基于边际理论的认罪认罚从宽制度优化研究

石芬琴

（兰州大学 法学院，兰州 730000）

摘要：认罪认罚从宽制度自设立起便展示出巨大的制度优越性，不仅有利于刑事程序的多元化，而且可以促进司法资源的优化配置。但该制度也会给现存司法体制带来挑战，传统法学视角在分析制度的科学合理性、正当性，以及实现效益最大化等方面存在局限，边际理论可以为认罪认罚从宽制度的优化提供理论支撑。文章从法经济学视角出发，采用边际理论，就认罪认罚从宽制度的"公正—效率"和"成本—收益"关系进行分析，并提出将错案率控制在较低水平、完善冤假错案治理的保障机制、协调诉讼程序简化与减刑比例的关系和明确诉讼程序简化的上限等优化构想，以提高刑事处罚的正当性。

关键词：认罪认罚从宽制度；法经济学；边际理论；成本收益

中图分类号：D912.29　**文献标志码**：A　**文章编号**：1672-349X(2025)05-0078-10

DOI：10.16160/j.cnki.tsxyxb.2025.05.011

Study on Optimization of the Leniency System for Pleading Guilty and Accepting Punishment Based on Marginal Theory

SHI Fenqin

(School of Law, Lanzhou University, Lanzhou 730000, China)

Abstract: Since its establishment, the leniency system for pleading guilty and accepting punishment has demonstrated significant institutional advantages. It not only facilitates the diversification of criminal procedures but also promotes the optimal allocation of judicial resources. However, this system also poses challenges to the existing judicial framework. Traditional legal perspectives have limitations in analyzing the system's scientific rationality, legitimacy, and the maximization of benefits, etc., while marginal theory can provide theoretical support for optimizing the leniency system. From the perspective of law and economics, this paper employs marginal theory to analyze the "justice-efficiency" and "cost-benefit" relationships within the leniency system. It proposes optimization strategies, including maintaining a low rate of wrongful convictions, improving safeguards against miscarriages of justice, balancing the relationship between procedural simplification and sentence reduction ratios, and defining the upper limits of procedural simplification, which aim to en-

收稿日期：2023-12-03

基金项目：甘肃省科技计划资助项目(23JRZA372)

作者简介：石芬琴(1999—)，女，湖南湘西人，硕士研究生，研究方向为诉讼法学。

hance the legitimacy of criminal punishment.

Key Words: leniency system for pleading guilty and accepting punishment; law and economics; marginal theory; cost-benefit

一、问题的提出

最高人民检察院 2023 年工作报告显示，2022 年检察机关认罪认罚从宽制度的适用率已超过九成，量刑建议采纳率高达 98.3%，一审服判率为 97%，高出未适用该制度案件服判率 29.5 个百分点[1]。认罪认罚从宽制度自设立以来其适用使得刑事案件上诉率、申诉率大幅减少，这体现了该制度显著的改革新意，并展示出巨大的生命力，有利于罪犯改造、促进社会和谐。传统法学研究侧重于规范研究和实证研究，在理论界，学者们对认罪认罚从宽制度的探讨主要聚焦于法学视角下如何完善制度实施的相关配套措施，如完善权利保障和部门之间的职权制约[2]、解决值班律师角色定位与协商功能[3]等。但是，有关如何促进司法资源的合理配置以实现认罪认罚从宽制度最大效益的研究则相对不足，对制度本身的科学性、合理性和正当性也缺乏令人信服的充分论证。

法律是调整人们在现实社会中行为关系的规则，经济学是研究资源的最优配置的学科。在实践中，法律通常涉及资源的最优配置和利益的分配问题，但仅仅依靠法律制度本身往往会忽略经济效率，无法实现社会效益最大化的目标[4]，因此，法经济学应运而生。法经济学可以效率为尺度衡量法律制度的“优劣”，不仅为法律制度提供了价值评判标准，还为改进法律制度提供了新的指导方法。边际理论是经济学领域的重要原理，用于衡量某个决策或行动的额外收益与额外成本及相互关系，从而探讨如何实现效益最大化。当以边际理论进行法学研究时，可采用一种量化分析的手段，如“成本一收益”分析。此种分析方式能使个体的净收益和社会的净福祉变得清晰可见，从而为法律制度的改进提供明确的指导。

从司法实践来看，认罪认罚从宽制度虽已得到广泛运用，但仍然存在如下问题亟需厘清：其一，在法经济学视角下，认罪认罚从宽与司法公正相比，制度本身的合理性如何得到保证？如何平衡公正审判和提高诉讼效率之间的关系？其二，是要百分百消除错案还是允许一定比例的错案存在？能否为了防范和纠正错案而肯定被追诉人的上诉权？其三，速裁程序能否作为一个独立的减刑情节？减刑比例应如何控制？诉讼程序简化是否会影响到减刑比例？两者之间是何关系？基于以上问题，本文以边际理论作为研究的理论基础，对认罪认罚从宽制度的“公正一效率”和“成本一收益”关系进行分析，并提出相应的优化构想，为制度完善提供参考。

二、认罪认罚从宽制度的“公正一效率”分析

（一）公正和效率：平衡司法资源投入和刑事错案率的关系

法学研究不可避免地会涉及对“公正”和“效率”两个范畴的讨论。经济学更加关注“如何将蛋糕做大”（总量和效率），而法学则侧重于“如何将蛋糕分好”（公正）。然而，从现实来看，公正和效率并非绝对二选一的选项，很多情况下公正需要依赖效率的支持才能真正实现[5]。概言之，在重视公平地“分蛋糕”的同时，也需要在一定程度上考虑如何有效率地“做大蛋糕”，以达到制度构建的最优设计。刑事诉讼贯彻着惩罚犯罪与保障人权两大理念，其更加强调对司法公正的追求，这在一定程度上很容易导致忽视诉讼效率。但是，当今我国正面临着刑事案件激增和审判人员精英化之间的矛盾冲突，以及“案多人少”的司法困境，加之我国正进行的司法体制改革强调不断优化简易程序和速裁程序，如何在保障司法公正的前提下提高诉讼

效率就成为需要重点解决的问题。在法学理论界,甚至有人主张将诉讼效率视为刑事诉讼的首要价值[6]。诚然,公正与效率两者均为刑事诉讼法的追求目标,认罪认罚从宽制度也应始终坚持“均衡”的价值理念,强调公正与效率的统一。但在司法实践中,公正与效率之间的冲突常常难以规避,解决问题的关键在于如何巧妙地化解两者之间的矛盾。

认罪认罚从宽制度的目标为通过有效利用司法资源实现对犯罪行为的惩罚以及对人权的维护。制度的运行需要资源,但资源往往是有限的[7]。在案件数量激增的背景下,司法资源的有限性导致司法主体需要对认罪认罚从宽制度的价值进行分析,探索实现制度效益最大化的途径。法经济学研究的重点在于,如何以有限的资源获得最大的利润,从而实现资源的合理配置。经济学领域的边际理论涉及两个重要的规律:一是“边际收益递减规律”[8],表明当增加某种资源或行为的数量时,其边际收益会逐渐减少,即每增加一单位资源所获得的额外收益会逐渐减少;二是“边际成本递增规律”[9],表明随着生产中产量的增加,每额外生产一单位其成本会逐渐增加。因此,要寻求边际收益和边际成本的“最佳均衡点”。

边际理论中的“公正—效率”模型在法律制度研究中能够提供独特视角,帮助司法主体评估法律制度的运行效果和资源利用效率,并提出改进措施,从而更好地助力刑事司法制度的完善。认罪认罚从宽制度通过司法资源的投入来实现司法公正的目标,如果秉承着过分追求效率的价值倾向就易导致冤假错案的发生,错案率越高,则该制度的收益就越低。因此,认罪认罚从宽制度如何平衡公正和效率的关系实质上即为如何平衡司法资源的投入和刑事错案率之间的关系。

(二)认罪认罚从宽制度的“公正—效率”模型分析

在经济学领域,普遍能够认识到资源是稀缺的;同样,在法学领域,司法资源也是有限的,维护司法体系正常运转需要耗费大量人力和财政资源,司法体制的正常运行面临着高昂成本的挑战[10]。如何以有限的司法资源获得最大的经济效益,法经济学提供了分析资源配置和效率关系的工具和方法,可以运用边际收益和边际成本模型对司法资源的投入量和刑事错案率的关系进行分析,构建认罪认罚从宽制度的“公正—效率”模型(见图 1 与图 2),探索实现效益最大化的方式。

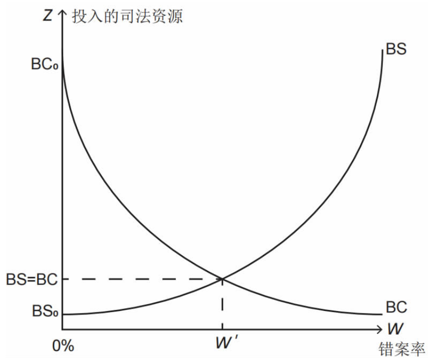

图 1 认罪认罚从宽制度的“公正—效率”模型

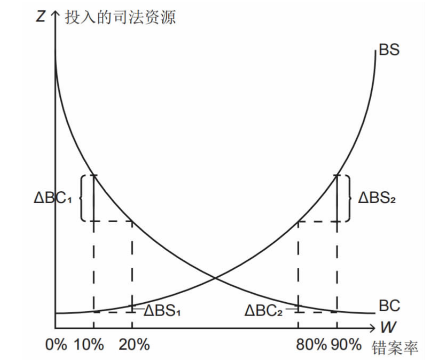

图 2 认罪认罚从宽制度“公正—效率”模型中的典型数值

以错案率(W)为横坐标,以投入的司法资源(Z)为纵坐标,构建认罪认罚从宽制度“公正—效率”模型(见图 1)。曲线 BC 代表减少刑事错案的边际成本,包括国家财政为司法机关工作人员支付的劳动报酬,以及处理刑事案件需要的设备、场所等资源。根据“边际成本递增规律”,若错案率越低,那么欲使错案率进一步降低所需耗费的司法资源和成本就越高。概言之,当错案率处于较高的区间时,一些成本较低且有效的改进措施可以非常显著地降低错案率,此时投入的司法资源相对较少,因此边际成

本较低，即 ΔBC_2 较小（见图 2）。然而，随着错案率的逐渐下降，剩余的刑事错案的纠正难度会不断上升，例如，错案率从 20％降到 10％所需耗费的司法成本比从 90％降到 80％多，即 ΔBC_1 明显大于 ΔBC_2（见图 2）。当错案率逐渐趋近于 0 的理想状态时，往往需要耗费庞大的司法资源才能继续降低错案率，因此，随着错案率的不断降低，边际成本呈现出递增趋势。

曲线 BS 代表减少刑事错案所带来的边际收益，包括对犯罪分子的惩治、司法人权的保障、司法权威的树立以及良好社会秩序的维护。根据“边际收益递减规律”，随着错案率的不断降低，认罪认罚从宽制度带来的持续收益也不断减少。在高错案率的情况下，每一步的降低都能带来显著的改善效果，因此当错案率初始降低时，即使降低 10％，获得的边际收益也是巨大的，即 ΔBS_2 较大（见图 2）。然而，随着错案率的进一步下降，边际收益会逐渐减小，例如，当错案率从 20％降低到 10％时，尽管仍然会带来收益，但该降低幅度所带来的收益相对于从 90％到 80％的降低幅度所带来的收益少，即 ΔBS_1 明显小于 ΔBS_2（见图 2）。这是因为，在错案率较低的情况下，采用法律制度、措施等手段改善刑事错案的空间相对较小，进一步的错案防治可能需要耗费更多的资源，从而造成收益相对较少的情况。

最后，当错案率降为 0 时，虽然边际收益仍然处于正值 BS_0（见图 1），但是此时的收益却难以弥补需要付出的成本。一方面，边际收益已经大大减少，因为当错案率趋近于理想状态时，任何进一步的错案率降低都只能在较小的范围内产生改善效果；另一方面，综合来看，即使边际收益处于正值，但是此时的边际成本已大大超过边际收益，即 BC_0 明显高于 BS_0（见图 1）。因此，根据图中曲线走向可知，一开始将 1％的错案率降低将带来巨大的社会边际收益，然而随着错案率的不断下降，边际收益会逐渐减小，最终接近于零。

要想实现效益的最大化，需要找到一个“均衡点”，即当边际成本等于边际收益时，资源配置达到最优。对于该模型而言，当 BS＝BC 时，即位于 W' 点时，边际收益等于边际成本，是社会减少刑事错案率的最优选择（见图 1）。具体而言，当认罪认罚从宽制度实行后错案率高于 W' 点时，BC＜BS，边际收益大于边际成本，因而应当继续贯彻实行该制度，进一步降低错案率；当认罪认罚从宽制度实行后错案率低于 W' 点时，BC＞BS，边际成本大于边际收益，因而应当停止实行该制度，以避免增加制度运行的成本。因此可以得出两个结论：W' 代表刑事司法的理想状态，错案率在低于 W' 的情况下①，进一步的刑事治理措施会导致边际成本的增加，且相对收益减少，投入更多的司法资源来减少刑事错案将无法产生相应的社会收益，资源的投入可能超出其价值，因此在司法实践中，完全消除错案或者错案率降为 0，并不是认罪认罚从宽制度的最优选择，受客观条件的约束，司法机关应尽可能将错案率控制在较低水平；错案率在高于 W' 的情况下②，进一步减少刑事错案率的措施会导致边际收益大于边际成本，因此司法机关应该继续采取措施减少刑事错案，因为会带来更多的社会收益，此时应当建立起相应的错案防范机制，努力减少刑事错案，从而实现认罪认罚从宽制度的效益最大化。

三、认罪认罚从宽制度的“成本—收益”分析

（一）认罪认罚从宽制度的成本和收益

通过对法律制度进行成本收益及经济效率分析，可以评价法律实施的效果，并对特定法律制度的社会价值作出评判[11]。认罪认罚从宽制度在运行的过程中，一方面带来了社会正收益，另一方面也会产生一定的社会成本。通过采用“卡尔多－希克斯”标准[12]，可以准确地比

① 即错案率趋近于 0 的理想状态。

② 即存在一定错案率的情况。

较认罪认罚从宽制度的收益与成本，从而探求该制度的最优效益。

1.认罪认罚从宽制度的总成本及边际成本

认罪认罚从宽制度的运行总成本包括两个方面：被追诉人刑罚的减少和无罪者受追究可能性的增加。其一，对于公诉机关而言，认罪认罚从宽降低了对犯罪分子的惩罚力度，可能会导致冤假错案的发生，因为适用认罪认罚从宽制度的案件，司法机关通常会依法对被追诉人从宽处理，刑罚的减轻使得刑事诉讼法惩罚犯罪的功能下降，刑罚的威慑作用减弱。其二，对于被追诉人而言，一方面，认罪认罚可能会导致被追诉人无罪辩护或量刑辩护权的丧失，进而使其失去被判无罪的可能；另一方面，认罪认罚从宽制度在某种程度上削弱了对人权的保障，尽管尊重和保障人权是刑事诉讼法的核心价值，但为了提高诉讼效率，司法机关常常会采用简化诉讼程序和缩短诉讼期限等措施，而这些措施往往会伴随着当事人诉讼权益的削弱，如有不慎，就易导致无罪者受到追究的可能性增加。

在此框架下，该制度的总运行成本可如此计算：假设在认罪认罚从宽制度实行之前，被追诉人可能面临的刑罚为常数 Y，而在制度实施之后，减刑比例为 L（其中 $0<L<1$）；另外，假设认罪认罚从宽制度可能导致无辜者受到追究的风险为 K，即无辜者受到追究的概率为 K，那么无罪者未受到追究的概率即正确判决的概率为（$1-K$）。因此，认罪认罚从宽制度的实施对被追诉人刑罚的减少量为（$1-K$）LY。当适用认罪认罚从宽制度后，判处的刑罚为（$1-L$）Y，那么认罪认罚从宽制度对人权保护效用的减少量可表示为 $K(1-L)Y$。如此，认罪认罚从宽制度的总成本 C 可以综合考虑被追诉人刑罚的减少和无罪者受追究可能性的增加，用式(1)表示：

$$C=(1-K)LY+K(1-L)Y。 \tag{1}$$

为得出认罪认罚从宽制度的边际成本，C 对 K 求偏导：

$$\frac{\partial c}{\partial k}=Y-2LY。 \tag{2}$$

因此，认罪认罚从宽制度的边际成本为 $Y-2LY$。

2.认罪认罚从宽制度的总收益及边际收益

认罪认罚从宽制度的运行总收益涵盖三个方面：司法机关的人力成本、时间成本的降低以及当事人长时间诉讼所带来负担的减少。首先，认罪认罚从宽制度的实施可以减少司法机关的人力成本，当被追诉人选择主动认罪认罚时，会减少审理过程中需要投入的人力资源。其次，认罪认罚从宽制度的实施还可以减少案件的诉讼时间，通过鼓励被追诉人在早期阶段认罪认罚，可以避免较长时间的审判过程，从而加快案件的结案速度。最后，认罪认罚从宽制度的实施可以减轻当事人的诉累，被追诉人可以避免长时间的诉讼拖延和可能增加的刑罚风险，追诉人能够获得及时的经济赔偿以及诉讼时间的缩减。

同时，这三方面的收益是相互关联的，存在着内在的正相关关系。例如，减少工作量会导致诉讼时间的减少，诉讼时间的减少又会减轻当事人的诉累。为了合理表达这种相关关系，可以采用以下的函数表达式：假设认罪认罚从宽制度的实施减少的人力成本为 T_1，减少的诉讼时间为 T_2，减轻的当事人诉讼负担为 T_3。假设存在两个正比例常数 a 和 b（$a,b>0$），使得 $T_1=aT_2=bT_3$ 成立，那么 $T_1+T_2+T_3=nT_2$（其中 n 为常数且 $n>1$）。另外，假设固定收益为常数 R'，认罪认罚从宽制度的总收益 R 等于固定收益 R' 减去无罪者受到追究概率下损失的收益 nT_2K，具体表示为：

$$R=R'-nT_2K。 \tag{3}$$

为得出认罪认罚从宽制度的边际收益，R 对 K 求偏导：

$$\frac{\partial r}{\partial k}=-nT_2。 \tag{4}$$

因此，认罪认罚从宽制度的边际收益为 $-nT_2$。

（二）认罪认罚从宽制度的“成本一收益”模型分析

为了寻求认罪认罚从宽制度的最大效益，

需要比较该制度的边际成本和边际收益，用经济学的公式表示为：

$$-nT_2=Y-2LY(\text{其中 } n>1)。\tag{5}$$

因为 Y 和 n 均为固定常数，因此 Y/n 也为固定常数，设为 $g(g>0)$，因此上述公式可以简化为：

$$T_2=2Lg-g。\tag{6}$$

进一步可简化为：

$$L=\frac{T_2}{2g}+\frac{1}{2}(0<L<1)。\tag{7}$$

以减少的诉讼时间 T_2 为横坐标，以减刑比例 L 为纵坐标，构建认罪认罚从宽制度的“成本一收益”模型，具体如图 3 所示。

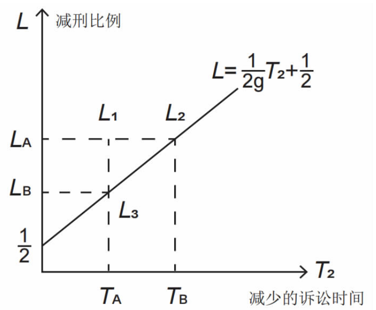

图 3　认罪认罚从宽制度的“成本一收益”模型

根据边际理论，当边际收益等于边际成本时，即处于“均衡”状态时，认罪认罚从宽制度的效益最大化。如图 3 所示，当减刑比例相同时，L_1 减少的诉讼时间为 T_A，L_2 减少的诉讼时间为 T_B，当且仅当数值位于直线上时，该制度的效益最大化，因此如果减刑比例同为 L_A 时，减少的诉讼时间最佳点应为 L_2，其对应数值为 T_B（见图 3）。同理可得，当减少的诉讼时间相同时，L_3 减刑比例为 L_B，L_1 减刑比例为 L_A，当且仅当数值位于直线上时，该制度的效益最大化，因此如果减少的诉讼时间同为 T_A 时，减刑比例最佳点应为 L_3，其对应的数值为 L_B（见图 3）。对比 L_3 和 L_2 可以发现，减少的诉讼时间 T_2 和减刑比例 L 呈现出正相关关系。

由于 $Y/n=g$，因此 $L=\frac{T_2}{2g}+\frac{1}{2}$ 可进一步转化为 $L=\frac{nT_2}{2Y}+\frac{1}{2}$，由此可分析 T_2、L 以及 Y 三者之间的关系。根据认罪认罚从宽制度的“成本一收益”模型，可得出以下结论：当判处的刑罚（Y）相对固定时，诉讼程序的简化程度（T_2）越高①，减刑比例（L）也应相应越高。一方面，根据经济学的原理，在应当判处的刑罚相对固定的情况下，诉讼程序的简化程度与减刑比例之间呈现出正相关关系，即在认罪认罚从宽制度的实行过程中，针对同一个或同一类型的案件，适用速裁程序或简易程序审理与适用普通程序审理相比，被追诉人应当享有更大幅度的量刑减让。另一方面，对于原定刑较重的案件，为了确保公平正义的有效实现，应在一定程度上限制诉讼程序的简化程度，从而控制减刑比例。综上所述，通过分析认罪认罚从宽制度的“成本一收益”模型可知，诉讼程序简化得越多，减刑比例就越大。因此对于原本应当受到较重刑罚的案件，在简化诉讼程序时应遵守一定的原则并受到相应的限制，以确保认罪认罚从宽制度的有效性和公正性。

四、认罪认罚从宽制度的优化构想

（一）强化对冤假错案的防范和治理

依据上文对认罪认罚从宽制度的“公正一效率”模型得出的两个结论可知，为了平衡司法公正和诉讼效率之间的关系，实现司法资源投入和刑事错案率之间的“均衡”，一方面需要将错案率控制在较低水平，另一方面需要完善冤假错案治理的保障机制。

1.将错案率控制在较低水平

在对待刑事错案的问题上，“宁可错放，不可错判”[13]的观点得到普遍支持，既体现了司法机关防范冤假错案的工作态度，也反映了司法运作系统的客观规律。从法经济学视角审视认罪认罚从宽制度，该观点也符合经济学模型分析的结果。如果司法机关选择“错放”，那么罪犯可能会继续危害社会，影响社会稳定和谐；

① 由于诉讼程序的简化即意味着诉讼时间的减少，因此减少的诉讼时间 T_2 可以表述为诉讼程序的简化程度。

如果司法机关选择“错判”，在放纵犯罪损害社会公共利益的同时，也会侵害无罪者的权利，即无罪者可能会被错误地追究刑事责任或者受到不必要的刑事处罚，从而对社会公共利益和个人权益造成损害结果。总体而言，“错判”所造成的损失往往比“错放”所造成的损失更大[14]。另外，“错判”相对于“错放”产生的更大的危害在于损害了社会公众对司法公正的信任度，降低了司法的公信力，该损失往往是难以弥补的。司法活动中刑事错案无法百分百规避，“不枉不纵”只能是法律制度的一种理想预期，会受到司法运行成本、科技发展水平、人类认知能力等条件的客观限制，刑事司法中完全杜绝错案是一种不现实的选择，“或枉或纵”也许才是客观存在的现实状态，所以我们不得不在“枉”和“纵”之间进行取舍[15]。

经济学模型也表明，在有限司法资源和人类认知局限等约束下，错案无法绝对消除，存在客观的残余风险。在给定司法资源总量约束下，追求绝对零错案可能导致案件积压、审理周期过长等问题，甚至可能因过度谨慎而抑制正当的司法能动性。但必须明确，有错必纠、追求每一个案件的公正是司法系统不可妥协的应然目标，任何错案都是对个体权利和司法公正的严重损害，必须彻底预防和纠正。经济学模型的作用在于：揭示在给定的资源和技术约束下，完全消除错案在现实中面临巨大挑战，过度集中于最后一层风险的消除可能导致整体司法效能的下降，因此司法机关应尽可能将错案率控制在较低水平，在维护司法公信力的同时实现制度效益的最大化。

具体而言，在认罪认罚从宽制度实施过程中，需要把无罪者认罪的比例控制在制度的边界范围内①，以确保刑事诉讼程序秉承的价值不丧失。那么，如何合理地确定错案率的最低边界？根据上文分析，当边际成本等于边际收益时错案率为 W'，即最优错案率为 W'，认罪认罚从宽制度适用之前的错案率假设为 W_1，那么适用该制度后错案率增加量的上限应为 $W'-W_1$，因此认罪认罚从宽制度错案率增加量须控制在 $W'-W_1$ 的限度内，如此才能实现司法资源的最优配置。最终，确定错案率的最低边界需要权衡各种现实因素，包括司法公正、诉讼效率和社会成本等。在综合考量这些因素的情况下，尽可能地降低错案率，从而确保获得最大的制度经济效益。

2.完善冤假错案治理的保障机制

在适用认罪认罚从宽制度审理刑事案件时，法官认定案件事实仍需要满足“犯罪事实清楚，证据确实、充分”的要求，并且达到“排除合理怀疑”的标准，因此仅将错案率控制在较低水平对于制度优化和公正审判而言是不足的，还需要完善冤假错案治理的相应保障机制。

其一，明确被追诉人的上诉权是完善冤假错案治理保障机制的重要举措。在认罪认罚从宽制度实施过程中，被追诉人是否应当具备上诉权存在学理上的争议。有学者主张，应限制被追诉人的上诉权，且适用速裁程序应当一审终审[16]；还有学者指出上诉权是基本诉讼权利，不应受到限制，否则将难以实现司法公正[17]；还有学者认为只有例外情况下才能赋予被追诉人上诉权，对上诉权应当严格限制[18]。从法经济学视角来看，肯定被追诉人的上诉权具备正当性基础。认罪认罚案件审理过程中通常会适用简易程序或者速裁程序，但在诉讼程序简化的同时也会变相增加“隐性成本”，即公诉机关和审判机关将需要更加审慎地查明案件真相，从而变相提高了查明真相的难度，增加了司法机关人员的工作量。另外，诉讼程序的简化在一定程度上也会提高冤假错案发生的概率。因此，完善认罪认罚从宽制度应适当控制“隐性成本”，一方面需要明确程序简化的上限，另一方面需要完善相应的治理错案的保障机

① 认罪认罚从宽制度的边界涉及程序简化限度的问题，后文会详细阐述。

制。被追诉人的上诉权可以被视为降低“隐性成本”和风险的保障措施,是控制诉讼成本的一种手段,是纠正冤假错案的风险分担机制。同时,在保证自愿性和真实性前提下的认罪认罚是被追诉人对自己实体权利和程序权利的有效处分,现代法律制度应当予以尊重。并且,在刑事诉讼中控辩双方的交涉能力并不平等,控方的诉讼地位和诉讼能力相对强势,肯定上诉权也是被追诉人认罪认罚自愿性的重要保障措施,有利于维护被追诉人的正当权益并减小冤假错案的发生概率。

其二,完善冤假错案治理的保障机制也需优化相关配套措施。首先,在明确肯定被追诉人上诉权的同时,需完善相应的配套机制,以防止恶意上诉和反悔等情况的发生,如严格设定被追诉人上诉的启动前提、明确上诉审查的适用案件范围、规定上诉审查的具体情形以及上诉审查的程序和标准等。其次,加强对被追诉人自愿性的审查和对办案人员的监督。在审查被追诉人是否认罪认罚的过程中,需要建立严格的审查程序,确保被追诉人的认罪行为真实、自愿且有效。审查程序应包括多个环节,如向被追诉人解释认罪认罚从宽制度规定的权利和义务,确保其能准确理解;对被追诉人进行多次询问和反复确认,以明确其真实意愿;提供适时的法律援助,确保被追诉人充分了解自己的权益和选择。另外,还应建立健全监督机制,对办案人员的行为进行监督和评估。监督机制可分为内部监督和外部监督两部分:内部监督包括对办案人员进行定期评估、对案件进行审核和督查等;外部监督可以由独立的机构或组织承担,以监督办案过程的合法性和公正性。最后,充分发挥值班律师的作用[19]。保障和强化值班律师的阅卷权,使其能够自由查阅与案件直接相关的文件、证据和材料等;值班律师应与检方开展积极合作,参与量刑协商的过程;为了保证值班律师高效履职,在考虑实际需求的基础上应设立相应的奖惩机制,这也是实现认罪认罚从宽制度司法正义和诉讼效率两大价值的应有之义。

(二)推进诉讼程序简化和量刑规范化

依据上文对认罪认罚从宽制度“成本-收益”模型得出的两个结论可知,为了实现制度效益的最大化,一方面需要协调诉讼程序简化与减刑比例的关系,另一方面需要明确诉讼程序简化的上限。

1. 协调诉讼程序简化与减刑比例的关系

从经济学的视角看,在认罪认罚从宽制度实施过程中,在判处的刑罚相对固定的情况下,适用速裁程序审理案件应当比适用简易程序审理案件获得更大幅度的量刑减让。而从法学的视角来看,现阶段我国并未明确规定可以根据适用诉讼程序的不同区分减刑比例。司法机关在对案件进行裁定时,通常会综合考虑多种因素,包括被追诉人的犯罪情节、主观悔罪态度、认罪认罚的时间、与司法机关的合作程度以及对追诉人的赔偿程度等。基于以上因素的考虑有助于建立多层次的量刑规则,以确保判决更为全面和公正[20]。

纵使经济学研究视角和法学研究视角得出的结论不一致,也并不意味着该矛盾不可调和,在认罪认罚从宽制度的实施过程中,速裁程序可视为一个独立的减刑情节,即适用速裁程序审理认罪认罚案件可以获得更多比例的量刑减让,主要基于以下理由。

其一,适用速裁程序获得的量刑减让,其减少的刑罚并非绝对的数值,而是相应的比例,更具有可操作性。例如,适用速裁程序审理的认罪认罚案件,可以在原定刑罚(Y)的基础上,明确规定减让 20%,即在不考虑其他量刑情节的基础上,适用速裁程序审理的认罪认罚案件最后判处的刑罚为 $Y\times80\%$;而适用简易程序进行审理的案件,则可以规定减让 10%,同样在不考虑其他量刑情节的基础上,适用简易程序审理的认罪认罚案件最后判处的刑罚为 $Y\times90\%$。

其二,速裁程序可以获得更大比例的量刑减让,其重要的原因在于速裁案件本身法定刑的裁量基准较低。根据《中华人民共和国刑事诉讼法》第 222 条①规定的速裁程序的适用条件可知,适用速裁程序审理的案件大多为轻刑案件,且案件事实清楚,证据确实、充分,案件争议不大,因此发生冤假错案的可能性较低。出于提高被追诉人认罪认罚率、提高诉讼效率的要求,速裁程序可以获得更大比例的量刑减让。

其三,速裁程序作为一个独立的减刑情节,是对被追诉人放弃部分诉讼权利的回报和奖励。诉讼程序的简化一方面提高了审判的效率,另一方面也意味着被追诉人放弃了部分的程序性权利,通过程序性权利换取实体性利益,被追诉人应当获得更大比例的量刑减让。司法机关通过给予选择刑事速裁程序的被追诉人一定幅度的“量刑优惠”形成“量刑奖励”,从而鼓励更多被追诉人自愿选择认罪认罚,提高认罪认罚从宽制度的适用率。在司法实践的过程中,存在部分地方检察机关在庭审之前对同意适用速裁程序审理案件的认罪认罚的被追诉人出具一份承诺书,承诺在基准刑上减小一定的量刑幅度[21]。因此,适用速裁程序审理的认罪认罚案件可以获得更大比例的量刑减让的做法在司法实践中具有一定的可行性。

2.明确诉讼程序简化的上限

在经济学中,“均衡”是指一种状态或情境,其中的各个经济变量达到一种相互协调的状态,没有内在的趋势导致其发生变化。“均衡”状态在经济学中是一种重要的分析方法,利于理解经济系统中各种变量之间的相互关系,并预测经济行为和结果[22]。在认罪认罚从宽制度中也需要找到诉讼程序的简化程度和减刑比例之间的“均衡”,为了兼顾司法公正和诉讼效率,不宜过度简化诉讼程序,也不宜过度进行量刑减让,需要明确诉讼程序简化程度的上限。

从法经济学视角出发可探索出认罪认罚从宽制度中诉讼程序简化的上限,根据前文,边际成本等于边际收益时最优错案率为 W',认罪认罚从宽制度适用之前的错案率为 W_1,因此适用该制度后错案率增加量的上限应为 $W'-W_1$。另外,由于程序的简化具有限度,假设程序简化的上限为 T_0,如此可以得出认罪认罚从宽制度的制度边界,即为认罪认罚从宽制度的总收益 $(R'-nT_2K)$ 等于减少的刑罚量 $(1-K)LY$ 加上适用认罪认罚从宽制度之后错案率增加量的上限 $(W'-W_1)$ 乘以认罪认罚之后判处的刑罚 $(1-L)Y$[23],即 $R'-nT_2K=(1-K)LY+(W'-W_1)\times(1-L)Y(T_2\leqslant T_0)$。

认罪认罚从宽制度的制度边界强调追求实质性公正而非仅是形式性公正,司法机关在实施该制度时应当注重确保公正性和合理性,而非过于追求程序上的简化和诉讼效率的提高。虽然认罪认罚从宽制度的实施需要提高效率、进而减少司法成本,但该做法应当受到一定的限制,以确保被追诉人的合法权益不受侵害。诉讼程序的简化过程需要在一定限度(T_0)内进行,超出该限度可能会导致程序不当简化,最终影响到案件的公正审理。从法学的视角来看,程序的简化也并非无限制,即使采用简易程序或者速裁程序进行审理,也同样需要经过开庭审理、检察机关出庭等程序。并且,当事人有权答辩、提交证据并进行辩论。同时,司法机关认定案件事实时不能降低证明标准,以确保审判结果的真实性和公正性,并保护诉讼中各方的权益。总之,认罪认罚从宽制度在司法实践中的适用边界需要进行控制,例如确保庭审程序的实质化、限制简化过程、在资源合理分配的同时控制错案率等,如此才能保证认罪认罚从宽制度的公正性和有效性,同时维护司法的权威和被追诉人的合法权益。

① 《中华人民共和国刑事诉讼法》第 222 条:“基层人民法院管辖的可能判处三年有期徒刑以下刑罚的案件,案件事实清楚,证据确实、充分,被追诉人认罪认罚并同意适用速裁程序的,可以适用速裁程序,由审判员一人独任审判。人民检察院在提起公诉的时候,可以建议人民法院适用速裁程序。”

五、结语

认罪认罚从宽制度作为一项兼具程序与实体内容的制度，在设立之初便展示了巨大的制度优越性，不仅有利于打击犯罪和尊重、保障人权，还有助于实现司法资源的优化配置并促进案件繁简分流。本文从法经济学视角出发，基于边际理论分析了该制度的各种理论和实践问题，以期能够推进认罪认罚从宽制度的优化。但法经济学理论在具体运用过程中可能存在一定的局限性，一方面，本文虽对认罪认罚从宽制度存在的成本和收益进行了总结，但并不能穷尽所有可能；另一方面，法经济学理论大多是以理性经济人的假设作为前提，但在复杂多变的司法实践中，很多主体难以做到绝对的理性。尽管如此，法经济学仍为研究法律制度提供了方法论指导，并拓展了研究思路和视角，帮助法学理论和实务界人士找到公正和效率的平衡点以及成本和收益的均衡点，并最终促进认罪认罚从宽制度的良性运转。

参考文献：

[1] 张军．最高人民检察院工作报告[N]．人民日报，2023-03-18(4).

[2] 王敏远．认罪认罚从宽制度疑难问题研究[J]．中国法学，2017(1)：17－34.

[3] 姚莉．认罪认罚程序中值班律师的角色与功能[J]．法商研究，2017，34(6)：42－49.

[4] 曲振涛．论法经济学的发展、逻辑基础及其基本理论[J]．经济研究，2005(9)：113－121。

[5] 罗尔斯．正义论[M]．北京：中国社会科学出版社，1988：25.

[6] 李文健．刑事诉讼效率论[M]．北京：中国政法大学出版社，1999：54.

[7] 加里·S·贝克尔．人类行为的经济分析[M]．上海：上海三联书店，1995：63.

[8] 梁益海．经济增长研究之研究：边际效益递增理论之胜出[J]．平原大学学报，2006(1)：11－14.

[9] 马歇尔·A．经济学原理[M]．朱志泰，译．北京：商务印书馆，1964：215－230.

[10] 保罗·A·萨缪尔森，威廉·D·诺德豪斯．经济学[M]．胡代光，余斌，张军扩，等译．北京：首都经济贸易大学出版社，1996：14.

[11] 罗宾·保罗·麦乐怡．法与经济学[M]．孙潮，译．杭州：浙江人民出版社，1999：34.

[12] 贾继鹏，刘卫斌．政府在社会保障责任中的卡尔多－希克斯有效性标准分析[J]．四川理工学院学报(社会科学版)，2009，24(2)：51－54.

[13] 沈德咏．我们应当如何防范冤假错案[N]．人民法院报，2013-05-06(2).

[14] 熊秋红．冤案防范与权利保障[J]．法学论坛，2010，25(4)：33－38.

[15] 林喜芬．论刑事司法错误的发生机理与现实危害：基于诉讼运行的分析[J]．南京大学法律评论，2011(1)：258－268.

[16] 汪建成．以效率为价值导向的刑事速裁程序论纲[J]．政法论坛，2016，34(1)：119－124.

[17] 陈瑞华．认罪认罚从宽制度的若干争议问题[J]．中国法学，2017(1)：35－52.

[18] 颜世征，张楚昊．认罪认罚案件被告人上诉应对机制[J]．人民检察，2017(15)：38－40.

[19] 汪海燕．三重悖离：认罪认罚从宽程序中值班律师制度的困境[J]．法学杂志，2019，40(12)：12－23.

[20] 陈卫东．认罪认罚从宽制度研究[J]．中国法学，2016(2)：48－64.

[21] 陈瑞华．"认罪认罚从宽"改革的理论反思：基于刑事速裁程序运行经验的考察[J]．当代法学，2016，30(4)：3－13.

[22] 张乃根．法经济学－经济学家视野里的法律现象[M]．北京：中国政法大学出版社，2003：45.

[23] 姜保忠，来宇．认罪认罚从宽制度的法律经济学分析[J]．河南财经政法大学学报，2021，36(4)：88－99.

(责任编校：冯兆娜)