法学研究 Chinese Journal of Law
Vol.43，No.4（2021）

量刑指导意见的司法实践与重构

——以盗窃罪为切入点

彭雅丽 *

内容提要：学界对量刑指导意见已有诸多研究，但依托计量分析模型评估其运行效果的实证研究，尚属空白；现有的规范研究，也难以从规范制定技术上满足对量刑指导意见的改进需要。以盗窃罪量刑为切入点，实证地考察了量刑指导意见的司法实践，并有如下发现：在有期徒刑量刑中，趋轻情节的作用力不及趋重情节；量刑指导意见对罚金刑的指导不足。通过借鉴英美量刑指南的立法技术，可以改进我国的量刑指导体系：其一，要将量刑起点确定为具体的点；其二，在以数额为量刑中心的现状下，要缓和边际数额的决定性作用；其三，从量刑起点出发确定基准刑时，应当合并考虑数额；其四，在量刑情节的编排上，要先确定责任刑情节、后确定预防刑情节，并在有期徒刑量刑中扩大趋轻情节的适用空间；其五，应当加强对罚金刑的量刑指导，构建更为细致的罚金刑量刑体系，并在与自由刑的搭配上做到实质公平。

关键词：量刑指导意见　盗窃罪　量刑基准　罚金

自20世纪70年代弗兰克尔法官指出“量刑无法”的问题以来，各国在量刑规范化上做出了积极尝试。〔1〕其中，美国明尼苏达州首先采用表格式量刑指南的形式，限制法官的裁量权。我国的量刑实体改革主要采用了量刑指导这一形式。2010年《人民法院量刑指导意见（试行）》标志着我国量刑规范化的全面推进。近年来，量刑指导意见虽几经修正，但其内容仍饱受学者诟病。〔2〕

刑罚立法是刑事立法必须考虑的一个基本方面，量刑指导意见作为刑罚立法的重要补充，

---

* 美国芝加哥大学法学博士（J. S. D.）研究生。
本文得到“上海交通大学凯原海外青年访问学者”项目的资助。

〔1〕 See Marvin E. Frankel, *Lawlessness in Sentencing*, 41 U. Cin. L. Rev. 1 (1972); Michael Tonry, *Punishment Policies and Patterns in Western Countries*, in Michael Tonry & Richard S. Frase (eds.), *Sentencing and Sanctions in Western Countries*, Oxford: Oxford University Press, 2001, pp. 3 – 28.

〔2〕 学界在量刑原则是否有必要出现在指导意见中、量刑基准的确定、量刑步骤、量刑情节的区分、犯罪数额的绝对化等方面，对量刑指导意见提出了批评与建议。参见周长军：《量刑治理的模式之争——兼评量刑的两个指导“意见”》，《中国法学》2011年第1期，第52页；刘军：《量刑如何实现均衡——以量刑规范性文件为分析样本》，《法学》2011年第8期，第148页；王瑞军：《体系性思考与量刑的规范化——以〈量刑指导意见〉及实践为分析对象》，《政法论丛》2014年第6期，第72页；李冠煜：《量刑规范化改革视野下的量刑基准研究——以完善〈关于常见犯罪的量刑指导意见〉规定的量刑步骤为中心》，《比较法研究》2015年第6期，第111页。

其制定及适用是制定机关和学者都应当仔细研究的问题。这首先当然是一个法教义学问题，但却无法仅凭法教义学就能够解决。量刑指导意见有没有效果、有什么样的效果，只有实证研究才能给出答案。有了答案之后，法教义学改进量刑指导意见的看法才有了根基。另外，法教义学的有关主张无法脱离立法技术或规范制度技术而径直被成文法或规范性文件吸纳。因而，改进量刑指导意见，不仅要讨论“道”即法教义学，也要研究“术”即规范制定技术。在这一问题上，借鉴域外立法是有益的。

就量刑指导意见的实际效果以及如何改进量刑指导意见这两个问题，本文试图通过定量研究和比较研究的方法，为已有研究提供一些补充性的和实用性的知识。本文的结构安排如下：（1）讨论量刑指导意见的实践问题以及重构困境，并引出本文所采用的方法论。（2）列明研究假设、数据与方法。（3）报告线性回归的结果，同时归纳总结量刑指导意见在具体适用上的问题。（4）对这些问题背后的结构性缺陷进行反思，并在比较研究的基础上提出重构的建议。（5）对整篇文章作简短总结，呼吁学界作更深入的研究。

一、实践问题、重构困境与研究方法

（一）实践问题与定量研究

目前，关于量刑指导意见的实际效果，存在两个主要问题：一是它到底有没有效果；二是如果有效果，有没有一些副作用。毕竟，量刑指导或量刑指南存在一些宏观上的通病。其一，难以完全解决量刑失衡的问题。虽然相比起没有量刑指导或量刑指南的无序状态，指导或指南几乎确定能够提高量刑的一致性，但其本身存在应用上的问题以及深层次矛盾。比如，“一些法院或者法官对《意见》的内容和精神理解的不够透彻，在实践中不免出现了这样那样的问题。其中，既有对‘定性定量法’运用不当的问题，也有机械量刑、‘数字化’量刑的问题”。[3] 更深层次的矛盾是，量刑的目标是同案同判且异案异判，二者缺一则难言均衡，但量刑指导或量刑指南在一个方面的改进，往往意味着另一个方面的牺牲。例如，明尼苏达州式的表格法量刑指南对法官的限制较多，在促进同案同判的同时，也降低了异案异判的可能。过高的审判确定性，也使美国检察官在辩诉交易中获得了不必要的优势。[4] 然而，一旦放松对法官的制约，必然会放大同案不同判的问题。因此，量刑指导或量刑指南的制定者必须小心地在这二者之间寻找适当的平衡。在这一维度上，我国量刑指导意见的实际效果仍有待考察。

其二，量刑指导或量刑指南可能实质上助长重刑主义。在美国，这主要体现在监禁率的提高上。实证研究表明，一些州的量刑指南恶化了过度监禁的问题。[5] 在英国，暴行罪的量刑指南公布后，对其的跟踪研究发现该罪的平均刑期显著提高。[6] 另外，量刑指导或量刑指南的制定依据若是既有的量刑实践，则可能吸收“沉疴”，阻碍未来的刑罚轻缓化。

---

〔3〕 最高人民法院中国应用法学研究所编：《量刑规范化典型案例（1）》，人民法院出版社 2011 年版，第 1 页。

〔4〕 See Albert W. Alschuler, *Sentencing Reform and Prosecutorial Power: A Critique of Recent Proposals for "Fixed" and "Presumptive" Sentencing*, 126 Univ. Pa. Law Rev. 550 – 577 (1978).

〔5〕 See Sean Nicholson-Crotty, *The Impact of Sentencing Guidelines on State-level Sanctions: An Analysis over Time*, 50 Crime Delinq. 395 – 411 (2004).

〔6〕 See Sentencing Council, *Assessing the Impact and Implementation of the Sentencing Council's Assault Definitive Guideline* (2015), http://www.sentencingcouncil.org.uk/publications/item/assault – final – resource – assessment/. 本文提及的“英国”指英格兰和威尔士，不包括苏格兰和北爱尔兰。

在上述通病之外，学界从量刑指导意见的内容出发，指出了我国特有的技术层面的问题。首先，在量刑基准上，由于目前的指导意见在量刑起点上主张“幅的理论”，法官难以准确把握量刑起点。[7] 其次，在量刑情节的确定与整理上，指导意见简单地按量刑情节的宽严进行分类和排列，法官对情节只能“同向相加、逆向相减”，在某些情节上“部分连乘、部分相加减”，而不区分责任刑情节和预防刑情节。

那么，我国量刑指导意见对量刑均衡和重刑主义的实际作用是什么？一些缺憾能否通过技术调整加以解决？这些问题，都需要进一步的实证研究。但遗憾的是，我国在官方层面并未对量刑指导意见的实际效果进行公开的跟踪研究。与此形成对比，英国明确规定，量刑委员会的重要职责是监督量刑指南的实施和效果。[8] 实际上，英国对每一罪行的量刑指南都提供了相应的评估报告。量刑指南的发源地美国明尼苏达州，也会每年对该年度的量刑实践进行梳理。

同时，我国学界也缺乏对量刑指导意见实施效果的实证研究。现有围绕量刑的研究，多是针对某个具体罪名量刑实践的检讨。[9] 在为数不多的文献中，对量刑指导意见作用的实证评估，一是通过走访调研，形成比较直觉的认识。[10] 但这种方法更侧重学者的主观感受，可能存在较大的个人偏误，因而难以成为有力的证据。二是通过实证研究探究结果类指标，[11] 比较典型的如被告人最终的宣告刑是否在量刑指导意见所容许的幅度内，即考察合规率。[12] 这种方法是英美评估量刑指南的重要办法，但并不适合我国。[13]

若将量刑过程想象成一个包含自变量、因变量和算法的公式，我国的量刑指导意见相较英美量刑指南有明显的区别。美国的表格式量刑指南针对各罪行等级，根据横轴（累犯分数）与纵轴（罪行轻重）给出量刑结果，直接提供了自变量和因变量。英国根据各具体罪行提供量刑指南，在这些指南的步骤一中提供自变量（应该考虑的量刑情节）以区分罪行轻重，并在步骤二中提供各类别的量刑表。综而言之，英国也提供了自变量和因变量。区别于英美，我国量刑指导意见并不提供因变量，而是提供了自变量（常见量刑情节）及算法（基于量刑情节所能调整的幅度）。

由于英美并不对算法进行规制，法官只能对量刑情节进行整体考量，所以研究者难以定量地区分各量刑情节的作用，只能以合规率为标准来确定量刑实践是否符合量刑指南。[14] 但是，由于我国量刑指导意见提供了明确的算法，定量地研究指导意见的作用机制，成为了一条切实可行的路径，也是更优越的路径。这是因为，在我国语境下仍使用合规率作为评估标准，

---

〔7〕 参见前引〔2〕，李冠煜文，第 117 页以下。

〔8〕 Coroners and Justice Act 2009, Sections 127 – 131.

〔9〕 例如文姬：《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；李冠煜：《贪污罪量刑规范化的中国实践——基于〈刑法修正案（九）〉生效后的案例分析》，《法学》2020 年第 1 期。

〔10〕 参见熊秋红：《中国量刑改革：理论、规范与经验》，《法学家》2011 年第 5 期，第 37 页。

〔11〕 左卫民的定量研究也比较了上诉率等指标，但其问题意识着重于程序改革而非实体改革。参见左卫民：《中国量刑程序改革：误区与正道》，《法学研究》2010 年第 4 期。

〔12〕 参见蔡曦蕾：《量刑失衡的克服：模式与选择》，《中外法学》2014 年第 6 期，第 1596 页。

〔13〕 See Neil Hutton, *The Definitive Guideline on Assault Offences: The Performance of Justice*, in Andrew Ashworth & Julian V. Roberts (eds.), *Sentencing Guidelines: Exploring the English Model*, Oxford: Oxford University Press, 2013, pp. 86 – 103.

〔14〕 这一结论以评估标准是狭义的“是否符合量刑指南”为前提。若涉及广义的对量刑指南作用的评价，则需要同时评估量刑指南是否带来“更好的”量刑，进一步探究在量刑一致性以外，量刑指南是否有其他副作用。实际上，英国量刑委员会也就公共意见、资源评估作了相关研究。参见 Nicola Padfield, *Exploring the Success of Sentencing Guidelines*，载上引 Ashworth 等编书，第 31 页以下。

存在两个问题：其一，由于量刑情节上机械的“同向相加、逆向相减（部分连乘）”并不符合学界在规范上要求责任刑作为量刑上限的共识，所以作为合规率评价基础的量刑标准本身就存在纰漏。简言之，更高的合规率并不一定意味着更优的量刑。其二，这种方法只能比较量刑结果，而无法解释量刑过程。它无法回答是什么在影响着合规率，量刑指导意见的作用机制仍然像是一个“黑匣子”，这种方法无法提供进一步改善量刑指导意见的重要信息。

由此，本文针对司法实践如何应用量刑指导意见这一问题，使用定量的回归分析模型，辅以对广州市基层法院法官的访谈，探究量刑指导意见的作用机制。与过往的评估路径不同，本文并不探究最终的合规率。本文关注的是，指导意见所规定的量刑情节在实践中如何影响量刑过程。这种方法可以直观地揭示量刑指导意见的具体效果，更好地判断法官量刑实践的优劣。

值得一提的是，在认罪认罚从宽制度背景下，检察官的量刑建议权得到了进一步凸显。2021 年最新修订的《关于常见犯罪的量刑指导意见（试行）》由最高人民法院、最高人民检察院联合印发，通知中明确规定各级检察院对于符合规范范围的 23 种常见犯罪的量刑应当按照该意见提出量刑建议。因此，研究量刑指导意见的作用机制，也有利于未来更好地规范检察院的量刑建议工作。

（二）重构困境与比较研究

量刑指导意见面临的重构困境是，学界的争鸣似乎对意见的修正并无影响。量刑基准问题是指导意见需要解决的重要问题，本文以这一问题为例说明学界影响实务的难度。学界过往对量刑基准问题有两种研究路径。实证研究者希望通过定量研究，从司法实践中直接归纳出量刑基准。[15] 然而，实然并不当然代表应然。在美国明尼苏达州确定其量刑指南时，过往的量刑实践并未发挥作用；其纯粹从公共政策出发，平地起高楼般地建立了一套全新的量刑体系。[16] 再者，确定量刑基准的标准（量刑均值、量刑中值、其他值）并无定论。因此，量刑指导意见难以将实证研究成果纳入体系。相较之下，法教义学者希望通过论理演绎式地对量刑基准提供指导。由于量刑基准理论本身就是一个舶来品，法教义学的研究在一定程度上有比较研究的成分，不过其并不探讨立法技术或规范制定技术，而仅借鉴理论。但是，法学理论的较量往往少有结果，因此相关理论难以被量刑指导意见所采纳。责任刑到底应当是一个范围（幅的理论）还是一个点（点的理论），[17] 如何对“幅的理论”和“点的理论”进行调和，[18]“相关理论比较烦琐，与实践不能很好对接。”[19] 或许正是因为这样，目前的量刑指导意见与学界的研究存在较大脱节。

在具备司法实践信息的基础上，本文通过与英美量刑指南立法技术的比较研究，尝试打破上述僵局。这一尝试的贡献有两个方面：其一，通过比较研究，提供能够被量刑指导意见直接吸收的规范制定技术。其二，根据定量研究所展现的司法实践情况，选择当下最能解决实践问

〔15〕 参见白建军：《量刑基准实证研究》，《法学研究》2008 年第 1 期；王利荣：《对常见犯罪量刑基准的经验分析》，《法学研究》2009 年第 2 期；李安：《量刑实证研究的方法论检视：从实证观念到实证技术》，《中外法学》2009 年第 6 期。

〔16〕 See Austin Lovegrove, *Judicial Decision Making, Sentencing Policy, and Numerical Guidance*, New York: Springer, 2012, p. 47.

〔17〕 参见周光权：《量刑基准研究》，《中国法学》1999 年第 5 期，第 127 页；张明楷：《责任主义与量刑原理——以点的理论为中心》，《法学研究》2010 年第 5 期，第 128 页。

〔18〕 参见前引〔2〕，李冠煜文，第 120 页以下。

〔19〕 参见周光权：《量刑的实践及其未来走向》，《中外法学》2020 年第 5 期，第 1150 页。

题的理论。

需要明确的是，英美在改进量刑指南上与我国处于不同阶段。英美量刑指南的改进，更多是在具体问题上“小修小补”，使其更精致、更准确。例如，英国量刑委员会在2011年对因走私毒品而被监禁的女性犯罪人（典型如人体藏毒者）进行了访谈。通过对这些犯罪人的形象刻画，让她们讲述自己的故事以及犯罪原因，量刑委员会决定降低该类犯罪人的刑期。[20]然而，我国的量刑指导意见尚有结构性调整的空间。基于这种考虑，本文侧重对英美立法技术的借鉴，而非量刑指导或量刑指南本身的比较。

本文选择盗窃罪作为研究的切入点。选择盗窃罪，是因为：其一，盗窃罪的发案率在我国刑事案件中高居前列，[21]对盗窃罪的量刑实践进行实证研究，具有相当的代表性。其二，在实证研究上，盗窃罪具备特别的优势。为了准确探究各情节与量刑之间的关系，需要尽可能控制其他因素。而盗窃罪中，盗窃数额是相当精确的量化指标；同时，盗窃罪的案情也相对简单。其三，盗窃罪样本在罪质和刑罚上存在人为缺失、错误的情况很少，如有疏漏，更大可能是源于裁判文书上传的误差。因而，该类案件样本较其他样本（如中国裁判文书网上的死刑案件样本）更少样本偏差。

二、研究假设、数据与方法

（一）研究假设

量刑统一有两个方面，一是结果统一，二是手段统一。美国明尼苏达州的量刑指南追求的主要是结果统一。从公共政策的立场出发，该量刑指南的重要任务是精确预测并控制监禁人数。[22]表格式量刑指南卓越地完成了这一任务，因为在知晓被告人罪行及犯罪历史的情况下，就能预估其将要受到的刑罚。[23]相较之下，我国更追求手段统一。量刑指导意见并未将管控监禁人数列为目的之一，而是旨在提供统一的量刑方法与步骤，期待法官在使用统一的方法后得到令人满意的结果。以盗窃罪为例，法官应当遵循的步骤是：首先，依据基本犯罪构成事实在法定刑幅度内确定量刑起点。其次，在量刑起点的基础上，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。最后，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。

法官的量刑实践应当在何种程度上受指导意见的规制，可以从意见的行文沿革中获得线索。2014年《关于常见犯罪的量刑指导意见》较2010年《人民法院量刑指导意见（试行）》，在常见量刑情节的适用上删去了“可以在相应的幅度内确定具体调节比例”的规定。[24]删去这一规定，意味着法官若选择适用指导意见所规定的某一具体量刑情节，就必须在所规定的幅

〔20〕 参见前引〔14〕，Padfield文，第43页。

〔21〕 参见2020年最高人民法院工作报告，http://www.court.gov.cn/zixun-xiangqing-231301.html，2021年7月11日最后访问。

〔22〕 See Dale G. Parent, *What Did the United States Sentencing Commission Miss*, 101 Yale L. J. 1773 (1991).

〔23〕 当然，这只是初步的推定结果，最终的宣告刑还要继续考量趋重情节与趋轻情节是否容许量刑偏离推定刑。

〔24〕 量刑规范化全面推进后，我国的量刑指导意见先后有2010年、2014年、2017年和2021年四个版本。2017年版较2014年版在常见量刑情节及调节幅度、盗窃罪量刑等方面几无区别（仅调整了立功条款的顺序）；2021年版也保留了上述有关方面的原有条款。本文实证研究选用的数据是2015年的数据，所以实证结果参照标准所指的量刑指导意见是2014年版。其他情况除列明外，通指2014年量刑指导意见的条款。

度内确定具体的调节比例。发布 2021 年量刑指导意见的通知中也再次强调，各级法院对纳入规范范围的案件应当按照该意见规范量刑。

据此，本文构建假设 A：量刑情节与量刑之间的相关性与量刑指导意见所规定的一致。例如，指导意见规定，累犯情节应当增加基准刑的 10%—40%，那么，累犯情节与刑罚之间应当具备显著相关性。同时，依据作用幅度对量刑情节排序，作用幅度大的量刑情节，其作用力应当大于作用幅度小的量刑情节。对于非法定量刑因素，由于其并非犯罪事实，也不属于任何法定和酌定量刑情节，故不应显著影响量刑。

另外，量刑指导意见要求，“量刑时要充分考虑各种法定和酌定量刑情节”，而其列明的只是“常见量刑情节”，这说明意见所涵盖的量刑情节并不周延，即量刑指导意见本就不期待其规定的量刑情节能百分百地解释量刑。不过，盗窃罪属于量刑指导意见所规制的常见犯罪之一，且盗窃罪案情比较简单。笔者在阅读判决书时也发现，指导意见所规定的量刑情节覆盖了影响盗窃罪量刑的绝大部分量刑情节。据此，本文构建假设 B：量刑指导意见所规定的因素应当解释大部分的盗窃罪量刑。

（二）数据与方法

本文的研究对象是一审盗窃罪案件。本文的样本均来自中国裁判文书网，选择“刑事案件”为案件类型，案由选择“盗窃”，文书类型限定为“判决书”，裁判年份为“2015”，法院地域为“广东省”，选择广州市辖下的各基层法院，剔除重复案件、无法辨认案件与最终判定为他罪案件，共获得 999 份有效一审判决书。若文书有多个被告人的，进行分项列出，共获得 1226 个样本，将未遂与判处管制、拘役的样本剔除，最终获得 916 个样本。[25] 排除管制样本，是因为量刑指导意见确定宣告刑的最后一条是“综合全案犯罪事实和量刑情节，依法应当判处无期徒刑以上刑罚、管制或者单处附加刑、缓刑、免刑的，应当依法适用。”这就意味着，指导意见所规范的对象不包括管制。排除拘役样本，是因为实务界并无如何统一拘役与有期徒刑的认识。诚然，白建军、胡昌明提出过构建刑罚量以统一拘役与有期徒刑的方法，[26] 但没有哪个法官在判案时会想着拘役是有期徒刑的 75% 抑或 90%。对法官而言，拘役就是拘役，无法与有期徒刑进行替换。本文要探究的问题是量刑指导意见的作用机制，而非某个因素或情节如何影响刑罚量，这就意味着本文的视野只能是法官的视野，如果将拘役和有期徒刑用法官所没有的观念统一起来，反而会得到不恰当的结果。同时，对拘役进行单独建模必要性不大，因为有期徒刑已足够代表自由刑，而且拘役的样本量较小。排除未遂样本，是因为在阅读判决书的过程中，笔者发现未遂样本中既有普通未遂，也有多次盗窃中的某一次盗窃未遂，后者既有未遂又有既遂，作用机制较为复杂，可能扰乱模型对真实系数的贴近。

在选定样本后，笔者阅读了所有判决书，并记录了样本的基本信息、裁判理由部分认可的所有量刑情节以及最终的刑罚。本文设置的因变量，一是有期徒刑刑期，二是罚金刑数额。[27]

〔25〕 中国裁判文书网的判决书一直处于更新状态，因而判决书数量或与现在用相同检索条件检索所得数量有出入。

〔26〕 参见白建军：《刑罚轻重的量化分析》，《中国社会科学》2001 年第 6 期，第 115 页；胡昌明：《被告人身份差异对量刑的影响：基于 1060 份刑事判决的实证分析》，《清华法学》2018 年第 4 期，第 93 页。

〔27〕 本文未将缓刑纳入因变量，是因为缓刑属于刑罚的应用，而非刑罚本身。量刑指导意见所起的作用应仅在法官确定实刑这一环节，之后，法官应根据刑法的规定确定是否适用缓刑并确定缓刑的考验期。再者，观察量刑指导意见的规定可知，缓刑并非量刑指导意见所规范的对象。2021 年量刑指导意见虽然具体规定了适用缓刑的原则，但其并未超出刑法的规定提供实质性指导。

本文设置的自变量有：（1）非法定量刑因素：案件所属行政区人均 GDP 水平是否高于全市人均水平、被告人是否有辩护人、是否外来人口；（2）裁判理由部分提出的量刑情节：盗窃数额、次数，是否入户、扒窃、携带凶器，是否使用破坏性盗窃手段，被告人是否有前科、累犯、特殊的身体状况、是否属于限制刑事责任能力人，是否达成谅解、退赃退赔缴回、未缴犯罪所得、自首、坦白、当庭认罪、立功，被害人是否弱势。[28]

本文选择多元线性回归方法，以公式表达如下：[29]

log 有期徒刑$_i = \alpha + \beta$ 非法定量刑因素$_i + \gamma$ 量刑情节$_i + \varepsilon_i$

以及

log 罚金刑$_i = \alpha + \beta$ 非法定量刑因素$_i + \gamma$ 量刑情节$_i + \varepsilon_i$

其中，ε 代表无法被观测到的情节，β 与 γ 代表系数。

三、量刑情节的实证检验

（一）有期徒刑模型

有期徒刑模型的线性回归结果如表 1 所示。

表 1　有期徒刑线性回归模型

	模型一	模型二	模型三
行政区	-0.0445(0.0264)	-0.0408(0.0263)	
辩护人	0.0158(0.0441)	-0.0000236(0.0432)	
外来人口	0.0214(0.0321)	0.0268(0.0318)	
盗窃数额	0.202***(0.00849)	0.202***(0.00846)	0.202***(0.00823)
盗窃次数	0.204***(0.0247)	0.201***(0.0246)	0.199***(0.0245)
入户	0.223***(0.0284)	0.224***(0.0284)	0.225***(0.0282)
扒窃	0.195***(0.0322)	0.196***(0.0323)	0.192***(0.0317)
携带凶器	0.263(0.137)		
前科	0.0893**(0.0315)	0.0851**(0.0315)	0.0873**(0.0314)
累犯	0.293***(0.0228)	0.290***(0.0228)	0.292***(0.0227)
特殊身体状况	-0.279(0.155)		
限制刑事责任能力人	-0.238(0.307)		
被害人弱势	0.671*(0.306)	0.671*(0.306)	0.641*(0.305)
破坏性盗窃手段	0.138(0.218)		

〔28〕 本文所指的前科不包括累犯。这些自变量除盗窃数额为数值变量外，均为名义变量。“行政区”自变量意图探究“富裕地区”与“落后地区”之间是否存在显著量刑区别。广州全市均为广东省高级人民法院所规定的“一类地区”，因而该自变量不应体现出显著性。另外，笔者虽然记录了共犯的信息，但记录时并未区分同等作用的共犯、主犯、从犯，因而，将共犯样本放入模型，可能会使部分系数失真。为了更准确地反映量刑指导意见的作用机制，使系数更贴近真实系数，本文剔除了有共犯的样本。

〔29〕 因变量的数额范围较大，通过对因变量取对数的方法，能够直观地观察各因素之间的相关性。对变量取对数并不影响模型的预设，但改变了对模型系数的解释：仅对因变量取对数的情况下，其他条件不变，自变量 X 的 1 个单位的变化可以引起因变量 Y 的 β（系数）$*100\%$ 的变化；对自变量、因变量均取对数的情况下，X 的 1% 的变化可以引起 Y 的 $\beta\%$ 的变化。当自变量为名义变量时，自变量从 0 到 1 可以引起因变量 Y 的 $100\ (e^{\beta}-1)\%$ 的变化。

（续表 1）

	模型一	模型二	模型三
未缴犯罪所得	0.0471(0.0853)		
谅解退赃退赔缴回	-0.0524(0.0288)	-0.0454(0.0287)	-0.0489(0.0280)
自首	-0.375***(0.0945)	-0.371***(0.0944)	-0.365***(0.0942)
坦白	-0.137**(0.0435)	-0.136**(0.0430)	-0.127**(0.0425)
立功	-0.244(0.176)		
当庭认罪	-0.0884(0.0524)	-0.0886(0.0520)	-0.0766(0.0513)
常数	0.483***	0.485***	0.492***
样本总数〔30〕	852	852	852
Adjusted R^2	0.543	0.541	0.541

注：括号内为标准差

$^{*}p<0.05$，$^{**}p<0.01$，$^{***}p<0.001$

表 1 中的模型一包含所有记录的自变量，模型二包含所有达到统计数量的自变量以及未达到统计数量但仍有显著性的自变量，模型三则去除了模型二中的非法定量刑因素。〔31〕模型的 Adjusted R^2 均在 0.54 左右，拟合优度较好。去除非法定量刑因素（行政区、是否有辩护人、被告人是否外来人口），模型的 Adjusted R^2 并无改变，意味着模型涵盖的所有量刑情节能解释约 54% 的量刑，且非法定量刑因素不仅没有显著性，解释力也很小。这佐证了假设 B，即量刑指导意见可以解释大部分的有期徒刑量刑。值得一提的是，盗窃数额的一元模型的 Adjusted R^2 为 0.36，意味着这一个情节就解释了 36% 的有期徒刑量刑，说明法官的量刑是以盗窃数额为核心的。

假设 A 的检验需具体讨论各量刑情节。经稳健性检验，非法定量刑因素与有期徒刑之间无显著相关关系。盗窃数额、次数、是否入户、扒窃这些犯罪构成事实，以及前科、累犯、被害人弱势这些趋重情节均表现出显著性。但在趋轻情节中，仅自首和坦白体现出显著性。

表 2 列举了量刑指导意见所规定的各情节的基准刑调节幅度。

表 2　量刑指导意见规定的基准刑调节幅度

前科	可以增加基准刑的 10% 以下
谅解退赃退赔缴回	可以减少基准刑的 20%—40% 以下
当庭认罪	可以减少基准刑的 10% 以下
累犯	应当增加基准刑的 10%—40%
自首	可以减少基准刑的 40% 以下；犯罪较轻的，可以减少基准刑的 40% 以上或者依法免除处罚（即减少基准刑的 100%）
坦白	可以减少基准刑的 20% 以下

基于表 1 的系数及注 29 给出的有关名义变量系数的解释方法，在其他因素不变的情况下，存在前科情节，增加了被告人约 9% 的刑期；存在累犯情节，增加了约 34% 的刑期；存在自首

〔30〕 这里的样本总数差别是对盗窃数额、有期徒刑取对数及剔除共犯样本所造成的。

〔31〕 如果具备某自变量的样本未达到统计数量，那么，该自变量不显著，可能仅因样本数太少，故模型一仅起参考作用。然而，如果某些量刑情节很少出现，但作用十分显著，这种显著性往往是稳定的，因此这些量刑情节可以继续留在正式模型中，但对其系数进行解释时仍应慎重。

情节，减少了44%的刑期；存在坦白情节，减少了约13%的刑期。结合表2可以发现，在调节幅度的选择上，法官对趋重情节的适用更为大胆。并且，“谅解退赃退赔缴回”和“当庭认罪”这两个趋轻情节的作用并不显著。但是，据表2，它们的作用力最起码不应低于前科。根据有期徒刑模型报告的结果，首先，存在不显著的量刑情节，这意味着量刑不均是有可能的。其次，上述发现说明，我国法官在趋轻情节的适用上较趋重情节保守。不过，不显著的趋轻情节以“可以”适用的情节为主，所以法官在形式上并未违背量刑指导意见的规定。但在实质上，量刑情节的作用力排序并不符合量刑指导意见“做到该宽则宽，当严则严”的要求。当然，“可以”适用的规定根源于刑法，量刑指导意见只能在刑法的规定之下进行具体幅度的规定。不过，未来对量刑指导意见的修正，应当考虑到法官保守适用趋轻情节的司法现实。如果忽略这一点，量刑指导意见的实际效果就可能会偏离期待，即虽然结果满足形式要求，实质上却只是“当严则严”，该宽时不够宽，从而无法有效因应重刑主义的量刑实践。

（二）罚金刑模型

表3　罚金刑线性回归模型

	模型一	模型二	模型三
行政区	-0.0611(0.0435)	-0.0675(0.0431)	
辩护人	0.0219(0.0717)	0.0340(0.0705)	
外来人口	-0.0414(0.0523)	-0.0466(0.0518)	
盗窃数额	0.229***(0.0139)	0.230***(0.0139)	0.232***(0.0135)
盗窃次数	0.272***(0.0406)	0.269***(0.0403)	0.270***(0.0402)
入户	0.184***(0.0467)	0.187***(0.0466)	0.181***(0.0464)
扒窃	0.127*(0.0529)	0.133*(0.0528)	0.116*(0.0519)
携带凶器	0.125(0.225)		
前科	0.173***(0.0517)	0.170**(0.0514)	0.170***(0.0513)
累犯	0.163***(0.0374)	0.165***(0.0373)	0.165***(0.0372)
特殊身体状况	0.144(0.228)		
限制刑事责任能力人	-0.627(0.506)		
被害人弱势	-0.418(0.503)		
破坏性盗窃手段	-0.352(0.359)		
未缴犯罪所得	0.226(0.141)		
谅解退赃退赔缴回	0.0463(0.0474)	0.0390(0.0470)	0.0374(0.0460)
自首	-0.296*(0.144)	-0.296*(0.144)	-0.275(0.143)
坦白	-0.234**(0.0710)	-0.231**(0.0700)	-0.222**(0.0692)
立功	-0.718*(0.290)	-0.721*(0.290)	-0.708*(0.290)
当庭认罪	-0.182*(0.0858)	-0.181*(0.0850)	-0.170*(0.0839)
常数	5.607***	5.611***	5.537***
样本总数[32]	859	859	859
Adjusted R^2	0.356	0.355	0.355

注：括号内为标准差

$^{*}p<0.05$，$^{**}p<0.01$，$^{***}p<0.001$

〔32〕 这里的样本总数差别是对盗窃数额、罚金取对数及剔除共犯样本所造成的。

表 3 报告的模型与有期徒刑模型的建模思路一致，模型的 Adjusted R^2 均约为 0.35，即所有法定因素能解释约 35% 的罚金刑量刑。这一结果并不符合假设 B，量刑指导意见对罚金刑的规制较弱。此数据较有期徒刑模型低，说明罚金刑在量刑上有更大的任意性。与有期徒刑模型类似，盗窃数额这一个情节就可以解释 28% 的罚金刑量刑，是解释力最强的因素。

对假设 A 的检验，各因素的显著性基本与有期徒刑模型一致。与有期徒刑模型不同的是，立功、当庭认罪等趋轻情节表现出显著性，而被害人弱势因素不再显著。基于表 3 的系数及注 29给出的有关名义变量系数的解释方法，累犯的作用大幅降低，只能增加约 17% 的罚金刑（有期徒刑模型中约 34%）；前科的作用大幅增加，达到 18% （有期徒刑模型中约 9%）；自首的作用有所减轻，约为 33% （有期徒刑模型中约 44%）；坦白的作用大幅增加，约为 25%（有期徒刑模型中约 13%）；当庭认罪的作用幅度约为 20%。

综而言之，罚金刑量刑整体上较轻。相较有期徒刑模型，有更多趋轻情节体现出显著相关性，并且趋轻情节的作用幅度也较大。但是，罚金刑量刑中递进的量刑情节（当庭认罪、坦白、自首；前科、累犯），幅度区别较小，一些情节的作用幅度甚至超越了量刑指导意见所容许的范围。另外，罚金刑模型的解释力较有期徒刑模型约少 19%。量刑指导意见以有期徒刑量刑为纲，其应当如何指导罚金刑量刑，或者如何构建罚金刑指导意见，乃未竟之难题。

（三）两个模型的进一步分析

回顾实证研究的结果，可以推断出量刑指导意见确实在影响着司法实践，但这种影响并不尽如人意。其一，有期徒刑量刑中，趋重情节的作用力大于趋轻情节。其二，量刑指导意见对罚金刑指导不足，其列明的量刑情节对罚金刑量刑的解释力较小，作用幅度也不理想。本文采用访谈方法，继续探究背后的原因。

1．趋轻情节与有期徒刑量刑

应当说明的是，法官的量刑步骤经常是“反推”（先估堆结果，再用量刑表确认，或者只估堆而完全不用量刑表）而非“正推”（通过量刑表计算出结果）。估堆的根据是法院过往判决的类似案件。这就意味着，法官对类似案件的判断、对类似案件量刑的印象、数学能力、估算方法、一时的手松手紧，均会影响最终的判决结果。这些因素均未体现在判决书中，是模型的不可观测项。

那么，不显著的量刑情节究竟缘何而成？有期徒刑量刑中，法官为什么偏重使用趋重情节，而非趋轻情节？首先，量刑情节不显著，并非量刑指导意见规定某些情节“可以”适用而非“应当”适用所导致。实际上，法官在访谈中表明其并不特别区分“可以”适用和“应当”适用，只要认定案件存在量刑指导意见所规定的情节，就会做相应的调整。对模型结果的一种有力的解释是，这可能是量刑起点靠近法定最低刑所造成的。例如，数额较大的盗窃罪，按照量刑指导意见，顶格量刑起点为 1 年有期徒刑，那么趋轻情节下，量刑为 1 年有期徒刑以下；趋重情节下，量刑为 1 至 3 年有期徒刑。这意味着在有期徒刑范围内，趋轻情节的调整空间仅为趋重情节的四分之一（6 个月 v. 2 年）。在这种情况下，趋重情节的实际调整幅度很可能大于趋轻情节。试想某法官承办了一起盗窃案件，其将数额较大的盗窃罪的量刑起点确定为 6 个月有期徒刑。如果该案同时具备当庭认罪、退赃退赔这两个趋轻情节，但该法官认为判处有期徒刑合适，他就必须对这两个趋轻情节的作用做打折处理。但是，如果该案具备的是前科、被害人弱势等趋重情节，法官完全可以按照其认为应当适用的幅度量刑，而无需对趋重情节的作用幅度打折。更何况，量刑指导意见规定的常见量刑情节大部分是趋轻情节，在调整空间本就局促的情况下，各趋轻情节必然要牺牲一定的作用幅度。这种结构性缺陷难以简单地

通过增加趋轻情节的数量、扩大其作用幅度来弥补。2021 年量刑指导意见在基本结构没有变化的前提下增加了数个趋轻情节，这能在一定程度上促进刑罚轻缓化，但应当注意的是，其作用效果在量刑实践中可能会被打折。

同时，法官的主观认知不完全等同于客观行为。当量刑结果主要由估堆得出时，法官“认为”自己从轻了，但实际上可能并没有。法官估堆的依据是自己过去处理的类似案件，即便是新任法官，其量刑也要依据所任职法院的集体经验。这些经验结果即便形式上符合量刑指导意见，其形成方式也可能并不与量刑指导意见完全一致。

2．罚金刑量刑指导体系的缺位

完美状态下，所有量刑情节应能解释绝大部分的量刑结果。不过，量刑过程中的变量纷繁复杂，且很多变量无法被记录，所以模型拟合值小是正常的。然而，在有期徒刑模型的拟合度达 54% 的情况下，罚金刑模型的拟合度却只有 35%，有着 19% 的差距。同时，在罚金刑量刑中，体现出显著性的量刑情节在作用幅度上，也不完全符合量刑指导意见的要求。

就量刑指导意见的内容看，虽然它确实着重于对有期徒刑量刑的规定，但并未否定其对并处罚金刑的适用。就此而言，量刑指导意见应当对有期徒刑和罚金刑的量刑发挥同等作用，但不可否认的是，将指导意见直接适用于罚金刑有其困难。在罚金刑上，法院并不存在对量刑起点与基准刑的共识。法官的回应是，罚金刑量刑的主要考虑因素并非量刑情节，而是罚金刑与自由刑的对应关系。某法官表示，1 年自由刑通常对应 2000 元罚金。这种对应并无法律依据，只不过是法院的内部经验。因此，这种对应并不绝对，该法官也曾对 1 年有期徒刑搭配并处 3000 元罚金。这一点也被实证数据所证实：有期徒刑与罚金刑一元模型的 Adjusted R^2 为 31.5%，远超盗窃数额的解释力。

另外，罚金刑量刑有更多影响量刑的不可观测项。对法官的访谈表明，在罚金刑量刑中，被告人的给付能力是仅次于盗窃数额的影响因素。倘若认为给付能力仅影响罚金刑量刑，且其他因素对罚金刑和有期徒刑量刑的影响程度一致，则给付能力对罚金刑量刑的影响可达到两个模型的 Adjusted R^2 之差，即 19%。在给付能力并未被量刑指导意见正式认可的情况下，司法实践注重对被告人给付能力的考量，根源于罚金刑的执行难问题，是量刑对执行妥协的结果。[33] 毕竟，罚金刑的执行率畸低。[34] 其中，流动人口犯罪的罚金刑执行难问题尤为突出。在本研究中，被告人中的 87% 是外来人口，且从显示学历、职业等被告人信息的部分判决书看，大部分被告人为无业人员或农村人口。

除上述以外，量刑指导意见是以有期徒刑量刑为核心构建的，其对罚金刑的指导意义不强。也正因此，法官在罚金刑的量刑中更容易受法外因素的影响。所以，建立罚金刑的量刑指导体系是有必要的。

四、量刑指导意见的深层次问题

基于实证研究结果，本文通过比较研究，继续检讨对量刑指导意见进行结构性调整的必要

〔33〕 2021 年量刑指导意见明确了罚金刑量刑应当考虑被告人的给付能力，但未明确其具体的考量方法。

〔34〕 例如，自 1997 年刑法实施以来，除 2005 年以外，1997 年至 2008 年陕西省西安市莲湖区人民法院的罚金刑执行率均在 18% 以下，其中有六年的执行率不到 10%。参见翟俊杰、冯妍：《罚金刑执行中存在问题及对策》，http://xalhfy.chinacourt.gov.cn/article/detail/2009/12/id/3079036.shtml，2021 年 7 月 11 日最后访问。

性和方法。

（一）量刑起点问题

量刑指导意见赋予法官的刑罚裁量幅度较大，其并未将量刑起点规定为一个具体的点，而是将其规定为一个幅度：盗窃数额较大的，在 1 年以下有期徒刑、拘役幅度内确定量刑起点；达到巨大起点的，在 3 至 4 年有期徒刑幅度内确定量刑起点；达到特别巨大起点的，在 10 至 12 年有期徒刑幅度内确定量刑起点。通常，法院对基本犯罪构成事实所对应的量刑起点存在共识，例如，某法院普通盗窃 3000 元所对应的量刑起点通常为 6 个月有期徒刑。如果按照通常量刑起点计算的结果无法让法官感到公正，一个便捷有效的做法就是调整量刑起点，例如从 6 个月调整到 8 个月，甚至 12 个月。实际上，笔者比较了一位法官提供的量刑表后发现，案件类似而量刑起点不一的情况并不罕见。

实践中，量刑起点往往是法官估堆反算的结果。由于量刑起点被规定为一个较大的幅度，所以估堆结果往往符合量刑指导意见的形式要求。这意味着量刑指导意见的实质约束力或许不足，趋轻情节的作用不显著也从侧面佐证了这一点。若法官对公正刑罚的经验一致，且贴近理想的量刑基准，问题也称不上严重。但棘手之处正是法官对公正的把握存在差异，且法官的能力、水平也难称一致。在这种情况下，一个被告人可能仅因分配的法官不同而在最终量刑上有数月之别。

面对司法实务从量刑终点反推量刑起点的现实，笔者赞同张明楷的意见，即“点的理论”于当今中国较“幅的理论”更优，因为“点的理论”能带给法官更为明确的指引，能使量刑指导意见的约束落到实处。[35] 实际上，法官从过往经验中获得确信，直接得到所处理案件的初步结论，并不当然是一件坏事。如若有良好的制度将此种经验上的确信规范在理想的量刑结果附近，自拟定的宣告刑出发，利用基准刑和量刑情节获得确信并进行修正，不失为一种经济便宜的做法，而只有精确的“点”才能使这种修正不沦为形式。实际上，法院内部对基本犯罪构成事实所对应的量刑起点的共识说明，法官在司法实践中已经对量刑起“点”有了一个粗浅的把握。英国的量刑指南就是将量刑起点确定为一个具体的点，并在此基础上依据量刑情节确定一个将该点包含在内的量刑上下限。例如，当被告人具备较高的有责性（例如是共同犯罪的主犯），且盗窃数额在 10 万英镑以上时，量刑起点是确定的 3 年 6 个月监禁刑，量刑范围是 2 年 6 个月到 6 年。[36]

（二）以盗窃数额为基础的问题

无论是有期徒刑还是罚金刑，盗窃数额都是影响量刑最重要的情节。可以说，盗窃罪的量刑指导是以数额为基础构建起来的。但是，类似我国规定普通盗窃情形达到数额较大起点方构成犯罪的做法，并非通常现象。在许多法域，盗窃数额只是区分轻罪与重罪的依据，而不是区分罪与非罪的材料。[37]

以广州市为例，盗窃数额与量刑起点的对应关系如表 4。从表 4 可见，我国盗窃罪量刑体系确定的三个基本点（数额较大、巨大、特别巨大）对应的量刑幅度没有重叠，这使得数额

〔35〕 参见前引〔17〕，张明楷文，第 130 页以下。

〔36〕 See Sentencing Council, *Theft General Guidelines* (*2016*), https://www.sentencingcouncil.org.uk/offences/magistrates-court/item/theft-general.

〔37〕 如美国肯塔基州修订法规第 514.030 节（KRS § 514.030）、密苏里州修订法规第 570. 030 节（RSMo § 570.030）、纽约州刑法第 155.25 节（NYPC § 155.25）均不要求达到特定数额才构成犯罪。

临界点的边际刑罚量在相邻区间内最大；而且数额越大，单位数额所背负的刑罚量越小（即第四栏数字的倒数），减小的曲线十分陡峭。简言之，当前的效果是，犯罪门槛的震慑力很高，而一旦行为人跨过门槛开始犯罪，门槛之后的震慑极其不足，并未形成逐步升高的震慑阶梯。

表 4　广州市盗窃数额与量刑起点的对照

法定刑档次	盗窃数额（元）	量刑起点	有期徒刑每月所对应的盗窃数额（元）
数额较大	3000	1 年有期徒刑以下	250—500
	3000—10 万	每 15000 元可以增加 3 个月至 6 个月，至高不超过 3 年	2500—5000
数额巨大	10 万	3 年至 4 年	2083.3—2777.8
	10 万—50 万	每 50000 元可以增加 6 个月至 1 年，至高不超过 10 年	4166.7—8333.3
数额特别巨大	50 万	10 年至 12 年	3472.2—4166.7
	50 万—100 万	可以增加 1 年以下	>4166.7
	100 万—300 万	可以增加 1 年至 3 年	55555.6—166666.7
	300 万以上	可以增加 3 年以上，或判处无期徒刑	>69444.4

我国当前的量刑规范思路是，确定的自变量（盗窃数额）对应确定的因变量（量刑起点范围），自变量、因变量均不重叠。这是因为我国刑法也规定了几乎不相重叠的法定刑档次，例如 3 年以下有期徒刑、3 至 10 年、10 年以上，仅在分界线上重叠。因此，边际数额具有决定性作用。然而，3000 元与 2999 元这一元之差，难以承担区分法定刑档次的重责。

对于数额临界点的问题，英国的处理方式是使临界点左右所对应的刑期互有重叠，而不是在临界点制造一个泾渭分明的界线（表 5）。根据量刑指南的使用方法，“量刑者应当确定这些因素（指量刑情节）或其他相关因素的组合对于量刑起点向上或向下的调整……在某些情况下，超出量刑范围的刑罚幅度可能是适当的。”〔38〕由此，表 5 中的量刑范围指代广义的量刑基准，而非基准刑。只有在例外的情况下，量刑基准才能超出量刑范围。边界点左右量刑范围的重叠给了法官更大的裁量余地：边界点的作用不再是决定性的，其更像是一个“路标”。美国明尼苏达州的量刑指南也采纳了类似的思路（表 6）。

表 5　英国量刑指南关于高有责性盗窃罪被告人量刑起点和量刑范围的设置〔39〕

数额大于 10 万英镑	量刑起点：3 年 6 个月；量刑范围：2 年 6 个月—6 年
1 万—10 万英镑	量刑起点：2 年；量刑范围：1 年—3 年 6 个月
500—1 万英镑	量刑起点：1 年；量刑范围：26 周—2 年
不大于 500 英镑	量刑起点：高强度的社区刑；量刑范围：中强度的社区刑—36 周

由此不难看出，通过使刑罚档次边缘范围重叠，可以缓和边际数额的决定性作用。这与前

〔38〕 Sentencing Council, *Using Sentencing Council Guidelines*, https://www.sentencingcouncil.org.uk/explanatory-material/magistrates-court/item/using-the-mcsg/using-sentencing-council-guidelines.

〔39〕 参见前引〔36〕。

述量刑起点应当为一个具体的点并不矛盾。因为量刑起点的问题是，当其为一个幅度时，量刑指导意见的约束力不足，所以应当补足其约束力。而边际数额的问题则是确定性太强，应当予以一定的缓和。

表 6　美国明尼苏达州量刑指南关于盗窃罪量刑起点和量刑范围的设置〔40〕

犯罪历史分数 / 数额（美元）/罪行等级	0	1	2	3	4	5	6 及以上
>5000/3 级	12^1	13	15	17	19 17—22	21 18—25	23 20—27
5000 及以下/2 级	12^1	12^1	13	15	17	19	21 18－25

注：刑期的单位为月。阴影部分为推定缓刑，非阴影部分为推定实刑。

12^1 = 1 年零 1 天

难题在于我国的法定刑档次有确定的分野，故无法对边际数额左右的量刑范围进行调整。如果无法直接对刑法进行修正，仍有解释空间的只有“数额较大”“数额巨大”“数额特别巨大”的含义。一个次好的方法是将数额较大、巨大、特别巨大的边缘范围重叠起来。例如，如果原本认为数额较大、巨大、特别巨大的范围应为 3000—10 万元、10 万—50 万元、50 万元以上，那么可以探索将其范围修改为 3000—12 万元、10 万—55 万元、50 万元以上。当然，容许多大的重叠范围可以达到法官自由裁量与规范确定作用的平衡，需要继续研究。这种方法的疏漏之处是，只要我国盗窃罪的成立依然需要达到一定的盗窃数额，就很难从技术上缓和该特定数额的决定性作用。因而，应当考虑直接修正刑法的可能性。2020 年美国明尼苏达州法典第 609.52 节第 3 分项对盗窃罪就仅规定了法定刑的上限。我国如果贯彻责任主义的思考方向，在法定刑中仅规定上限，反而可以使量刑指导意见更为系统且灵活地满足刑事政策的需要。

对于单位数额所背负的刑罚量随着数额增加而减少这个问题，首先需要说明这是各国共同面临的问题。此种情况似乎有悖罪刑均衡原则。〔41〕由于人的生命有限（且我国的有期徒刑存在上限）而可能的盗窃数额无限，盗窃罪的自由刑量刑必然带有单位数额的刑罚量随着数额增加而减小的特质。一种可能的改良方案是，在自由刑贬值的区间内搭配更严厉的罚金刑以达到惩戒目的。

（三）量刑步骤问题

量刑指导意见列明的盗窃罪量刑步骤为：（1）根据基本犯罪构成事实确定量刑起点；（2）根据其他影响犯罪构成要件的事实（盗窃次数、数额、结果等）在量刑起点的基础上增加刑罚量，确定基准刑；（3）根据量刑情节调整基准刑，并综合全案确定宣告刑。照此方法，一个通常的盗窃罪量刑公式可以被理解为：

$$\text{宣告刑}_i = \frac{\text{量刑起点}_i\text{（步骤一）} + \text{其他影响犯罪构成要件的事实}_i\text{（步骤二）}}{\text{（基准刑）}} \pm \text{量刑情节}_i\text{（步骤三）}$$

〔40〕 See Minnesota Sentencing Guidelines Commission, *Minnesota Sentencing Guidelines and Commentary 2019*, p. 79, https://mn.gov/msgc-stat/documents/Guidelines/2019/MinnSentencingGuidelinesCommentary.pdf. 需要注意的是，在我国定为盗窃罪的，在美国可能是单独罪名，例如扒窃（theft from person），刑罚等级为 4 级。

〔41〕 这并非盗窃罪的专利，在其他犯罪中也存在类似现象。参见林竹静：《受贿罪数额权重过高的实证分析》，《中国刑事法杂志》2014 年第 1 期，第 114 页。

这是一个历经三步，从量刑起点到基准刑再到量刑终点（宣告刑）的方案。具体到个案，在普通盗窃8000元的情形下（以3000元为数额较大的起点），依据量刑指导意见，第一步为使用盗窃3000元为确定量刑起点的事实，在第二步用剩余的5000元作为调节刑罚量、确定基准刑的事实。但是，量刑指导意见并未明确确定基准刑的具体方法。狭义的量刑基准针对抽象的个罪，是不考虑量刑情节的前提下，仅就犯罪构成的基本事实所判处的刑罚。[42] 如果认为狭义的量刑基准是依据个案的整个犯罪构成事实而确定的，例如，应当对“普通盗窃8000元”进行整体性地量刑基准考量，则狭义的量刑基准论会要求合并考虑第一、二两个步骤。如果认为量刑基准应依据基本犯罪构成事实来确定，则狭义的量刑基准论会主要讨论第一个步骤，即量刑起点问题。

与狭义量刑基准论的两种分野对应，司法实务中对第一、二两个步骤是分开考虑还是合并考虑，也有两种取向。合并考虑的做法是，依照盗窃数额在法定刑区间内按比例确定基准刑。例如普通盗窃8000元，合并考虑和分开考虑的计算方法（若认为数额较大的起点为3000元，量刑起点为6个月；每10000元增加6个月刑期）分别如下：

合并考虑：$\frac{8000}{(10000+3000)}\times(6+6)\approx 7$

分开考虑：6（步骤一）$+\frac{8000-3000}{10000}\times 6$（步骤二）$=9$

合并考虑的好处是，其更加有利于实现刑罚谦抑；并且，每1元数额之于刑罚量的边际价值是一致的，前3000元和后5000元在合并考虑的方案中对刑罚量的影响相同。在分开考虑的场合，一个明显的问题是，在同一个犯罪中，定罪所考虑的数额（前3000元）的边际价值明显高于量刑调整数额（后5000元），而这种价值上的区别并无充分依据。

当然，合并考虑也有问题。一是计算更为复杂，需要分档后再根据各档次的算法进行计算，而分开考虑的算法是固定的。二是合并考虑的算法决定了分界数额以上一定范围内（分母数额的二分之一以下）的盗窃数额都对应着基准刑期。例如在本例中，盗窃3000元至6500元的基准刑就都为6个月。但是，存在这两个问题并不意味着分开考虑就更好。首先，合并考虑的算法并未复杂到法官无法掌握的程度，而且对各档次都固定的算法（即分开考虑的计算方法）往往意味着对公正性的部分牺牲。其次，决定法定刑档次的边际数额以上一定范围内的数额对应基准刑期，并不当然是一件坏事。如上所述，我国用盗窃数额划分法定刑档次的构造，本来就造成了一些问题。合并考虑的算法实质上拉长了边际数额的范围，呼应了前述降低边际数额决定性作用的主张，有利于平衡单位数额所背负的刑罚量。而且，美国联邦量刑指南中初犯的前八个罪行等级，也都对应着同一个量刑范围。但在我国司法实务中，分开考虑的方法占据主流。因此，在修订量刑指导意见时，笔者建议应当明确规定对第一、二两个步骤进行整体性考量。

（四）量刑情节问题

有学者认为，量刑指导意见违背了责任主义，因其将责任情节与预防情节一概而论，未对

〔42〕 参见前引〔17〕，周光权文，第132页；张向东：《从量刑基准到基准刑——量刑方法的革新》，《中国刑事法杂志》2011年第3期，第38页。

二者作考量步骤上的先后区分。[43] 我国量刑指导意见至今有关量刑情节的构建似有英美量刑指南的影子，即不特意区分责任刑情节与预防刑情节，仅做趋轻情节与趋重情节的区分。但是，英美并不在量刑指南中确定各情节的作用范围，而是直接规制量刑结果。英美不给出量刑算法的原因主要在于两个方面：其一，两类情节的考量过于复杂，应当交由法官裁量。其二，对趋重情节和趋轻情节的考量应当着重罪犯的个人改造因素，而英美量刑指南主要基于责任刑情节。但是，有学者明确反对这两个论点，认为在量刑指南中进行明确规定、用集体智慧对量刑情节进行规制，将有助于量刑的统一性。[44] 从这个角度看，我国其实比英美走得更远一些。但我国对量刑情节一概而论的做法，确有值得探讨的地方。

量刑基准理论，无论是“点的理论”还是“幅的理论”，均要求只能在责任刑以下考虑预防刑，故基于有责性的情节（例如携带凶器）是第一位的，基于预防性的情节（例如累犯、自首）是第二位的，实质上形成了量刑上的“二阶层论”。公式表达如下：

$$\text{宣告刑}_i = \frac{\text{构成要件要素}_i\text{（步骤 a1）} \pm \text{责任刑情节}_i\text{（步骤 a2）}}{\text{（责任刑）}}\text{（步骤 a）} - \text{预防刑情节}_i\text{（步骤 b。注意：这里只能是减号）}$$

量刑指导意见与量刑基准理论并非完全不可调和。指导意见所规定的第一、二两个步骤合并起来即是步骤 a1，只需在第三步注意应当先考虑责任刑情节，从而获得责任刑的点，之后基于预防刑情节只能降低刑罚量即可。要想达到这样的目标，只需删去“同向相加、逆向相减”这一机械适用的原则性规定。当然，更理想的解决方案是对量刑指导意见进行彻底调整，将有关责任的情节与有关预防的情节进行分类，并采用先责任刑情节、后预防刑情节的编排模式，使其完全满足在责任刑范围内确定预防刑的理论要求。

在这个基本框架下，针对趋轻情节作用不显著的司法现状，应当考虑对量刑指导意见的内容进行调整。导致这种现状的两个可能原因是，量刑起点靠近法定最低刑以及法官的前见，这二者都无法轻易改变。如果为了让趋轻情节和趋重情节的调整幅度一致而将量刑起点拉升至中点位置，无疑会加重量刑，此乃舍本逐末之举。法官的前见依托于过往审判的案件，不以人的意志为转移。在这两个前提不变的情况下，可以探索如何将拘役和有期徒刑（本质上都是监禁刑）有机地衔接起来，扩大法官的视线范围。同时，为了让量刑指导意见真正限制法官的量刑估堆结果，可以考虑缩小量刑情节的幅度范围，真正区分“可以”适用和“应当”适用的情节。例如，将当庭认罪情节的量刑调整范围规定为“可以减少基准刑的 5%—10%”。这样一来，在有当庭认罪情节的案件中，法官既有根据案件实际情况不减少刑罚量的灵活选择，同时也能受到有效约束，因为其一旦选择减少刑罚量，就必须体现出该情节应有的分量。

（五）罚金刑的量刑指导问题

过往对罚金刑的研究集中在执行难与罚金刑的适用范围这两个问题上。[45] 我国量刑指导意见主要着眼于有期徒刑，在罚金刑量刑的指导上作用不突出。罚金刑模型较低的 Adjusted R^2

〔43〕 参见前引〔17〕，张明楷文，第 139 页。

〔44〕 See Andreas von Hirsch, *Foreword*, in Julian V. Roberts (ed.), *Mitigation and Aggravation at Sentencing*, New York: Cambridge University Press, 2011, pp. xiii – xvi.

〔45〕 参见冯亚东：《罪刑关系的反思与重构——兼谈罚金刑在中国现阶段之适用》，《中国社会科学》2006 年第 5 期；钱叶六：《论中国罚金刑的改革与完善——以探寻罚金刑执行难之解决方案为视角》，《法学论坛》2006 年第 4 期；周光富：《罚金刑执行难之克服》，《政治与法律》2003 年第 6 期。

也佐证了这一点。2021 年量刑指导意见发布之前，量刑指导意见并未对罚金刑作明确指引，相关规定散见于个罪的司法解释。例如，盗窃罪罚金刑的量刑指导见于 2013 年最高人民法院、最高人民检察院《关于办理盗窃刑事案件适用法律若干问题的解释》第 14 条："因犯盗窃罪，依法判处罚金刑的，应当在一千元以上盗窃数额的二倍以下判处罚金；没有盗窃数额或者盗窃数额无法计算的，应当在一千元以上十万元以下判处罚金。"该条为 2021 年量刑指导意见所吸纳。类似地，该意见在部分常见犯罪中（如抢劫罪、诈骗罪）也设置了相应的罚金刑条款，但遗憾的是，这些条款较为宽泛，并未给出如何适用罚金刑的具体指引。

张明楷认为，量定罚金数额时应当考虑以下因素：犯罪情节、支付能力、行为人将来的职业状况与其他情况。[46] 高铭暄等人认为："罚金刑数额立法原则应遵循数额相对确定原则、罚金数额适度原则、与自由刑形成比值原则、不断适应通货膨胀原则、相对平等原则。"[47] 学界过往曾对判处罚金刑时是否应当考虑被告人的给付能力进行过探讨。[48] 值得肯定的是，2021 年量刑指导意见明确了"应当以犯罪情节为依据，并综合考虑被告人缴纳罚金的能力，依法决定罚金数额"，承认了被告人给付能力的作用与独立价值。

然而，2021 年量刑指导意见并未提供考量被告人给付能力的方法。这可能导致对其的考量会让位给各法院的实际需要，使得以提高执行率为目的的实用逻辑取代以促进公平公正为主的道德逻辑，加重量刑不均。首先，在没有道德逻辑指引的情况下，法院自发适用的实用逻辑可能会有很大的偏差值。其次，野蛮生长的实用逻辑并不一定能解决实践中的问题，因为这种逻辑能够成立的前提是：当罚金较低时，被告人更可能如期缴纳罚金；如果罚金完全超出被告人的支付能力，被告人当然不可能缴纳罚金。但值得注意的是，法院因此而调低的罚金额，或许并不是从一个超出给付能力的罚金额调整到匹配其给付能力的罚金额，而是在两个匹配被告人给付能力的罚金额中选择较低的那一个。[49] 例如在本文样本中，罚金额大部分在 2000 元以下，这个数额似乎并未超出一般被告人及其家庭的给付能力。即便如此，罚金刑执行难的问题依然存在。因此，若任由罚金刑量刑的实用逻辑野蛮生长，则可能既降低了罚金刑量刑的公正性和严肃性，也不能从根本上改善罚金刑的执行。

综而言之，我国有必要建立更为细致的罚金刑量刑体系。具体到盗窃罪，英国的做法是将可能的罚金分为六组，每组确定量刑起点和量刑范围。如 A 组的量刑起点是 50% 的周薪，量刑范围是 25%—75% 的周薪；B 组的量刑起点是 100% 的周薪，量刑范围是 75%—125% 的周薪。但是，英国的罚金是单处罚金，在盗窃罪量刑中主要适用于低有责性的被告人。[50] 而且，英国仅在例外的情况下可以考虑并处罚金。[51] 相较之下，美国的做法与我国更为相似。根据美国联邦量刑指南第 5E1.2（a）节的规定，罚金为必并罚金，"法院应当在所有案件中

〔46〕 参见张明楷：《刑法学》上，法律出版社 2016 年版，第 536 页。

〔47〕 高铭暄、孙晓：《宽严相济刑事政策与罚金刑改革》，《法学论坛》2009 年第 2 期，第 5 页。

〔48〕 持否定意见的观点，参见刘明祥：《论解决罚金刑执行难题的立法途径》，《法学家》2009 年第 2 期，第 102 页以下。

〔49〕 罚金刑执行难是因为罚金超出了被告人的实际给付能力这一论断，尚无大数据支持。即便有零星个例，也难以表明这种情况的普遍性。参见赵一荣：《刑事案件中财产性判决"执行难"问题研究》，《法律适用》2019 年第 23 期，第 129 页。

〔50〕 参见前引〔36〕。

〔51〕 See Sentencing Council, *Imposition of Fines with Custodial Sentence*, https://www.sentencingcouncil.org.uk/explanatory-material/magistrates-court/item/fines-and-financial-orders/approach-to-the-assessment-of-fines-2/11-imposition-of-fines-with-custodial-sentences.

判处罚金，除非被告人证明他现在且未来并无任何支付能力。”联邦量刑指南为盗窃罪提供了统一的罚金表，部分内容见表 7。

表 7　美国联邦量刑指南盗窃罪罚金表〔52〕

罪行等级	罚金最小额（美元）	罚金最大额（美元）
3 以下	200	9500
4—5	500	9500
6—7	1000	9500
8—9	2000	20000
10—11	4000	40000
12—13	5500	55000
14—15	7500	75000
16—17	10000	95000
18—19	10000	100000

在接受罚金刑量刑时被告人的给付能力应当发挥调节作用的基础上，美国盗窃罪罚金表所提供的启示是：应当区分盗窃数额档次并确定其对应的罚金范围；考虑到被告人的给付能力天差地别，罚金范围应当设置的比较大，罚金范围也应互有重叠。同时，为了保证刑罚与社会危害性的最基本均衡，应当谨慎设置罚金档次的最小值。当然，如果被告人能够证明其现在与未来并无给付能力（如其本身即需要他人扶养的重度残疾者），那么可以超出本档次最小值量刑。考虑被告人给付能力对罚金刑量刑的调节作用，意味着需加强相关配套措施，如被告人银行账户、经济收入的审查等。从精细化司法的角度出发，进一步的措施是允许判处有整有零的罚金。本样本中，罚金相差不低于 500 元，多为 1000 元以上，强求整数罚金反而可能丧失罚金刑本身的灵活性。

当罚金刑与自由刑并处时，还需考虑其搭配问题。美国联邦量刑指南规定，“在量定罚金刑时，必须保证其数额与其他刑罚的搭配是惩罚性的”，〔53〕其采用的方法是将各罪划分等级，使各等级对应不同的自由刑和罚金刑。由于我国没有对各罪划分等级，所以很难直接借鉴美国的方法。实践中 1 年自由刑对应 2000 元罚金的经验估堆方法显然过于粗糙，如果量刑指导意见无法给出单独的罚金刑量刑指导，最起码也要对如何搭配罚金刑和自由刑这一问题提供明确指引。

结 论

本文以盗窃罪为样本实证性地探究了我国量刑指导意见下法官的量刑实践，发现量刑指导意见尚缺乏对罚金刑的充分指导；而在有期徒刑的量刑中，趋轻情节的作用不如趋重情节。这些发现表明，量刑指导意见虽然在我国的量刑规范化进程中发挥了重要作用，但仍有很大的改进空间。在使用量刑指导意见促进量刑均衡这一道路上，我们必须不断从实践中汲取养分，及

〔52〕 United States Sentencing Commission, Guidelines Manual 2018, § 5E1.2 (c) (3), at p. 437.

〔53〕 United States Sentencing Commission, Guidelines Manual 2018, § 5E1.2 (d), at p. 438.

时对指导意见进行调整。

本文在实证研究的基础上，结合与英美量刑指南的比较，对量刑起点、以数额为量刑中心、量刑步骤、量刑情节、罚金刑的量刑指导等这些量刑指导意见的深层次问题，进行了细致检讨。对于量刑起点，本文综合实践情况，建议将其规定为一个具体的点而非幅度。在数额问题上，本文借鉴英美量刑指南的构造，建议通过重叠边际数额范围的方式，模糊边际数额所背负的单位刑罚量，以缓和边际数额的决定性作用。在量刑步骤上，本文基于单位刑罚量的考量，建议合并考虑确定基准刑的两个步骤。在量刑情节上，本文认同在责任刑范围内确定基准刑的做法，并针对法官在趋轻情节适用上的保守，建议量刑指导意见作出更为积极的调整。在罚金刑的量刑指导上，本文吸取域外经验，呼吁构建更细致的罚金刑量刑体系。

Abstract: Although studies on sentencing guidelines are abundant in China, no empirical research has been carried out to quantitatively evaluate the actual effects of the guidelines. As a result, the existing doctrinal researches are unable to provide the legislative techniques necessary for the improvement of the guidelines. This article empirically examined the judicial practice of sentencing guidelines through the lens of crime of theft. The Stata empirical analysis shows that the effect of mitigating circumstances is generally weaker than that of aggravating circumstances in the sentencing of fixed-term imprisonment. Moreover, the sentencing guidelines failed to provide sufficient guidance on fines. On the basis of the above analysis, this article draws on the legislative techniques of the British and American sentencing guidelines, and put forward suggestions for improving the sentencing guideline system in China. Firstly, the starting point for sentencing should be determined as a specific point instead of a range; secondly, the decisive role of the marginal amount should be reduced; thirdly, when determining the benchmark sentence from the starting point, the amount should be considered together; fourthly, in the arrangement of sentencing factors, those relating to retribution should be determined before those relating to deterrence, so as to limit the final sentence under just deserts; and finally, it is necessary to strengthen the guidance on fines and to ensure that the fines, taken together with other sanctions imposed, are punitive in nature.

Key Words: sentencing guidelines, theft, sentencing benchmark, fines