的量刑裁量权,所以检察机关量刑建议的采纳率已高达90%以上,高采纳率的出现并不意外。

(三) 抑制法官庭审中证据调查的积极性

根据现行《刑事诉讼法》第201条第1款的规定,只有当被告人确属无罪或免除刑事责任、认罪不自愿、否认有罪、诉审之间罪名分歧、影响公正审判五种情形下,法院才有权不采纳量刑建议。假定以30%的量刑偏差幅度来设定“明显不当”的认定标准,即使法院发现新的可能影响量刑结果的量刑事实,只要该事实对量刑结果的影响比例不超过30%,该量刑事实的审理则无必要,法官通常也不在判决书中予以阐明或给予量刑说理论证。在开庭审理阶段,如果案件主要事实或量刑情节发生重大变化,公诉机关的定罪证据尚未达到“排除合理怀疑”标准,需要重新鉴定或证据调查,其实意味着之前量刑建议的采纳基础已发生动摇。不过,在现实中,由于检察机关将“采纳率”作为年度绩效考核的标准之一,一旦法院不采纳量刑建议,则可能会引发检察机关的抗诉或其他类型冲突。所以,法院可能出于避免诉审机关的冲突而放弃对事实的严格审查,即在被告人认罪的情形下,仅仅依据卷宗进行形式上的审查,导致认罪认罚从宽案件的审理过程流于形式。例如,法院在开庭审理后发现自首、立功、共犯作用等量刑事实认定错误等影响量刑结果的情形,但未偏离量刑建议的30%幅度,如果检察机关拒绝修正,法院难以认定为“明显不当”而只能采纳检察机关的量刑建议。如果量刑事实发生了变化,通常会引起起刑点与基准刑的变化。即使变化造成的量刑影响幅度不超过30%,如果因此剥夺法官的量刑裁量权,将会影响量刑结果与刑罚公正的实现。在认罪速裁程序中,为维持量刑建议的效力稳定,牺牲法官的自由裁量权,不仅会影响刑罚公正与刑罚均衡的实现,也会导致法官开庭后对新的量刑事实难以裁判而选择沉默或刻意回避。随着认罪认罚从宽制度的实践展开,对于不属于“明显不当”的量刑情节,法官往往会避免在判决书增列起诉书之外的新的量刑情节,以尽量维护检察机关的量刑建议。

(四) 对庭审实质化的消极影响

依照认罪从宽程序审理的案件,由于控辩审三方均聚焦于“从宽”的幅度上,事实的审理、有罪证据的审查判断、程序性辩护及罪名争议反而处于边缘化地位,实际的开庭审理过程则流于形式,“重量刑、轻审理”的庭审模式造成“审判中心主义”改革目标的落空。认罪认罚从宽程序的高比例适用成为常态化,造成普通程序与速裁案件的比例呈现结构性失衡,绝大多数刑事案件的庭审过程以简化、快速方式结束,“实体从宽、程序从简”模式下,法院的开庭欠缺实质化的审理过程。

“审判中心主义”则要求审理程序的实质化,尤其是第一审审理程序被立法机关设定为刑事诉讼程序的重心。在审判中心主义的诉讼模式下,事实与证据的调查认定、罪名适用、量刑及程序争议均应当通过实质化的开庭审理程序裁判。当前,认罪认罚从宽制度的运行呈现“双高一快”的特征,即认罪从宽案件的高适用率、量刑建议高采纳率、审理程序的快速化。在被告人签署认罪具结书提起公诉后,案件的开庭审理程序相对简化,三年以下徒刑的轻罪案件甚至可以省略法庭调查与法庭辩论程序。^[14] 在法

[14] 参见《刑事诉讼法》第224条。认罪认罚从宽案件以简化审理方式终结,其适用率与言词、集中审理原则之间,存在此消彼长关系。如果绝大多数被告人认罪而放弃普通审理程序,通过简化审理快速审结案件,实质化审理方式反而成为例外(只有不到10%的案件通过普通审判程序审理)。即使是不以认罪速裁程序审理的案件,由于一审对质、辩论及超低的证人出庭率等原因,现实的庭审与2016年认罪速裁改革之前的庭审在开庭审理时长、对质辩护模式、有罪判决率等表征上并无明显差异。

庭调查与法庭辩论被过度简化的背景下,第一审庭审程序中的法官较难对事实争议问题进行实质性审理,人民法院往往以卷宗为基础进行裁判,这事实上形成了检察机关主导认罪速裁案件审理结果的机制。对于合议庭依照普通程序审理的认罪认罚从宽案件,由于卷宗并送制度、辩护人地位的相对弱势、证人出庭率低等原因,“实体从宽、程序从简”成为人民法院审理认罪从宽案件的基本模式。除非出现被告人认罪不真实、不自愿等特殊情形,量刑幅度事实上成为开庭审理阶段的主要聚焦点与重心,证据调查与法庭辩论程序的价值因认罪认罚从宽的推行而减损。在“程序从简、量刑从宽”的诉讼模式影响下,在现实中,以言词辩论、集中审理为特征的“审判中心主义”反而沦为例外,检察主导认罪认罚从宽案件的结果则成为原则并常态化。绝大多数的刑事案件将缺乏实质化的审理过程,对于依照认罪速裁程序审理的案件,法院将审理的重心从事实、证据、罪名、程序争议转向量刑问题,开庭审理程序虽快速但易流于形式,法官极少对公诉进行实质性审查,“审判中心主义”的改革目标甚至可能落空。

二、“一般应当采纳”条款与罪刑相当原则的冲突

量刑均衡是刑事审判的目标之一,法官需要依据法定刑、基准刑、酌定情节等诸多权重因素行使量刑裁量权,科刑轻重得宜、罪当其罪。刑事审判既要实现罚当其罪,又要兼顾刑罚的公平、均衡性,还须考量一般预防与特殊预防等刑事政策,刑罚学理论、量刑科学与量刑方法程序相当复杂,检察官的功能定位是审查起诉、出庭公诉,不宜担任量刑官。作为量刑公正的守护者,法官不应当盲从检察官的求刑权。从刑罚学原理分析,“一般应当采纳”条款与实现量刑公正之间存有诸多矛盾,背离了罪刑相当原则,而且反噬了法官的量刑裁量权。

(一) 审检之间的量刑“算法”分歧:常态化与必然性

在刑法教义学上,不仅罪名的认定、犯罪竞合时罪数形态认定、量刑情节的权重比例、量刑方法等问题争议不断,责任刑与预防刑的优先顺位亦难有定论。在审查起诉阶段,检察机关的重心与聚焦点往往集中在事实、证据与罪名认定问题,个案中的量刑公正、均衡问题往往只作为从属问题。在重大的争议案件中,不仅在诉审机关之间会产生量刑争议,而且不同检察官、法官,对同一案件的罪名认定、罪数形态与量刑算法均可能存在重大分歧。

1. 不同量刑理念所必然导致的量刑结果差异

责任刑与预防刑是当代国家的主要量刑基准,前者将责任作为量刑的基础,而预防刑则是将预防犯罪作为量刑基准。但是,责任刑与预防刑之间往往会产生一定的冲突,甚至会出现刑罚目的的的二律背反。^[15] 即使是以责任刑为基础,刑法学术界亦有“积极的责任主义”与“消极的责任主义”之争,前者主张存在责任必须处罚,后者主张即使存在犯罪仍可减轻、免除处罚。^[16] 而消极的责任主义又可区分为犯罪论“归责上的消极责任主义”与“量刑中的消极责任主义”,前者是指定罪阶段中没有责任的行为不得认定为犯罪,后者是指量刑阶段中严格依照责任判定刑罚,刑罚的量不得超出行为人应负责任的量。^[17] 同时,对于想象竞合、牵连犯、结果加重犯、转化犯、结合犯、事后不可罚行为等诸多行为的犯罪竞合问题,刑法教义学上难有明确清晰的规则,检察官是否能够

[15] 例如在对刑罚目的的理解上,日本刑法学界关于报应刑与目的刑顺位的学派之争未有定论,在现实判例中,往往是责任主义与预防主义的折中。参见[日]高桥则夫:《刑法总论》,中国政法大学出版社2020年版,第480页。

[16] 参见[日]城下裕二:《量刑理论的现代课题》,黎其武、赵珊珊译,法律出版社2016年版,第266页。

[17] 李冠煜:《量刑基准的研究——以责任和预防的关系为中心》,中国社会科学出版社2014年版,第101页。

避免重复评价,正确区分一罪与数罪,司法实践中颇具争议。^[18] 在行为单复数难有绝对区分标准与定论的前提下,不同检察官之间、法官与检察官之间、不同法官之间,量刑算法与量刑结果将会产生各种分歧,要求法官庭审中原则上采纳检察机关确定刑的量刑建议,有违现代刑罚学原理与量刑科学。

由于不同检察官对于刑罚目的、罪数形态、归责标准、量刑方法、量刑基准等诸多问题存在理念、经验与方法上的差异,不同案件不同检察官会出现“量刑相场”上的巨大差异,导致争议案件中量刑建议的公正性易受到公众与媒体的质疑。^[19] 在检察系统内,即使是通过案例指导制度、检察委员会决议制度、科室案例会诊讨论制度甚至控审协商量刑标准等,仍然难以完全实现个案量刑的客观、公正。与其片面追求司法的数字化绩效,盲目提高量刑建议的高采纳率,不如扩张法官的量刑自由裁量权,通过正式的庭审与独立的量刑听证程序由法官司职量刑。开庭审理程序采用控辩辩论的方式,更宜于法官通过权衡个案中的犯罪事实情况、被告人人身危险性、法定与酌定的量刑情节、责任刑与预防刑的兼顾等诸多刑罚要素,做出罪刑相当的公正量刑判决。检察官行使求刑权,法官对量刑建议进行司法审查并行使终局量刑权,是当代刑事司法的基本特征之一,还能够体现法官的量刑理念与自由裁量权,亦是审检制衡原则的体现。

2. 量刑理论中“点的理论”与“幅的理论”之争所导致的量刑分歧

即使不考虑预防刑,在责任刑思维下,量刑仍然有“点的理论”与“幅的理论”之争。^[20] 量刑幅的理论亦称为裁量空间理论,认为存在一个与罪责相适应的幅度空间,法官最终适用的刑罚,只可能是一个幅度,在幅度刑范围内,再考量预防刑。^[21] 而点的理论则认为,法官的量刑应当是一个精准的量刑点,刑罚与责任之间存在着客观的一一对应关系,法官量刑不能偏离点的范围。^[22] 如果采用幅度理论,检察官的量刑建议亦应当为一个具体的量刑幅度;如果采用量刑点的理论,检察官的量刑建议只能以精准量刑方式提起。在司法实务中,检察官的量刑建议采用双轨制,在三年以上的重罪案件中的量刑建议表现为一个量刑幅度,不精确定到月;对拘役及三年以下有期徒刑

[18] 例如轰动一时的深圳“快播”公司涉嫌传播淫秽物品牟利案,检察机关以情节加重犯起诉快播公司及五名被告人“牟利”,量刑建议为十年以上有期徒刑,北京海淀区法院认为不构成“牟利”,判处五名被告人三年不等有期徒刑。再例如,天津赵春华非法持有枪支案,对于不能击穿人体皮肤动能极低的气枪能否视为刑法中的“枪支”,辩护人与公诉人法庭上持不同的意见。天津市中级人民法院二审量刑改判缓刑,天津河北区检察院的三年以上的量刑建议明显存在错误。一审法院如果盲目因循量刑建议,有可能导致个案中的罪刑关系失衡。犯罪行为事实的变化、重大量刑事实的变化、证据存疑、犯罪竞合等诸多情形,均会导致量刑幅度的重大调整,如果法院盲目迎合检察机关的量刑建议,只能是以牺牲量刑公正为代价。对于轻罪案件,司法机关则易形成“小案小办、简案简办”之思维,忽视轻罪案件的裁判质量。“繁简分流”的前提要件之一是保持案件基本的程序公正与量刑公正,如果为了实现案件分流片面追求诉讼效率,法院对量刑建议的审查流于形式,轻罪刑事案件的裁判品质无疑会受影响。

[19] “量刑相场”概念来源于日本刑法学界,意指法院量刑的行情表或司法实证上的量刑倾向,即某一法院对特定罪行的量刑相场在一定的周期内能够根据判例或量刑规范进行预测。参见李冠煜:《量刑基准的研究——以责任和预防的关系为中心》,中国社会科学出版社2014年版,第101页。

[20] Andrew von Hirsch, *Proportionality in the Philosophy of Punishment: From Why Punish to How Much*, 25 *Israel Law Review* 549 (1991).

[21] 参见[德]汉斯-约格·阿尔布莱希特:《重罪量刑:关于刑量确立与刑量阐释的比较性理论与实证研究》,熊琦等译,法律出版社2017年版,第42页。

[22] 参见张明楷:《责任主义与量刑原理——以点的理论为中心》,载《法学研究》2010年第4期,第134页。我国刑法典采用并合主义刑罚观,兼顾责任刑与预防刑,有期徒刑以上案件,我国刑罚立法采用幅度刑。犯罪行为轻重与刑罚轻重很难精准,客观上可能不存在一个精确的量刑点,归责与量刑中的量刑极难量化,点的理论追求精准的数字化、量化的量刑结果,往往是不可能的任务。

刑的案件,则倾向于提起精准量刑建议,刑罚精确到月。^[23]从责任与刑罚关系而论,“点的理论”假定任何犯罪行为的刑事责任均能对应一个精准的量刑点,实质上是法官确定刑罚的过程等同于物理世界的规律探寻,具有一定的哲学先验论的色彩,既不符合当代刑事司法的真相与实践,亦违反我国刑法典的立法精神及刑事政策。

受“点的理论”思维定式影响,检察机关审查起诉后所设定的精准量刑点,可能背离刑罚公正。对于三年以上有期徒刑的重罪案件,刑法典大多规定了较大的刑罚幅度,赋予法院较大的自由裁量权限,检察机关提起精准量刑建议或将量刑建议幅度设定在一年以内并要求法院原则上采纳,实质上是在个案中限缩了法院的量刑裁量权,否定了幅度刑理论。

3. 多项量刑因素合理赋值与量刑算法透明、公正的实现困境

刑罚公正、平等与均衡的实现,涉及诸多的要素,即使通过独立的量刑听证程序,法官开庭审理后仍然难以绝对地实现刑罚正义。每年数百万计的刑事案件中,量刑因素呈现出相当的多元性与复杂化:一是犯罪行为自身因素,二是犯罪行为人的人身危险性要素,三是被害人因素,四是特殊预防与一般预防的刑事政策,五是各项法律未明文列举的法外要素。^[24]我国法院实际的量刑权重因素相当复杂多元,既要兼顾报应刑与预防刑,又要兼顾诸项犯罪事实、犯罪行为情节,亦要考量刑事政策,还要力图在一定的时空下保持同案同判。

如果认罪认罚从宽案件量刑建议以“精准刑为主、幅度刑为辅”方式提起,法官则“一般应当采纳”量刑建议,势必将由检察官承担量刑官角色。但是,即使不考虑案件事实、证据、程序与罪名认定上的争议,检察官难以在有限的审查起诉期限内启动完成全面的量刑调查程序,难以厘清所有的量刑要素并最终确定一个非常精准科学的基准刑。由于检察机关未能充分、全面、系统地考虑犯罪事实情节、人身危险性、法外因素、刑事政策等诸多权重因素,导致量刑建议违背罪刑相当原则的情形并不鲜见。例如,天津赵某某非法持有枪支案、吉林白城黄某某私建浮桥案、深圳王某贩卖野生鸚鵡案等诸多案件中,检察机关对于被告人认罪认罚的案件并未能完全评估量刑因素的多元化与复杂性,法院盲目地高比例采纳量刑建议反而会导致社会与公众的质疑。

(二) 检察官难以胜任“准量刑官”的职能

除非检察官精准量刑建议的公正性毋庸置疑,或者不逊色于法官的科刑公正,“一般应当采纳”条款才具有正当性。但是,理想与期待毕竟不等同于现实。从审查起诉阶段量刑建议的实际确立方法观察,即使是有长期司法经验的员额检察官,要在有限的办案期限内将量刑作为重心,在不同类型的案件中客观、全面评估犯罪情节以合理设定刑度,确定不同案件中责任刑与预防刑的顺位与权重比例,向法院提出公正、精准的罪刑相当的量刑建议,难度系数相当高,甚至个案中会

^[23] 《人民检察院刑事诉讼规则(2019)》第275条规定,犯罪嫌疑人认罪认罚的,人民检察院应当就主刑、附加刑、是否适用缓刑等提出量刑建议。量刑建议一般应当为确定刑。对于新类型、不常见犯罪案件,量刑情节复杂的重罪案件等,也可以提出幅度刑量刑建议。检察机关倾向于扩张确定刑量刑建议的范围,认罪认罚从宽案件由检察机关主导量刑结果。随着认罪认罚从宽制度改革的展开与实践,在我国法学学术界,支持精准量刑的主张似乎目前占主导地位。参见陈国庆:《量刑建议的若干问题》,载《中国刑事法杂志》2019年第5期,第8—10页;陕西省人民检察院课题组:《认罪认罚案件量刑建议精准化》,载《法律科学》2021年第3期,第142—152页。

^[24] 依据最高人民法院、最高人民检察院2021年《关于常见犯罪的量刑指导意见(试行)》第1条量刑“指导原则”之规定,量刑应当根据:“(一)犯罪的事实、性质、情节和对于社会的危害程度,决定判处的刑罚。(二)量刑既要考虑被告人所犯罪行的轻重,又要考虑被告人应负刑事责任的大小,做到罪责刑相适应,实现惩罚和预防犯罪的目的。(三)量刑应当贯彻宽严相济的刑事政策,做到该宽则宽,当严则严,宽严相济,罚当其罪,确保裁判政治效果、法律效果和社会效果的统一。(四)量刑要客观、全面把握不同时期不同地区的经济社会发展和治安形势的变化,确保刑法任务的实现;对于同一地区同一时期案情相似的案件,所判处的刑罚应当基本均衡。”

出现罪刑关系的失衡。^[25]

重罪案件的量刑幅度与裁量空间较宽,要实现刑罚公正,往往需要在庭前完成公正、中立的量刑调查程序,做出相对独立的量刑调查报告。首先,检察官担负追诉职能,以追诉犯罪为首要任务,对于有利于被告人的从轻量刑情节,往往缺少关注与尽职调查。审查起诉阶段,以审查事实、证据是否达到定罪标准以及罪名认定为主要任务,量刑问题在公诉环节只是作为次要问题存在。其次,即使检察官力图对各项量刑情节进行尽职调查,各项量刑情节的权重比例仍难以确定,个案中“责任刑—预防刑”的优先顺序存在不确定性。采用“点的理论”还是“幅的理论”,应当根据个案的具体犯罪情节、犯罪人犯罪前后情况、被害人和解状况等诸多情节综合全案情况来确定,相较不进行量刑公开听审的检察官,法官的开庭审理更具有正当程序机制上的优势。最后,即使检察机关委托第三方机构进行犯罪调查评估,例如调查被告人的社区品格、矫正可能性、犯罪成因、刑事和解状况等,调查评估的公正客观性仍需要由法官进行司法审查。无论是缓刑考察官还是委托量刑第三方评估机构进行量刑调查,法院作为中立的审判机关,相较检察机关更具优势,由法院委托更体现量刑调查的相对中立性。

其实,“简案”未必真简,一些轻罪案件中的量刑因素仍然可能相当复杂,个案中量刑方法、参数、相場及预防刑的确定与演算仍然呈现严重的分歧。即使是三年以下的轻罪刑事案件,个案间的量刑幅度仍然有相当的差异,检察官要在有限的审查起诉期限内达到接近于或优于法官的量刑公正,几乎是不可能的任务。量刑的因素非常复杂、多元化,例如是否适用缓刑、如何评估社区矫正的效能、被告人的人身危险性与社会危险性、被害人利益的保护以及其与被告人之间的关系修复程度等等,检察官通常要在非常短的诉讼期间内评估诸多量刑因素,几乎是难以实现的目标。司法机关如果将认罪认罚的案件视为“简案”而采取简办的思维,审查起诉阶段通常会演变为一个“小案小办、快办”的过程,提起的量刑建议往往难以充分地体现刑罚公正。轻罪案件仍然关系社会正义、影响刑罚公正以及实现轻罪有效治理,85%以上的轻罪案件呈现出一定“长尾效应”,但并非意味着所有的轻罪案件的起诉、审理、定罪量刑均“简案简办”。^[26]认罪认罚从宽制度的确立,主要目的是实现案件“繁简分流”与“宽严相济”的刑事政策。目前,三年以下徒刑的轻罪案件占比已高达85%以上,通过认罪速裁程序快速起诉、审理轻罪刑事案件,有利于案件的繁简分流进而提

^[25] 2021年最高人民法院《关于常见犯罪的量刑指导意见》规定:“……(人民法院)根据犯罪的事实、性质、情节和对于社会的危害程度,决定判处的刑罚。量刑既要考虑被告人所犯罪行的轻重,又要考虑被告人应负刑事责任的大小,做到罪责刑相适应,实现惩罚和预防犯罪的目的。”可见,我国法官既要考量犯罪行为因素,亦要考量犯罪人的人身危险性与社会危险性,还要兼顾宽严相济的刑事政策。仅以犯罪行为因素而论,包括行为人犯罪年龄、犯罪动机目的、是否激情犯罪、犯罪手段、犯罪时间地点、犯罪后态度等几十项法定、酌定的情形情节。量刑基准的确定、量刑方法或算法的运用、预防刑的权重比例大小等诸多复杂的要素,检察官在有限的审查起诉期限内,进行职权调查与查明所有量刑情节并遵循量刑程序,几乎是不可能的任务。在目前卷宗化、书面审查化、封闭化的审查起诉机制下,诉讼效率的过度偏重,必定会影响量刑公正的实现。

^[26] “长尾理论”(long Tail Theory)原属于经济学上的概念,是指在商业销售中,占比80%以上的“长尾商品”反而利润有限,精品商品的数量虽少但利润率却远超过绝大多数的长尾商品。表面看来,量大的商品价值利润远不如精品商品。经济学者提醒,在重视热门商品的销售时,对特定的长尾商品也要予以关注,现在一时冷门的商品将来可能成为热门的商品。我国轻罪案件与重罪案件的办案成本配置,目前呈现一定的长尾效应,假定司法机关将80%的精力投入到重罪案件,“小案小办、简办”思维下,检察机关可能投入20%的司法资源办理80%的轻罪案件。长尾理论的相关论述,可参见[美]克里斯·安德森:《长尾理论》,乔江涛、石晓燕译,中信出版社2019年版,第55—61页。在刑事司法中,案件的真正价值与社会影响力不能仅仅根据轻、重罪来判定,而仍然应当是一案逐案判断,一些轻罪案件量刑公正仍然会关系刑罚公正的底线。例如著名的天津赵春华非法持有枪支案、北京余金平交通肇事案、深圳王某贩卖珍稀野生鸚鵡案等等,即使是轻罪案件,检察官不当的量刑建议,仍然会引发公众对刑罚公正的质疑。

升司法效率。与三年有期徒刑以上的重罪刑事案件有所不同,被告人认罪的轻罪刑事司法程序以诉讼效率的实现为优先目标,赋予司法机关充分的量刑裁量权与控辩协商量刑空间。

(三) 检察机关的“从宽”幅度受制于传统“责任刑—预防刑”二元刑罚体系

无论是德国、日本的刑罚理论,还是从我国刑法典及量刑基准观察,“责任刑—预防刑”的二元刑罚论,目前仍然是传统刑罚理论的核心内容与思维定式。如果一方面维持传统的刑罚理念,苛求法官、检察官对所有认罪速裁案件做到罪刑相当与罚当其罪,量刑建议只能采用幅度刑而不能盲目采用确定刑的方式;另一方面,由于认罪速裁制度的施行,被告人及辩护人往往期待检察官的量刑建议能够精准到月,被告方试图最大限度地期待量刑减轻、从轻的可预测性,精准刑的量刑建议亦能提前让被告人有合理的量刑结果预期。不过,量刑结果固然重要,但更值得关注的应当是量刑过程中的“算法”,是刑度减让的比例决定量刑结果而非相反。对认罪行为与诉讼契约提升量刑增值的比例,扩张检察机关的量刑裁量权才是问题的关键。对被告人认罪的轻罪案件,由控辩双方平等协商量刑,超越传统的二元刑罚观,引入认罪契约刑的刑罚理念并适度提高认罪契约行为的量刑赋值,控辩双方达成的量刑合意才能充分体现被告人与国家刑罚权之间的沟通正义。相反,检察机关以传统“责任刑—预防刑”二元刑罚观所确定的量刑建议,由于对认罪行为的量刑赋值过低,难以充分体现量刑公正。

1. 确定刑量刑建议与传统“责任刑—预防刑”体系的冲突

传统的刑罚理论是建立在“责任刑—预防刑”的二元刑罚论体系之下,以报应刑为基础兼顾预防刑。2016年后认罪速裁程序的试点与实践,在轻罪刑事司法程序中呈现了一定的“契约刑”色彩,实质上是对传统刑罚观提出了一定的挑战。如果不引入“契约刑”理念,起诉阶段由控辩协商方式来商定刑罚即会缺失法理基础。

从认罪行为的具体内容来看,被告人通常可以选择发出三种类型的要约。一是自愿供述其罪,客观上能够减轻公诉机关的举证责任,属于定罪证明意义上的认罪行为;^[27]二是主动放弃申请排除违法证据、管辖异议、申请回避甚至上诉权等诉讼权利,一定程度上有利于诉讼周期的缩短与案件的简化审理,具有诉讼契约的色彩;三是纯粹的实体法量刑要约,即通过具结悔过、积极赔偿、完成企业合规计划、接受社区矫正等犯罪后的认罪、悔罪及社区矫正等非刑罚处罚方法的实效观察,决定是否予以减轻甚至免除刑罚。前两类契约的性质属于诉讼契约,第三类认罪要约则是刑法意义上的量刑裁量要约。不同类型的司法要约,量刑的减轻比例与幅度应当有所不同,方能真正体现认罪协商制度的契约公平原则。应当根据不同的要约类型,设置量刑从宽的具体幅度,赋予司法机关更大的裁量权空间。

即使是轻罪案件,量刑因素仍然复杂、多变与多元化,检察官与法官之间的量刑“算法”差异也较大,如果强行要求办案检察官提起“精准量刑建议”,只能是给出参考性的量刑建议,对法官难以有刚性拘束力。如果要求法院“一般应当采纳”检察机关的量刑建议,唯有修正传统的刑罚观,增立“契约刑”的量刑因子与合理比例,方能真正消解审检机关之间的量刑分歧与冲突。速裁程序中“认罪”的内涵已不再是传统“坦白从宽”或“宽严相济刑事政策”的子内容,而是控辩双方达成的新型的量刑契约。“认罪”不仅仅体现为被告人犯罪后的态度,而是诉讼契约、量刑契约与证据契约的综合体。从定罪证据视角而论,被告人的认罪减轻了侦查、起诉机关的审讯压力,缩短了案件的

^[27] 根据最高人民法院及地方高级人民法院的司法解释,侦查、起诉、审判阶段,如实供述罪行的,可以减轻基准刑罚的30%、20%、10%不等。在强治罪模式下,司法机关的量刑裁量权受到一定的限制,最高的量刑从宽幅度不得超过30%。

处理周期;从诉讼契约视角而言,被告人放弃了申请非法证据排除、管辖权异议、回避甚至上诉权等诉讼权利,通过认罪换取检察官的从轻量刑。

不过,无论是从立法理念还是从司法实践的现实情况观察,量刑建议仍然是检察机关主导的职权主义模式,目前仍然缺失一个控辩平等协商刑罚的过程。一方面,检察官依据传统“责任刑—预防刑”的刑罚理念,只能给出幅度刑;另一方面,对于轻罪案件,检察机关为提升诉讼效率与加快办案周期,要求量刑建议必须“精准化”。只有引入“契约刑罚观”的刑事司法理念,允许控辩达成合意的量刑契约,检察机关以月为单位的精准刑量刑建议,才更具有正当性与合理性。

2. “认罪”行为的量刑低赋值影响量刑“从宽”的幅度与量刑结果

如果只是传统的宽严相济刑事政策下的坦白裁量从宽,量刑从宽将仍由司法机关完全主导;如果是控辩协商型的诉讼契约型从宽,被告人与辩护人在审查起诉阶段的量刑诉求将被合理赋值,公诉机关拥有更多的起诉裁量权与量刑从宽幅度。如果引入诉讼契约理念,建构控辩平等协商机制,超越传统的检察官、法官主导量刑模式,扩张量刑减轻的幅度,轻罪的刑罚模式与治理模式将因此而改变。^[28]

如果仍秉持传统的治罪模式,由检察官完全主导量刑建议而非通过控辩协商确定,检察官的量刑建议仍然是传统“责任—预防刑”刑罚体系的体现,“从宽”的幅度通常不得超过法定刑的30%范围。由于传统的量刑基准体系仍然制约着检察机关量刑裁量幅度,所以,审查起诉阶段检察官最终单方确定的量刑建议,未必符合认罪认罚从宽制度的立法目的与改革初衷。由于在审查起诉阶段,目前仍然缺乏有效的控辩协商的过程、量刑调查与量刑听证程序,检察机关的量刑建议较难体现被告人与辩护人对“从宽”的合理期待,甚至在一些轻罪案件中量刑结果呈现出重刑化的趋势。

所以,控辩平等协商从宽代替检察主导裁量从宽之前,检察机关主导下所确定的精准量刑建议,较难体现认罪速裁程序的立法初衷与新情势,此时再禁止法院予以变更,反而适得其反,最终造成量刑失衡。总体观察,目前的量刑建议更似传统“坦白从宽”机制的升级版,检察官的量刑建议往往建立在报应刑基础上,预防刑与契约刑的权重赋值较低。所以,在正式开庭后立法再禁止或限制法院变更量刑建议,较难以充分体现被告人及辩护人的量刑诉求,盲目采纳检察机关的量刑建议,意味着法院的量刑结果未能充分地兼顾预防刑与契约刑。

三、回归审检制衡与量刑公正:辩证理性对待量刑建议的效力

我国2016年以来确立的认罪认罚速裁程序,虽与美国辩诉交易制度有本质的区别,但认罪速裁制度的初始渊源无疑来自美国。滥觞于美国的辩诉交易制度虽已有百年的历史,但对该制度的批评与改革从未停止。^[29]例如哈佛大学法学院院长罗斯科·庞德甚至批评辩诉交易不过是“违

^[28] 据学者实证研究,依照认罪认罚从宽程序审理的案件,法院的量刑结果有加重的趋势,实际的量刑结果从宽幅度低于改革理想与预期。“与2018年相比,在全面实施认罪认罚从宽制度后的2019年和2020年,全国的缓刑率、微刑率、轻刑率均略有下降;而在认罪认罚从宽制度尚未实施的2015年、2016年,全国的缓刑率、轻刑率都高于2019年和2020年(尤其是缓刑率),重刑率则低于2019年和2020年。这说明,从全国范围来看,认罪认罚‘从宽’制度的普遍适用对法院的量刑几乎没有什么影响。目前2021年的司法统计数据尚未公布,如果新数据不能表明量刑结果朝着有利于被告人的方向发生明显变化,那么被告人通过认罪认罚所能够获得的‘从宽’利益至少从审判结果来看只是一个虚幻的承诺。”孙长永:《中国检察官司法的特点和风险——基于认罪认罚从宽制度的观察与思考》,载《法学评论》2022年第4期,第75—76页。

^[29] See George Fisher, *Plea Bargaining's Triumph*, 109 (5) *Yale Law Review* 857 (2000).

反法律的许可证”。^[30] 中美两国刑事司法背景与诉讼结构相当迥异,不宜盲目效仿美国辩诉交易制度。从辩诉交易制度百年运作情况大体观察,辩诉交易遭受的批评远远超过赞扬。^[31] 如果轻罪案件仍然以司法公正与量刑公正为优先目标,德国认罪协商模式反而可能是我国认罪速决程序的较佳参照蓝本。即使是在辩诉交易广泛适用的美国,检察官量刑建议对审判法官只具有参考性而不具有刚性效力。^[32]

(一) 法院采纳量刑建议的前提条件:庭前对公诉事实的必要审查

庭审中心主义的内涵即包含对事实、证据进行实质化审理的内容,亦包括对量刑因子、量刑方法、量刑均衡进行合理化审查的程序。当前,认罪认罚从宽程序的适用比例高达90%以上,如果均以“小案小办、简案简办”之思维片面强调诉讼效率,将导致绝大多数刑事案件的开庭审理过程形式化。繁简分流、明案速审、认罪从宽的立法初衷固然无可厚非,但如果以牺牲法院开庭审理阶段事实、证据调查权与量刑听审权为代价,则是对司法权威的揶揄与司法公正的背离。量刑建议以案件事实证据清楚、罪名认定正确、认罪真实自愿、量刑因子赋值准确、量刑方法科学为前提基础,法院在庭前会议阶段有必要对之予以审查。即使是被告人认罪认罚的轻罪案件,庭前的公诉审查制度仍然具有正当性与必要性。从我国《刑事诉讼法》第201条明文列举的不采纳情形观察,如果出现被告人认罪不真实、被告人可能不构成犯罪或免除刑罚、诉审存在罪名分歧等情形,法院有权不采纳量刑建议。2018年最高人民法院确立了庭前会议制度,法院如果认为案件存在罪名认定、定罪证据、认罪自愿真实性、管辖回避等程序争议等特定情形,在正式开庭之前应当对公诉进行必要的审查。只有在事实证据清楚、案件定性准确、基准刑设定、量刑方法步骤科学等诸多要件具备的情形下,检察机关的量刑建议才能真正实现罪刑相当。如果开庭前缺乏公诉审查机制,或者公诉审查形式化,一审开庭后量刑建议的极高采纳率可能会偏离罪刑相当与量刑均衡原则。^[33]

(二) 量刑建议“一般应当采纳”条款效力的重新设定

被告人认罪认罚的案件,法院是否采纳检察机关的量刑建议,应当视以下情形而定:(1)“明显不当”判断标准的修正。审查起诉阶段的量刑建议是否公正均衡,不仅要依据责任刑情况审查判断,还要审查检察机关是否充分兼顾了预防刑与契约刑。责任刑以报应刑理念为前提,量刑权重的重心体现在犯罪行为与刑事责任间的均衡度是否相当,预防刑则重在考察犯罪行为人的身危险性与社会预防可能性,契约刑则重在考察被告人犯罪后配合司法机关审讯、起诉、审判的程度。所以,量刑建议的透明度、方法、步骤与量刑基准至关重要,如果检察机关要求法院采纳量刑建议,应当说明其已在轻

[30] See Albert W. Alschuler, *Plea Bargaining and Its History*, 79 (1) Columbia Law Review 1 (January 1979).

[31] See Stephen J. Schulhofer, *Plea Bargaining as Disaster*, 101 Yale Law Journal 1979, 1989 (1992).

[32] See Lauren Sudeall Lucas, *Criminal Procedure — Sixth Amendment — Alabama Supreme Court Upholds a Death Sentence Imposed by Judicial Override of a Jury Recommendation for Life Imprisonment without Parole — Ex Parte Hodges*, 117 Harvard Law Review 1283 (2004); Daniel Frome Kaplan, *Where Promises End: Prosecutorial Adherence to Sentence Recommendation Commitments in Plea Bargains*, 52 University Chicago Law Review 751 (1985).

[33] 盲目追求司法绩效数字化,行政压力传导司法,运用传统治罪模式对待认罪认罚,诸多原因造成了量刑建议的高采纳率。量刑建议的高采纳率,已经受到我国学者的批评。参见孙远:《一般应当采纳条款的立法失误及解释论应对》,载《法学杂志》2020年第6期,第112—113页;孙皓:《量刑建议的高采纳率误区》,载《中外法学》2021年第6期,第2512页;陈兴良:《程序与实体双重视野下认罪认罚从宽制度的教义学反思》,载《政法论坛》2023年第5期,第14页。

罪案件中充分考量了预防刑与契约刑的合理权重。^[34] (2) 情势变更原则。如果开庭后出现事实、罪名、认罪自愿性、重大量刑情节等诸多情势上新变化,即使是轻罪案件,仍应当允许法院变更量刑建议。一般应当采纳条款的适用基础是诉审之间无明显的量刑分歧,即诉审之间对案件事实、证据、罪名适用、起刑点无重大分歧。如果法院审查事实、证据、罪名、量刑后,认定的结果与公诉机关有明显的冲突,应当通过庭审程序重新认定事实,依法改判新罪名或重新确定量刑结果。反之,如果案件的事实、证据、认罪自愿性、定性以及量刑基准无争议,法院则采纳检察机关与被告人之前已达成的量刑协议。(3) 设定不宜适用认罪速裁程序的案件范围。侵犯人身权利、危害国家安全、恐怖活动、有组织犯罪等特定的刑事案件,法院仍然应当谨慎审查量刑建议的正当性与合理性。例如,对于性侵儿童的案件,由于未成年被害人表达能力与意志自由能力的特殊性,即使被告人认罪认罚,司法机关仍然应当以刑罚公正为优先目标,不予从宽量刑,尤其是被告人存在性犯罪前科或严重违法记录的案件。特定案件,基于刑罚正义的优先诉求,即使公诉机关行使求刑权给出明确的量刑建议,法院仍然有权拒绝采纳。(4) 不采纳量刑建议的程序机制。如果诉审之间存在严重的量刑分歧,法院应当遵循正当程序,有义务告知控辩双方将不采纳之前的量刑建议,避免诉讼突袭影响控辩双方的量刑辩论权利。法院依照普通程序审理案件,被告人之前的认罪认罚行为,仍然可作为量刑权重因子而量刑从宽。

(三) 量刑“算法”的透明度与量刑因子的合理赋值

检察机关的量刑建议,无论是以精确刑还是幅度刑方式提起,无论是报应刑调节量刑基准还是依据预防刑、契约刑推行刑事合规缓起诉,只有确保量刑“算法”的公正、透明,法院采纳量刑建议才符合罪刑相当原则。^[35] (1) 审查起诉阶段“量刑听证”程序的必要性。即使是三年以下的轻罪认罪案件,量刑算法仍然应当透明,审查起诉阶段可以给予被害人、被告人及辩护人相应的量刑建议“听证”的权利。量刑过程,通常是“起点刑—基准刑—宣告刑”的三阶段过程,基准刑的确立与计算方法复杂、多变而且不同个案间的差异明显。^[36] 在辩护人不在场听证的情形下,要求被告入接受认罪具结书及附带量刑建议,不仅会影响量刑的透明指数,也会影响到开庭审理后法官 20% 的自由量刑调节权的行使。^[37] 审查起

[34] “从宽”究竟是报应刑罚观下的从宽抑或预防刑、契约刑罚观下的从宽,对于轻罪案件量刑方法、结果尤其重要。被告人认罪认罚并且能够真正悔悟自新的轻罪案件,通常能够反映被告人与国家刑罚之间的“沟通正义”,重刑与加重量刑的必要性减轻,和缓刑刑罚幅度的适用概率更高。传统报应刑思想已被各国刑事司法实践修正,将犯罪行为背后的社会性因素及其他影响性因素纳入量刑考量因素。即使是以报应刑为量刑基础,纳入量刑视野的也不局限于犯罪行为自身,而是要考量所有可能影响罪责、可非难性的因素,例如日本等国甚至将侦查人员刑讯逼供、违法采集血液等违法侦查行为情形也作为法院酌定减轻量刑的因素。参见范耕维:《以违法侦查作为量刑因子之正当化根据:一个刑罚理论的观察》,载《台大法学论丛》2019年第4期,第2182页。

[35] 目前,欧美国家中检察官量刑建议采纳率最高的是荷兰。其将犯罪因子量刑化点数,制定全国统一、透明、条理化的“检方求刑准则指南”(Frame of Prosecutorial Sentencing Guidelines),同时通过资讯科技辅助技术,为检察官量刑建议提供数据支持。不过,即使如此,荷兰法官的量刑采纳率只达 80%,也不及我国目前 90% 以上的量刑采纳率。参见郭豫珍:《量刑与刑量——量刑辅助制度的全观微视》,元照出版公司 2013 年版,第 70—71 页。

[36] 例如,2021 年最高人民法院、最高人民检察院印发《关于常见犯罪的量刑指导意见(试行)》(以下简称“量刑《指导意见》”)规定:“……具有多个量刑情节的,一般根据各个量刑情节的调节比例,采用同向相加、逆向相减的方法调节基准刑;具有未成年人犯罪、老年人犯罪、限制行为能力的精神病人犯罪、又聋又哑的人或者盲人犯罪,防卫过当、避险过当、犯罪预备、犯罪未遂、犯罪中止,从犯、胁从犯和教唆犯等量刑情节的,先适用该量刑情节对基准刑进行调节,在此基础上,再适用其他量刑情节进行调节。”

[37] 依据量刑《指导意见》的相关规定,即使基准刑确定后,宣告刑要“综合考虑全案情况,独任审判员或合议庭可以在 20% 的幅度内对调节结果进行调整,确定宣告刑。当调节后的结果仍不符合罪责刑相适应原则的,应当提交审判委员会讨论,依法确定宣告刑”。可见,量刑因素相当复杂、多元,量刑背后的宽严相济刑事政策具体内容也易变。量刑《指导意见》赋予审判法官 20% 的裁量权幅度,如果一律采纳检察机关的量刑建议,《指导意见》中为法官预留的量刑弹性条款易被虚置。

诉阶段,量刑建议如果已经过一个正式的听证或准听审过程,引入被害人、被告人、辩护人在场的异议机制,更能够提升量刑过程的透明度与公正度,也为法院审查、采纳量刑建议提供庭前正当程序基础。(2) 区分轻、重罪,量刑建议效力实行双轨制。从刑罚原理而论,重罪的刑度浮动弹性远远高于轻罪案件,检察机关重罪案件提起的量刑建议原则上只具有参考性,对法院不应当具有强制约束力。例如,从德国、日本及我国台湾地区的“认罪协商”适用范围考察,为了保证刑罚公正,重罪案件不得适用协商程序,诉讼效率并非重罪刑事司法的优先目标。^[38] 如果将高采纳率作为司法绩效考核的优先目标之一,要求法院遵循一年的量刑裁量区间而“一般应当采纳”检察机关的量刑建议,不仅侵蚀了法官的量刑裁量权,也背离了刑罚公正。与轻罪案件不同,重罪案件量刑关系刑罚正义的底线,量刑权重因子也更为复杂多元化,量刑起点刑的跨度从六个月有期徒刑至死刑。一旦法院采纳量刑建议有误,将抵触罪刑相当原则与刑罚正义。从刑罚原理考察,预防刑、诉讼契约刑在轻罪案件能够有更多的适用空间,重罪的惩罚原则上以报应刑为基础,较少甚至不考虑预防刑的比例。所以,三年有期徒刑以上的重罪案件,法院应当有更广泛、弹性的量刑裁量权,检察机关的量刑建议只具有参考效力,不受检察机关一年量刑裁量区间的限制。(3) 提高认罪认罚行为的量刑赋值。精准刑抑或幅度刑,并非认罪速裁案件量刑建议的关键问题,被告人认罪认罚行为的量刑权重赋值才是关键。控辩协商达成从宽协议后,可以在量刑裁量空间内确定量刑的具体从宽比例。即使正式开庭后,出现罪名变更、量刑事实变化甚至被告人反悔等情况,法院仍然可以认可庭前控辩双方的认罪从宽幅度的具体比例。即被告人与检察机关之前达成的诉讼契约,法院原则上推定仍然有效,仍然遵循控辩双方从宽量刑的减轻比例。认罪认罚的量刑赋值比例可被视为过程、方法上的公正,即使庭审中出现罪名变更、犯罪事实情节变化、刑事和解等新情形,只要法院认定控辩双方之前的认罪量刑赋值与减轻幅度比例合理,仍然可以依照认罪速裁程序审理、量刑,控辩双方之前协商达成的量刑从宽的比例仍可以维持不变。如前所述,随着刑事合规、未成年人犯罪附条件缓起诉等新的改革举措的推进,“认罪”包括至少三种类型的认罪要约,仅依传统的宽严相济刑事政策而从宽处罚,其从宽幅度低于犯罪嫌疑人、辩护人的诉讼预期,轻罪刑事司法中给予司法机关更自由弹性的从宽比例,方能真正实现控辩平等与案件的繁简分流,亦为审判中心主义改革与重罪案件的庭审实质化创造条件。

(四) 确定刑量刑建议扩张适用的误区与修正

对于轻罪刑事案件,检察机关以确定刑的方式提出量刑建议(量刑精确度以月计算),受《刑事诉讼法》第201条“一般采纳条款”约束,法院通常采纳确定刑,因此检察机关审查起诉阶段确定的精准量刑建议已近乎等同于法院的量刑结果。支持精准刑的论点主要是基于精准刑的确定性,能够在审查起诉阶段将量刑结果预期确定,促使被告人积极认罪而量刑从宽。^[39] 同时,一个月以上拘役到三年有期徒刑刑罚以下的轻罪案件,量刑跨度区间较小,量刑建议精准到月,似乎也利于检察机关庭前即能确定准确的刑度,强化检察官在认罪速裁程序中的职权主导地位。然而,“确定刑为主、幅度刑为辅”的量刑建议方式,既会限缩法官的量刑裁量空间,亦会使公诉检察官客观上行

[38] 参见[德]托马斯·魏根特:《德国刑事诉讼法原理》,江溯等译,中国法制出版社2021年版,第322页。[日]土本武司:《日本刑事诉讼法要义》,董璠舆、宋英辉译,五南书局1997年版,第264页。王兆鹏、张明伟、李耕荣:《刑事诉讼法(下)》,新学林出版公司2023年版,第125页。我国台湾地区,重罪案件不得适用认罪协商程序,只有缓刑、两年以下有期徒刑、拘役或罚金刑的案件,才能以认罪协商程序结案。

[39] 参见余剑:《精准化量刑建议的法律性质及其展开》,载《东方法学》2023年第4期,第184页;董坤:《认罪认罚案件量刑建议精准化与法院采纳》,载《国家检察官学院学报》2020年第3期,第34页;周新:《论认罪认罚案件量刑建议精准化》,载《政治与法律》2021年第1期,第21—33页。

使了庭前“准量刑法官”的功能。“确定刑为主、幅度刑为辅”的量刑建议方式存在以下风险：（1）牺牲刑罚公正。在量刑技术、量刑方法、量刑听证程序、量刑调查制度等诸多配套机制尚不具备或不完善的前提下，由检察官担任量刑法官，只能以部分牺牲刑罚公正为代价。（2）确定刑方式的量刑建议不应当适用“一般应当采纳”条款。量刑建议如果以确定刑方式提起，即以月为精度单位来设定自由刑刑期，实质上即是对法官量刑裁量权的限制。由于《刑事诉讼法》第 201 条将量刑建议“明显不当”作为法院不采纳的条件，确定刑实质上侵蚀了法官个案中的量刑裁量空间。（3）增加纠正量刑错误的难度。检察机关在审查起诉阶段依职权单方设定量刑建议后，即使在正式开庭后被证明失当，对于量刑偏差在 30% 比例内的量刑建议，由于并非属于“明显不当”，我国法院如对量刑建议予以调整，将与第 201 条相抵触，法官可能仍然采纳量刑建议。如果一审判决结果出现量刑不当，上诉审法院对之改判的概率相当低，这事实上变相使原审检察官“非明显不当”的量刑建议产生了一定的终局效力。所以，即使是轻罪案件，检察机关的量刑建议仍然应当以六个月以上的幅度刑为主，确定刑只能适用于事实、证据、定性无争议的案件。总之，量刑建议应当是责任刑、预防刑与契约刑的综合评估、量刑因子赋值与合理权重，在量刑调查、量刑评估、量刑规范化、量刑大数据辅助技术等条件尚不具备的情况下，以幅度刑方式提出量刑建议更符合当代刑罚科学，也遵循当代刑事司法诉审分离制衡的基本原则，也更有利于法官量刑裁量权的充分行使，在个案中积极实现刑罚正义。

Abstract Actually, Chinese prosecutors' sentence recommendations are approaching and effective on the judge's final decisions on the courts according to the Criminal Procedure Law (Article 201), which also means restricting the judge's proportional judicial power in the sentencing process. As a result of the negative influence that Article 201 has been widely coming into force, some unexpected reactions have happened, whether on procedure justice or on penalty justice. In fact, the defendants' rights of pretrial hearings and punishment arguments are generally denied at the price of guaranteeing prosecuting. Meanwhile, Article 201 has some negative influences on criminal procedural justice on the background of guilty plea process, which literally robust the prosecutors with sentence recommendations. In as so much public prosecutors usually cannot investigate all the factors on crime with due diligence, finally the sentence recommendations by the prosecutors cannot represent full justice on criminal punishment. Some necessary judicial pretrial reviews on sentence recommendations case by case are important to reject some irrational recommendations for the court. By establishing the calculus methodology of criminal punishment and rational proportionality on sentence circumstances, sentencing justice will come into reality and the principle of separation of powers in criminal justice surely will also be sustained again.

Keywords Sentence Recommendations, Judge's Exclusive Power of Minus Civil Rights, Special Quanta of Punishment, Range-Only Theory of Punishment, Sentencing Justice

（责任编辑：樊传明）