

逮捕审查判断中犯罪嫌疑人 社会危险性的量化评估

周 翔*

内容提要：通过对2020年以来13万余份起诉书和不起诉书的机器学习，研究发现：我国司法人员在判断羁押必要性时，会评估犯罪嫌疑人的社会危险性，但相应评估侧重于罪行危险性因素，却忽略了人身危险性因素；引入量化评估方法，主要有利于改善仅依靠基本案情信息难以准确判断是否需要羁押的“复杂”案件的羁押必要性判断，量化方法有可能显著降低此类案件的羁押率。当前在判断羁押必要性时，办案人员主要依靠罪行危险性因素的主观综合判断，规范改革路径着力于规则细化和要件重构，但这无力化解社会危险性判断信息不足的问题。引入大数据建模方法，有助于系统统合零散的社会危险性判断信息。此外，可通过问卷、量表、数字化设备等方式扩充模型的训练数据，通过深化对社会危险性发生机制的理解，区分案由和社会危险性类型，构建多个子模型，以进一步提升量化工具的准确性。

关键词：逮捕 社会危险性 量化评估 羁押必要性审查

引言

我国学界和实务界在羁押必要性审查的改革方向上已初步形成共识，即应当加强对犯罪嫌疑人社会危险性要件的审查，以降低羁押率。纵览国内已有做法，加强对社会危险性要件的审查主要有规范和实证两条路径。前者是不断细化适用羁押措施和非羁押措施的情形，〔1〕并加

* 浙江大学光华法学院特聘研究员。

本文系国家社科基金项目“智慧司法背景下犯罪嫌疑人社会风险的大数据评估方法研究”（21CFX068）、浙江省哲学社会科学规划之江青年专项课题“数字法治背景下生成式大模型技术对裁判说理的辅助价值”（24ZJQN064Y）的阶段性成果。感谢浙江大学光华法学院硕士研究生陈瑞为数据分析提供的协助。

〔1〕 比如，2012年刑事诉讼法将社会危险性条件细化为五种情形；2015年最高人民检察院、公安部《关于逮捕社会危险性条件若干问题的规定（试行）》要求检察机关办理审查逮捕案件，应当严格审查是否具备社会危险性条件；2018年刑事诉讼法将“涉嫌犯罪的性质、情节，认罪认罚等情况”纳入社会危险性的评估要素；2023年《人民检察院、公安机关羁押必要性审查、评估工作规定》再次细化了审查评估的内容和方式。

强对社会危险性要件的阐释和要件关系重构。^{〔2〕}后者则采取定性和定量的方法,对犯罪嫌疑人的社会危险性作出相对实证的评估,以期提高审查的客观性。^{〔3〕}从实践情况看,各地也已经在加强社会危险性的量化评估工作。^{〔4〕}

虽然国内已提出多套量化评估方法和工具,并在北京、江苏无锡、江西萍乡、四川乐山等地的检察机关有所试验,^{〔5〕}但是,对这些方法和工具的质量以及降低羁押率的效果,目前尚未有文献作过实证分析。造成这类缺失的主要原因是,传统基于统计学的方法缺乏用于测量实践中社会危险性评估的现状、依据以及加强量化评估方法的效果的显性指标。以机器学习为代表的人工智能技术的兴起,为该领域的实证研究提供了新的发展契机。有鉴于此,本文利用算法和性能衡量等方法,对社会危险性要件的司法适用状况、评估依据和效果等,进行了相对系统的实证检验。本研究旨在分析,如果采取此种量化评估方法,其一,会对强制措施的类型判断产生何种实质性影响;其二,与规范改革路径相比,量化方法有何种优势;其三,如何实现两种路径的兼容。

一、研究假设和研究设计

(一) 研究假设

1. 研究假设一

值得验证的首要问题是,在我国当前的司法实践中,进行羁押判断时是否已经考虑了犯罪嫌疑人的社会危险性。虽然近年来开始强调社会危险性要件的审查,但该规范要求是否落实到了司法实践,在不同案件类型、地区落实的效果是否存在差异等,均缺乏实证检验。

由此提出假设一:在我国的司法实践中,羁押判断已经考虑了犯罪嫌疑人的社会危险性。如果实证检验发现,在案件基本事实、罪名等基础信息之外引入社会危险性评估信息,能够提升模型对现实中羁押判断过程的还原度,则说明假设一基本成立,即在我国的司法实践中,目前的羁押判断已经考虑了社会危险性的有关因素。

2. 研究假设二

既有文献批判了一种社会危险性评估理念和方法,即把罪行危险性和社会危险性相等同,

〔2〕 参见李训虎:《逮捕制度再改革的法释义学解读》,《法学研究》2018年第3期,第160页以下;杨依:《以社会危险性审查为核心的逮捕条件重构——基于经验事实的理论反思》,《比较法研究》2018年第3期,第132页以下;谢小剑:《审前未决羁押率下降:基本特点与成因解析》,《中国刑事法杂志》2021年第4期,第97页以下;秦宗文:《羁押必要性判断中的“一贯表现”证据研究》,《法学研究》2023年第6期,第189页以下。

〔3〕 参见高通:《逮捕社会危险性量化评估研究——以自动化决策与算法规制为视角》,《北方法学》2021年第6期,第132页以下;江溯:《论刑事司法中的危险评估》,《南京社会科学》2021年第5期,第91页以下;张吉喜:《逮捕社会危险性条件中犯罪嫌疑人逃跑风险评估研究》,《中国法学》2023年第4期,第286页以下。

〔4〕 例如,上海引入英国保释程序风险评估机制,对未成年人非羁押措施进行风险量表评估(See E. Li & M. Su, *From Punishment to Control: Assessing Juvenile Diversion in China*, 45 (2) *Law & Social Inquiry* 372-297 (2020));北京检察机关对流动人口轻微刑事犯罪不捕建立了风险评估项量表(See X. Yuan, *Risk, Risk Assessment, and Community Corrections in China*, 63 (14) *Int. J. Offender Ther. Comp. Criminology* 2466-2482 (2019));北京市石景山区人民检察院基于业务经验开发了社会危险性评估的直接主观赋值公式(参见杨秀莉、关振海:《逮捕条件中社会危险性评估模式之构建》,《中国刑事法杂志》2014年第1期,第63页)。

〔5〕 2024年4月,最高人民法院第一检察厅在湖北召开社会危险性量化评估试点工作座谈会,调研2023年9月最高人民检察院《关于深化推进社会危险性量化评估试点工作的通知》的推进情况,会上有部分试点单位作了报告。参见史兆珉、魏秋晨、曾林芸:《社会危险性量化评估,关键在于科学性精准性实质化》,《检察日报》2024年4月8日第1、2版。

忽略了以“一贯表现”为代表的人身危险性评价。^{〔6〕}若司法实践中果真存在重罪行危险性、轻人身危险性的现象，这是一种社会危险性评价的偏废，一定程度上说明社会危险性评估方法在科学性方面有待提升。本文借鉴该分类法，提出侧重于罪行危险性因素和侧重于人身危险性因素的两类评估。罪行危险性因素衡量的是案件本身的性质和严重程度，包括案件类型、情节和轻重等方面；人身危险性因素衡量的是犯罪嫌疑人的特征，包括社会属性、案后表现等方面。^{〔7〕}

由此提出假设二：在我国司法实践中，当前主要根据罪行危险性因素而不是人身危险性因素来判断社会危险性。如果实证检验发现，罪行危险性相关因素构造的模型能够更好地拟合社会危险性值，则说明假设二基本成立，即当前的司法实践更加偏向于采用罪行危险性因素进行社会危险性判断。

3. 研究假设三

本研究还希望探讨引入社会危险性量化评估对整体羁押率的影响，即引入量化评估方法能否真正合理地降低整体羁押率。基于前期的研究成果和认知，笔者认为，社会危险性的量化评估主要作用于“复杂”案件，主要影响这部分案件犯罪嫌疑人是否羁押的判断结论。^{〔8〕}引入社会危险性量化评估工具，能够控制乃至降低的应当是这部分案件的羁押率。2018年刑事诉讼法第81条第3款、2023年《人民检察院、公安机关羁押必要性审查、评估工作规定》第17条、第18条，明确规定了应当予以逮捕的情形和不予羁押的情形，这些案件属于本研究所称的“简单”案件。剩余案件，即在适用相关规定时存在较大裁量空间的案件，就是本研究所称的“复杂”案件。有研究主张，我国应当参考审判程序的繁简分流改革，推定大多数轻罪案件犯罪嫌疑人妨碍诉讼的社会危险性较低，而把社会危险性调查集中于重罪案件。^{〔9〕}本研究有理由相信，社会危险性量化评估方法和工具的引入，如果有助于改善整体的羁押状况，也主要作用于“复杂”案件。

由此提出假设三：社会危险性量化评估的引入，将降低部分难于判断的“复杂”案件的羁押率。为检验这一假设，本研究首先定义和识别出“复杂”案件，并以“复杂”案件的实际羁押情况作为现状的对照组，以社会危险性量化评估信息为输入的模型所预测的羁押情况是预测的实验组。比较实验组的羁押率和对照组的羁押率之后，如果发现加入社会危险性量化评估信息之后的案件适用羁押措施的比例下降，则说明假设三成立，即社会危险性的量化评估有助于改善整体的羁押状况。

（二）变量设置

本研究的目的是检验哪些因素会影响司法人员的社会危险性评估，社会危险性评估能否影响犯罪嫌疑人的羁押状态。故而，变量的设置包括确定社会危险性评估的量化方法、量化结果所依据的影响因素、司法实践的羁押判断结果等三个方面。

1. 社会危险性值的测量方法

国内既有的社会危险性量化评估方法设置了不同的社会危险性评估体系。高通在轻罪领域

〔6〕 参见前引〔2〕，秦宗文文，第187页。

〔7〕 本文就此提取了13万余个案件中有关的32个指标，将之分为“罪行危险性因素”（共20个）和“人身危险性因素”（共12个）两组，并就与既有文献分类法的对照关系作了说明。详见下文“变量设置”部分。

〔8〕 前期研究已经发现，在对案件采取繁简分流后，对简单案件采取简单的模型拟合，并未对模型预测羁押状况的总体性能造成减损（See X. Zhou, Q. Liu, Y. Wu et al., *LK-IB: A Hybrid Framework with Legal Knowledge Injection for Compulsory Measure Prediction*, *Artificial Intelligence and Law* (2023)）。该繁简分流的方法，亦可用于本文的实证研究。

〔9〕 参见前引〔2〕，秦宗文文，第187页。

提出了五分法,^[10]张吉喜采取的也是五分法,^[11]而杨秀莉等人采取的是四分法,^[12]王贞会采取的是三分法。^[13]造成这些差异的原因可能是:(1)采集的数据语料不同,从而关注视角有差异。比如,张吉喜的研究主要关注取保后重新犯罪、有逃跑风险的风险细分类型。(2)得到指标的方式不同。比如,杨秀莉等人主要采取梳理、归纳业务经验的方式,其他人则是从具有显著相关性的变量中挖掘有价值的指标。

关键是如何实现以上多种因素的综合评价,即把以上指标合并进社会危险性的单一指标中。在国内开展社会危险性评估工作的困难在于,我国司法办案部门并未过多开示社会危险性的评估过程,也未全面记录社会危险性的评估结果。本研究通过归纳“集体经验做法”来实现变量的可操作化。在社会科学的量化研究中有一种“去量纲化操作”的方法,近年来该方法被引入法学研究,比如用于测量妨害公务罪的严重程度。^[14]本研究采取类似的去量纲化处理方法,实现了社会危险性评估的可量化操作。^[15]

限于检察机关对逮捕和不予逮捕的理由说理不足和公开程度不够,本研究只能采用不批准逮捕理由说明书中的“无社会危险性”理由信息作为“集体经验”的来源,将其中不予逮捕的理由按照降序排列,以理由出现的频次作为社会危险性计算公式的变量和权重依据。为此,共收集并获得154份不批准逮捕理由说明书。^[16]对这些文书所使用的理由进行提取,共涉及30个相关变量,并对这些变量依据在总体文书中被提及的次数占比进行降序排列。本研究认为,某个不批准逮捕的因素被提及的次数越多,该因素在社会危险性判断中就越发重要,该因素对应的权重也应当越大。同时,当数据具有不同的量纲和范围时,直接进行比较可能会产生误导,因此将数据归一化到相同的比例范围,可以确保数据之间的比较更加准确和可靠。本研究将对应的频次进行百分比归一化处理后,作为模型的权重系数。^[17]计算公式中的X根据案情进行赋值,例如当犯罪嫌疑人自动投案时,“自动投案”项 $X=0$;犯罪嫌疑人未自动投案

[10] 包括犯罪嫌疑人自然情况的自变量,反映人身危险性的自变量,反映犯罪行为社会危害程度的自变量,反映妨碍刑事诉讼顺利进行可能性的自变量,反映人道主义考虑的自变量。参见高通:《轻罪案件中的逮捕社会危险性条件研究——以故意伤害罪为例》,《政法论坛》2021年第2期,第73页。

[11] 包括人口统计学指标、吸毒、涉嫌犯罪的类型和可能判处的刑罚、前科、程序性因素。参见张吉喜:《统计学方法在评估“逮捕必要性”中的运用》,《广东社会科学》2014年第6期,第221页。

[12] 包括犯罪情况、嫌疑人个人情况、罪后情况、被害人情况。参见前引〔4〕,杨秀莉等文,第63页。

[13] 包括人身危险性、社会危害性和诉讼可控性。参见王贞会:《审查逮捕社会危险性评估量化模型的原理与建构》,《政法论坛》2016年第2期,第70页。

[14] 去量纲化,就是借助各种数学工具实现从事物的经验值向抽象值的转换。其中,对不同事物进行权重分配,就是研究者将一定理论、评价标准注入研究对象的过程。参见白建军:《法秩序代偿现象及其治理——从妨害公务罪切入》,《中外法学》2020年第2期,第418页。

[15] 杨秀莉等人采取访谈、查阅卷宗等方式构造社会危险性的赋值方法。他们在北京市石景山区人民检察院统计了各类“具有社会危险性”的判断理由,但是未作为无量纲化操作的依据,而是就此归纳了几种常见的具有社会危险性的模式(参见前引〔4〕,杨秀莉等文,第63页)。他们计算各类社会危险性判断理由的频次和占比的方法,为本文所借鉴。同时,本文认为,较之主观归纳几种模式,不如把所有理由按照频次作为权重依据全部纳入一个计算公式,可能更为简单和科学。

[16] “不批准逮捕理由说明书”数据库文书收集自某省的检察机关官网,总计426份,经过数据去重,其中因社会危险性审查而不批准逮捕的数量为154份,本研究采用这部分文书作为构造公式的依据。值得注意的是,该样本数量并不充分,也不具有全国代表性,因此所拟合及反映的“司法实践如何作出社会危险性评价的集体经验”依然有待检验。但是,学术上提倡该方法具有指导实践的效果,检察机关可以利用内网或者全库的不批准逮捕理由说明书,进一步检验本研究提出的方法和结论是否得当。

[17] 百分比归一化是一种数据处理方法,通常用于将数据按照比例进行调整,使其处于0%到100%之间。这种方法可以帮助比较不同数据集之间的相对大小和趋势,而不受其原始数值范围的影响,便于比较和分析。

时, $X=1$ 。

表 1 社会危险性评估的集体经验和影响因素权重计算 [18]

影响因素	占比	百分比归一化后占比	影响因素	占比	百分比归一化后占比
如实供述	64%	12.7%	认罪认罚	55%	10.9%
无前科	51%	10.1%	积极赔偿	46%	9.1%
被害人谅解	45%	8.9%	情节较轻	36%	7.2%
悔罪表现	35%	7.0%	偶犯	32%	6.4%
主动投案	29%	5.8%	证据已固定/证据足以认定	21%	4.2%
退还赃物	18%	3.6%	在本地有固定住所	10%	2.0%
...

据此拟合出社会危险性值的测量公式:

$$Y = (0.127 \times X_1) + (0.109 \times X_2) + (0.101 \times X_3) + (0.091 \times X_4) + (0.089 \times X_5) + (0.072 \times X_6) + (0.070 \times X_7) + (0.064 \times X_8) + (0.058 \times X_9) + (0.042 \times X_{10}) + (0.036 \times X_{11}) + (0.020 \times X_{12}) \dots + (0.002 \times X_{30})$$

基于以上公式,本研究对 13 万余个案件的数据库进行了社会危险性的数值测量。

2. 社会危险性的影响因素

为发现司法实践中社会危险性评估的实然机制,必须对影响社会危险性的因素进行大规模样本测量。在测量前,首先要定义可能影响社会危险性的因素,并构造出影响因素的指标体系。研究假设二的主要目标是实证分析司法实践中社会危险性评估的有关机制,进而确定在当前的社会危险性评估中主要依据的信息类型。对于社会危险性评估的依据,既有文献给出了不同的分类法,包括但不限于以下几类:人身危险性、社会危害性、诉讼可控性的三分类法;〔19〕犯罪情况、嫌疑人个人情况、罪后情况、被害人情况的四分类法;〔20〕犯罪嫌疑人自然情况、反映人身危险性、反映犯罪行为社会危害程度、反映妨碍刑事诉讼顺利进行可能性、反映人道主义考虑的五分类法;〔21〕犯罪嫌疑人基本情况、违法犯罪史、犯罪行为、犯罪后表现的四分类法。〔22〕这些分类方法为本研究构造分类指标体系提供了指引。

如前所述,本研究采取二分类法,将判断社会危险性的罪行危险性和人身危险性两大类因素分为四小类,把犯罪类型、行为和情节归入罪行危险性因素,把犯罪嫌疑人人口信息、妨碍诉讼的可能性和诉讼可控性等归入人身危险性因素。本研究共提取了案件的以上因素共计 30 项。罪行危险性因素一共 11 项,其中“法益侵害”信息 9 项,包括情节轻微,从犯,金额较少,是否为家庭邻里纠纷,被害人过错,犯罪未遂,未造成损失,证据是否足以认定,是否可能被判处管制、拘役、缓刑;“法益修复”信息 2 项,包括被害人谅解、积极赔偿。人身危险性因素一共 19 项,其中“人口学”信息 12 项,包括有无前科、偶犯、本地固定住所、无业、企业家、吸毒史、一贯表现良好、限制刑事责任能力人、户籍所在地是否一致、严重疾病、家

〔18〕限于篇幅,仅列出了频次占比超过 10%、判断权重较高的 12 个影响因素,但在构造“集体经验”的公式时,依然采用了共计 30 个因素。

〔19〕参见前引〔13〕,王贞会文,第 70 页。

〔20〕参见前引〔4〕,杨秀莉等文,第 63 页。

〔21〕参见前引〔10〕,高通文,第 73 页。

〔22〕参见前引〔2〕,杨依文,第 130 页。

庭困难、性别；“案后表现”信息7项，包括如实供述、认罪认罚、悔罪表现较好、自动投案、立功、退缴赃物、抗拒抓捕。两大类四小类的分类法，对前述三分法、四分法和五分法作了一定的合并处理，但依然遵循了过去的习惯，比如社会危险性的典型三分法——犯罪性质、罪后表现和犯罪嫌疑人自身情况，本研究的罪行危险性因素等同于第一类，人身危险性因素等同于第二、三类，且基本对应人身危险性因素的两个子类型。^[23]

3. 社会危险性的用途指标

本研究特别关注的是，社会危险性量化评估的直接效果——决定是否对犯罪嫌疑人予以羁押。本研究主要的社会危险性用途指标，是针对犯罪嫌疑人采取“逮捕”或者“取保候审”的强制措施类型。

为了更全面反映司法实践审查强制措施的状况，本研究采用起诉书、不起诉书等审查起诉阶段文书，测量社会危险性评估对强制措施审查的影响。这类文书能够更客观地反映强制措施的类型分布，如果采用判决书，则可能遗漏部分不起诉案件，从而低估取保候审的适用比例。本研究收集了B区、Z区、G区、Y区等四个地区2020年1月1日至2021年12月31日间的全库不包含未成年犯罪嫌疑人的起诉书和不起诉书，经数据清洗后共计132645份。

(三) 分析方法

为了加强数据分析结论的丰富性和研究假设的检验，本研究还引入主题模型的算法，^[24]以分析不同社会危险性值之下的案件特征；引入LSTM深度学习的算法，实现了可直接采用文本作为模型的输入格式。为验证研究假设，并能够更好地从数据层面分析、揭示司法实践中的羁押判断过程和结果，本研究并行采用两种衡量模型还原能力的方法。^[25]

方法一：统计学的 R^2 评价方法。本研究的主要变量为“社会危险性值”，这是一个连续数值变量，本研究构建了多元线性回归模型并以 R^2 指标说明自变量的影响程度。 R^2 ，也称决定系数（coefficient of determination），是用来评估一个回归模型的拟合优度的指标。它的取值范围在0到1之间，越接近1表示模型的拟合效果越好。在统计学中， R^2 是一种常用的评估回归模型拟合优度的指标。^[26]

方法二：计算机科学通用的F1评价方法。F1值是精确率和召回率的调和平均数，其取值范围在0到1之间，数值越大，模型性能越好。^[27]通俗地说，该数值越高，意味着所构造的模型越是接近（拟合）训练数据背后的人类判断经验。机器学习的方法还有一些特别的优势，

[23] 值得注意的是，这里的影响社会危险性的因素提取和社会危险性值的测量是两个独立的操作。社会危险性值的测量依据是不予逮捕的理由，是一种集体经验的反映；提取影响社会危险性的因素，则是用于检验社会危险性的发生机制。

[24] 主题模型，是一种能够大致刻画不同社会危险性值的案件特征和差异的无监督机器学习方法，该方法的本质是利用词频的高低、词语的共线性聚合而成的不同主题，除了主题的个数、虚词的停用需要人为干预外，其余均为机器自动生成。因此，这种方法对于实证研究中的描述性分析是极有价值的，避免了人力的大量投入，且能够在一定程度上表征出案件的重要信息。关于主题模型的通俗介绍和应用示例，参见刘知远、崔安硕：《大数据智能》，电子工业出版社2016年版，第97页以下。

[25] 传统的法学定量研究主要采用统计学为主的分析和检验方法，当前计算机科学兴起的机器学习、深度学习等方法为法学研究开辟了新的路线。两种方法的比较，参见唐亘：《精通数据科学：从线性回归到深度学习》，人民邮电出版社2018年版，第64页以下。

[26] 参见孙敬水、马淑琴：《计量经济学》，清华大学出版社2022年版，第18页；王小毅、张继红：《统计学基础》，浙江工商大学出版社2016年版，第242页；赵松山：《对拟合优度 R^2 的影响因素分析与评价》，《东北财经大学学报》2003年第3期，第56页。

[27] 该数值既考虑了精确率，也兼顾了召回率，能够全面衡量模型预测的准确性和全面性。

比如可以用非结构化的语料作为输入；数据集分为训练集和测试集，用训练集拟合模型，在测试集中进行测试。以上两点，改变了过去以统计学为基础的定量实证研究。〔28〕

虽然本研究引入了机器学习中通用且成熟的评价方法，并且和统计学检验方法进行比较，但后文采用的分析方法，通俗地说，是在遵循“量化模型是一种对人类判断集体经验的反映”的假设的基础上，采用集体经验构造并测量社会危险性值，在 13 万余个案件样本中对我国的羁押审查判断是否已经开展社会危险性评估、评估所依据的因素指标、引入量化评估工具对羁押率的影响等三个方面进行评价。

二、社会危险性量化评估的依据和效果

（一）社会危险性值的测量结果

本研究分析了社会危险性值的测量结果，首先根据社会危险性值的高低进行分组，将 13 万余份样本按 0.3、0.7 的阈值分为高于 0.7 的高风险组、介于 0.3 与 0.7 之间的中风险组、低于 0.3 的低风险组，一共三组。接着通过主题模型、词云等方法，构造“社会危险性值区分度”的新指标。分析发现，不同等级的社会危险性值集中于不同的案件类型，不同案件类型表现出的社会危险性值分布集中度差异较大。

1. 依社会危险性值等级的分组情况

本研究对主题模型的分类结果进行合并和命名后，观察到不同社会危险性程度的组内情况差异。低风险组中，主要的案件类型为危险驾驶、故意伤害和盗窃。从地区间的差异看，G 区的案件集中于危险驾驶（占比 80%，高于总体的 47.5%），盗窃案件的社会危险性评价则相对严苛（占比 0%，没有被评价为低社会危险性的案件）；与之相反，Y 区对盗窃案件犯罪嫌疑人的社会危险性评价相对比较宽容（占比 11.7%，在各地区中最高）。

表 2 低风险组的案件类型

	主题一：危险驾驶	主题二：故意伤害	主题三：盗窃
总体	 <p>占比：47.5%</p>	 <p>占比：27.3%</p>	 <p>占比：8.1%</p>
B 区	占比：46.9%	占比：29.1%	占比：8.2%
Z 区	占比：53.4%	占比：14.5%	占比：9.5%
G 区	占比：81%	占比：19%	占比：0%
Y 区	占比：62.4%	占比：25.4%	占比：11.7%

再看中风险组的社会危险性值测量结果。前三位是危险驾驶、盗窃、赌博。与低风险组比较，危险驾驶案件虽然横跨低、中两组，但呈现出占比下降的趋势，在各地区间则有占比

〔28〕 关于以大数据、机器学习为代表的新兴技术对法学定量研究的方法论意义和具体变革路径，参见周期：《作为法学研究方法的大数据技术》，《法学家》2021 年第 6 期，第 60 页以下。

15.6%—43.4%的较大波动；盗窃跃升至第二位，占比呈现上升趋势，而且盗窃的财物价值更高；赌博进入了高频罪名的序列，主要分布于G区；B区占比第三位的仍然是故意伤害。

表3 中风险组的案件类型

	主题一：危险驾驶	主题二：盗窃	主题三：赌博
总体	 <p>占比：37.4%</p>	 <p>占比：12.5%</p>	 <p>占比：3.5%</p>
B区	占比：37.6%	占比：13.7%	(故意伤害) 占比：7.0%
Z区	占比：26.6%	占比：11.4%	占比：5.2%
G区	占比：15.6%	占比：10.7%	占比：9.4%
Y区	占比：43.4%	占比：10.1%	占比：6.7%

最后看高风险组的情况。高频案件类型变为盗窃、毒品犯罪（包括贩卖、运输毒品等）、诈骗。从案件类型看，毒品犯罪和诈骗是低、中风险组中未见到的。高频案件类型在各地区间的分布存在差异，比如盗窃案件在B区占比最高，毒品犯罪案件在Y区占比最高。

表4 高风险组的案件类型

	主题一：盗窃	主题二：毒品犯罪	主题三：诈骗
总体	 <p>占比：52.1%</p>	 <p>占比：10.3%</p>	 <p>占比：7.9%</p>
B区	占比：53.6%	占比：9.3%	占比：7.6%
Z区	占比：46.5%	占比：8.2%	占比：6.8%
G区	占比：42.3%	占比：9.8%	占比：7.4%
Y区	占比：47.1%	占比：11.3%	占比：10.5%

从组间的案件类型分布情况看，高频案件类型的覆盖率不一。低风险组的分布最集中，覆盖案件类型的八成；中风险组的分布相对较为分散，主要为危险驾驶，盗窃和赌博的覆盖率不足六成；高风险组的案件类型分布，集中度介于低风险组与中风险组之间。

2. 依罪名的分组情况

为进一步说明社会危险性测量结果的分布情况，可切换至“罪名”的视角分析社会危险性与犯罪行为之间的联系。本研究认为，只有案件基于社会危险性值归入高风险、低风险两组，才可能给司法人员提供充分的参考价值——归入高风险组，即可作出予以羁押的判断；归入低风险组，即可作出不予羁押的判断。为此，本研究设计了一个特别指标——社会危险性值区分度，用于衡量社会危险性量化评估方法在不同案件类型中对社会危险性的区分度。计算公式为：社会危险性值区分度 = (低风险组的案件数 + 高风险组的案件数) / 中风险组的案件数。本研究计算了高频罪名的社会危险性值区分度，结果发现，区分度较高的罪名有偷越国（边）

境罪、盗窃罪、贩卖毒品罪等，区分度不足的罪名有危险驾驶罪、妨害作证罪、开设赌场罪等。

表 5 不同罪名的社会危险性值区分度

罪名	社会危险性区分度
偷越国（边）境罪	0.68
盗窃罪	0.65
贩卖毒品罪	0.64
...	...
开设赌场罪	0.11
妨害作证罪	0.07
危险驾驶罪	0.06

从罪名间分布情况看，社会危险性值的分布区间存在较大差异。一些罪名的社会危险性值测量结果处于中间地带的较多，一些罪名则分布于社会危险性值的高、低两极。本研究认为，社会危险性的量化评估方法对羁押判断的贡献，重点在于后者，即社会危险性值区分度高的案件类型。社会危险性量化评估的结果若要给予司法人员足够清晰的指引，社会危险性值要么足够高、要么足够低，如此办案人才有充足理由作出予以羁押或者不予羁押的判断。从给出的罪名间社会危险性值区分度的排序看，盗窃罪、贩卖毒品罪、寻衅滋事罪等案件的犯罪嫌疑人，其社会危险性值要么处于高区间、要么处于低区间，能够给予司法人员更加明确的信号。相对而言，对于危险驾驶罪、交通肇事罪等罪名，社会危险性的量化评估则并非必要或者并不是有效的。^[29]反之，就提升社会危险性量化评估的科学性而言，重点应当放在社会危险性值区分度不够高的罪名。

（二）社会危险性要件的司法适用状况

本部分实验，利用机器学习的方法还原了当下的羁押判断方式，用于检验当前实践是否实际考虑了社会危险性要件。本研究默认在缺乏社会危险性要件审查的情况下，亦能进行强制措施审查；单纯依靠案件基本事实作出的强制措施类型判断，不一定全面反映司法现状；缺乏精确的量化评估也不代表司法实践没有考虑社会危险性要件。^[30]本部分的对照实验，可以说明单纯依靠案件基本事实、罪名等基础信息，无法全面还原司法实践进行羁押判断的现实情况，而社会危险性评估信息是一种有助于还原羁押判断现实情况的增益性信息。

1. 单纯依靠案件事实信息难以还原羁押判断的现实情况

本部分实验对单纯依靠案件事实信息能否完全拟合司法实践的羁押判断现状以及发生判断错误的原因展开检验。本研究分析了单纯依靠案件事实信息发生判断错误的案件规律，通过构建以“案件事实段”为输入、“是否羁押”为输出的模型，并比较模型预测的判断结果和实际的判断结果，从而筛选出模型“预测错误”的案件子集。

[29] 这里值得注意的是，不同的高中低社会危险性的划分阈值，会影响最终的分组和社会危险性值的区分度计算。但是，这并不影响该方法本身的实践价值，实践部门可以反复调试不同阈值，从而分析得到应当开展社会危险性量化评估的重点案件类型。

[30] 这里要承认，本研究构造的社会危险性量化公式不一定能够完全拟合现实中的社会危险性考量因素，本研究只是当前数据条件下的一种实验策略。

从表6可以看出,单纯依靠案件的基本事实预测羁押状态的F1值约为0.83,不能较好地预测实际的羁押判断。进一步分析预测错误案件的要素特征,可以发现如实供述、情节严重程度、认罪认罚、前科等指标偏高,预测错误案件在以上方面呈现出共性。

表6 单纯依靠案件事实信息错误预测羁押状态的案件特征

如实供述	99.8%
认罪认罚	90.1%
情节严重	79.1%
前科	29.8%
F1值	0.83

2. 社会危险性评估信息有助于还原实际的羁押判断

本研究检验司法实践进行羁押判断是否考虑了社会危险性有关信息的方式是,比较有无社会危险性评估信息的前后,模型对司法实践判断羁押状态的还原度是否存在差异。如果单纯依靠案件事实信息,模型未能较好地拟合司法实践进行羁押判断的现状,就说明实际的羁押判断还考量了案件事实以外的其他信息;加入社会危险性评估信息之后,如果模型的预测性能变得更优异,则说明社会危险性评估信息对于还原司法实践的羁押判断而言是一种增益性信息。

本研究的分析结论认为,加入社会危险性值作为模型的输入,有助于强化模型还原真实司法决策场景的能力。实验表明,与单独以案件事实信息作为输入相比,添加罪名信息能够提高模型预测强制措施类型的能力;继续添加社会危险性值之后,模型的性能进一步提升。单纯依靠案件事实信息的模型表现为F1值0.83,说明未能完全还原司法现状。实验组模型较之对照组,性能分别提升了5%和3%;分地区看,提升幅度从0%到6%不等。

在默认“数据模型是对人类决策过程的拟合”这一前提下,模型性能越好,说明数据模型的还原能力越强。数据上体现为,在加入社会危险性有关信息之后,模型对实践中羁押必要性审查方式的还原程度更高。

表7 不同输入项的模型对羁押判断的还原度

F1值	总体	B区	Z区	G区	Y区
对照1:案件事实	0.83	0.79	0.69	0.93	0.80
对照2:案件事实+罪名	0.85	0.80	0.75	0.93	0.84
实验组:案件事实+罪名+社会危险性值	0.88	0.82	0.75	0.93	0.85
性能提升(较之单纯以案件事实为模型输入)	5%	3%	6%	0%	5%

值得特别说明的是,本实验仅能说明社会危险性评估信息有助于还原司法实践的羁押判断现状,以上实证结果证明了研究假设一成立的可能性,即从“模型是对实然审查模式的反映,社会危险性因素的加入提升了模型的性能”这一点看,司法实践在进行羁押判断时确实考虑了社会危险性,只是不一定采用严格的量化方法,在不同罪名间采用的方法也有所不同。实践中对社会危险性是予以量化评估,还是停留于主观判断,评估所依据的具体因素如何,地区间存在的量化方式和量化程度上的差异等,均有待后续的定量检验。

(三) 社会危险性的司法判断依据

本部分实验把社会危险性值作为因变量,研究在羁押判断中司法实践如何确定犯罪嫌疑人

的社会危险性。本研究首先默认通过前述无量纲化处理得到的公式，是对社会危险性判断集体经验的恰当反映，在此基础上实验对 13 万余个案件构成的数据库进行社会危险性数值测量，并提取可能影响社会危险性判断的因素，最后对社会危险性判断的内在机制予以分析。本研究分别采取统计学的 R^2 值、机器学习的 F1 值两种衡量方法，对“究竟是人身危险性因素还是罪行危险性因素对社会危险性的判断影响更大”的研究假设二进行检验。

1. 统计学方法的检验

统计学方法的检验结果表明，司法实践在评估社会危险性时更加倚重罪行危险性因素。本研究采用所提取的 30 个因素作为自变量、以社会危险性值作为因变量，多元线性回归模型中 $R^2 = 1$ ，说明以上指标能够较好地拟合社会危险性的判断模型，侧面验证了采取“频次法”赋值社会危险性和提取影响因素具有合理性。进一步分析发现，罪行危险性因素组的 R^2 大于人身危险性因素组的 R^2 ，从而初步检验了研究假设二：当前的司法实践依然十分倚重案件情节的严重程度、案后是否积极退赃退赔、是否取得被害人谅解等罪行危险性因素，对社会危险性作出评价。

表 8 罪行危险性和人身危险性对社会危险性评估的影响权重 R^2 比较

人身危险性因素组的性能评价	人口信息	案后表现	人口信息+案后表现
R^2	0.210	0.398	0.576
罪行危险性因素组的性能评价	法益侵害	法益修复	法益侵害+法益修复
R^2	0.135	0.730	0.863
人身危险性因素组+罪行危险性因素组	$R^2 = 1$		

2. 机器学习方法的检验

深度学习的建模和衡量方法亦表明，罪行危险性因素的作用更加显著。实验采用 LSTM 的深度学习算法，分别采用当事人信息段、基本事实信息段作为模型的输入，测试哪一类信息能更好地预测社会危险性值。^[31] 实验结果表明，“基本事实段”信息作为输入的模型，预测社会危险性值的效果更好。考虑到采用文本信息作为模型的输入，较之采用结构化要素作为输入，所遗漏的信息更少，该结果可加强“当前的社会危险性判断更加依赖罪行危险性因素”的判断的可信度。

表 9 罪行危险性和人身危险性对社会危险性评估的影响权重 F1 比较

输入的信息	当事人信息段	基本事实信息段
深度学习方法 F1	0.355	0.481

（四）引入量化评估工具对羁押率的影响

本部分实验把社会危险性值作为自变量，主要检验引入量化评估工具对整体羁押率的影响。根据既有的认识，社会危险性的量化评估主要作用于“复杂”案件。本部分实验希望检验，与单纯依赖案件事实信息进行强制措施类型判断的审查相比，使用考虑了社会危险性评估

[31] LSTM 是一种基于长短期记忆递归神经网络的深度学习算法，实验中把起诉书和不起诉书中的“当事人段”和“基本事实段”分词后作为输入，分别构建了 2 个 LSTM 模型，预测 Labels 为社会危险性值（连续因变量），探究“当事人段”和“基本事实段”哪个文本信息能够更好地预测社会危险性值。

信息的模型预测强制措施类型，案件的整体羁押率能否得到控制。本部分采用两种方式定义“复杂”案件：其一，在采用“基本案情信息”预测强制措施类型时，模型预测的结果与实际结果不一致的案件（以下称“预测错误案件”）；其二，在采用“基本案情信息”预测强制措施类型时，模型预测的结果不能明确地归于某一种强制措施类型的案件（以下称“预测困难案件”）。

1. 预测错误案件

本研究在控制使用同一最优模型（GBDT 梯度提升决策树）的条件下，首先用“案件事实 + 罪名”作为输入信息预测羁押判断结果，把其中 53801 个预测错误案件作为本部分的研究对象。对于这部分案件，有理由认为，单纯用案件事实和罪名来判断是否羁押是不够的，笔者将对比通过“案件事实 + 罪名 + 社会危险性值”预测的羁押率和实际的羁押率。

对比结果表明，把社会危险性值作为输入信息后，模型预测的整体羁押率为 45%，而同样测试集中的实际羁押率约为 81%。从分析结论看，在“复杂”案件中引入社会危险性量化评估，也许有助于得出“非羁押”的结论，^[32]从而初步验证研究假设三。

表 10 预测错误案件中引入社会危险性值的羁押率变化情况

组别	羁押率 (N=53801)
预测羁押率	45%
实际羁押率	81%

2. 预测困难案件

本研究以深度学习 LSTM 算法，将“基本事实段”文本信息作为模型的输入进行羁押判断。每个案件的预测结果是从 0 到 1 的连续变量，算法一般默认 0.5 为二分标签的阈值，本场景中“0.5 以下”判定为“非羁押”，“0.5 以上”判定为“羁押”。从科学性的角度看，判断值临界于 0.5 附近的案件其实属于“预测困难”的情形。因此，本研究把预测结果介于 0.45—0.55 的案件定义为“复杂”案件，共计筛选出 4152 件。笔者将这部分案件作为测试对象，对比通过“案件事实 + 罪名 + 社会危险性值”预测的羁押率和实际的羁押率。

对比结果再次表明，研究假设三成立，即具备社会危险性评估信息之后，能够有助于得出“非羁押”的结论。这体现为整体的羁押率得到了控制，机器预测的羁押率为 40%，比实际羁押率低 14 个百分点。

表 11 预测困难案件中引入社会危险性值的羁押率变化情况

组别	羁押率 (N=4152)
预测羁押率	40%
实际羁押率	54%

本部分提供的实证数据，为引入社会危险性量化评估提供了一定指引。一方面验证了既有理论和实践的一些认识，另一方面证实了“逮捕必要性要件是抑制逮捕适用的关键”，^[33]

[32] 值得说明的是，表格中的“羁押率”统一使用性能最好的 GBDT 梯度提升决策树模型计算。该结论是在地区案件集合意义上而言的，在个案层面，不是每个案件都遵循该结论。

[33] 参见前引 [2]，杨依文，第 130 页。

“将社会危险性条件视为适用逮捕的核心要件”〔34〕等改革建议具有一定的合理性。

三、社会危险性审查的实践模式和改革路径

前文实证研究聚焦于司法实践中羁押判断的现状和依据，并对加强量化评估以控制案件羁押率的前景作了初步评估。基于以上分析结论，本部分从理论角度总结现有的社会危险性审查模式，反思当前做法的形成背景及局限，进而评析当前的主流改革路径。

（一）实践中社会危险性的审查模式

关于强制措施类型在司法实践中的审查判断方式，有不少学术归纳，包括对“构罪即捕”“以捕代侦”等现象的概念提炼和局限性批判，〔35〕并且试图在刑事程序的更大视域内考察逮捕的审查方式及功能。〔36〕这些基于经验主义的逮捕审查模式归纳，有助于理解逮捕审查判断的实然运转方式。本研究在此基础上，基于实证分析的发现以及针对检察机关一线办案人员的访谈，〔37〕拟进一步总结归纳社会危险性要件的审查方式。笔者发现，当前针对社会危险性要件的审查，司法机关遵循“繁简分流”、后置于其他要件的审查模式，大致依照如下步骤。

第一步：先审查“证据要件”和“刑罚要件”，再审查“社会危险性要件”，社会危险性的评估处于程序的最后一环。司法实践中的真正做法是在充分阅卷、讯问犯罪嫌疑人等必要程序之后，先判断其是否构成犯罪、是否可能判处有期徒刑以上刑罚，只有在满足以上两个要件时，才会启动社会危险性要件的审查程序。

第二步：根据案件的“繁简”进行社会危险性判断的分流。这里的“繁简”不完全等同于案件情节的轻重，而是与影响社会危险性评估依据的因素多寡有关。“简单”案件，是指那些依靠单一因素即可作出决定的情形，包括法条中明确认定为具有社会危险性情形。〔38〕实践中，办案人员一般径行将这些案件归入有较高或较低社会危险性的类别，这一点在本研究的实证部分得到过检验。其余案件则属于“复杂”案件，这些案件不能通过单一因素直接作出判断，而是要根据多种因素综合判断才能得出结论。

第三步：针对“复杂”案件，办案人员将根据案件的不同类型，开展差异化的评估工作。不同案件类型意味着所侵犯的法益、是否具有被害人、犯罪嫌疑人是否可能实施新的犯罪的概率都有所不同。办案中高效区分案件类型的方法可能是基于罪名作初步分流。本研究针对所收集的不批准逮捕理由说明书中“无社会危险性”的理由进行统计，发现不同罪名的审查重点有很大差异，比如盗窃案件和故意伤害案件的社会危险性评估，在模式上就存在明显区别。

第三步之一：再犯可能性的评估。对于某些再犯情形高发的罪名，审查者重点考虑的是，如果采取非羁押强制措施，犯罪嫌疑人再次犯罪的可能性大小。本研究统计了盗窃案件中“无社会危险性”的理由，发现较高频次的理由包括如实供述（占比66%）、无前科（占比

〔34〕 参见前引〔2〕，李训虎文，第114页。

〔35〕 参见韩东：《构罪即捕现象的解析及应对》，《河南社会科学》2015年第11期，第8页。

〔36〕 参见王彪：《刑事诉讼中的“逮捕中心主义”现象评析》，《中国刑事法杂志》2014年第2期，第72页。

〔37〕 本研究的访谈主要是验证性的，基于笔者曾有过的审查逮捕办案经验，先归纳出审查方式的步骤，再向原任职单位的同事进行验证，由他们对这份审查方式的归纳文本进行完善和补充。

〔38〕 主要包括“有证据证明有犯罪事实，可能判处10年有期徒刑以上刑罚的”，“有证据证明有犯罪事实，可能判处有期徒刑以上刑罚，曾经故意犯罪或者身份不明的”，具体的法律依据是2018年刑事诉讼法第81条第3款。

66%)、认罪认罚(占比54%)、退还赃物(占比49%)。以上情节基本能够说明犯罪嫌疑人的认罪悔罪态度以及重新犯罪的可能性。从司法办案经验看,盗窃的确也是再犯可能性较大的犯罪类型,有盗窃前科的盗窃案件犯罪嫌疑人不在少数。

第三步之二:被害人态度的评估。对于那些有被害人的案件,审查者还可能重视被害人及其家属的态度。本研究统计了故意伤害案件中“无社会危险性”的理由,发现较高频次的理由是取得被害人谅解(占比86%)、积极赔偿(占比72%)。这说明在故意伤害案件中,办案人员可能认为通过事后赔偿等方式可以实现双方和解,这会被看作犯罪造成的社会关系破坏情况已经有所修复。^[39]

需要特别说明的是,实践中社会危险性的审查过程不一定严格按以上步骤进行,办案人员完全有可能改变审查顺序,甚至“一气呵成”直接进行判定。对实践做法区分出步骤,能够聚焦社会危险性量化评估的着力点。本研究认为,量化评估主要在第三个步骤中对“复杂”案件发挥作用。当前,这部分案件的社会危险性评估以定性的、主观的判断为主。由于未全面采用以算法公式为内核的量化评估工具,当前的实践做法容易简化社会危险性要件的审查过程,从而给人留下结论主观性强、说理不充分的印象。

(二) 实践做法的形成背景及局限

其一,审查者把罪行危险性因素作为社会危险性判断的主要依据,与当前社会危险性判断依赖侦查案卷信息有关。从实际办案情况看,检察机关和审判机关依然严重依赖侦查机关所收集的证据,特别是大部分案件很少出现审查逮捕中检察机关要求侦查机关补充取证的情况。这一点在当下似乎很难改变,在“案卷中心主义”之下,检察机关在审查和决定审前羁押时,主要依赖侦查案卷,这就很难做到全面、客观、中立的审查和决定。^[40]总之,单纯围绕侦查机关制作的案卷开展社会危险性评估,本身即有偏颇性。

其二,审查者在侦查案卷的基础上,并无足够动力、意愿继续拓展罪行危险性因素以外与社会危险性有关的信息。根据有关办案期限的规定,案件初次羁押的审查判断期限较短,审查者要在一周左右的时间内完成阅卷、提讯和文书制作等工作,并且审查者还可能同时办理多个案件。局促的办案期限、“复杂”案件难于定性、事后问责的潜在风险等等,使得办案人员并无意愿、也无能力扩充有关社会危险性的信息。

其三,在社会危险性判断中采取主观心证方法,是一种理性选择。在处理那些社会危险性判断困难的案件时,办案人员只能结合案卷中已知的因素,得到主观上认为恰当的综合判断,似乎没有其他可以借助的工具。这一点从本研究的社会危险性值等级分组中,各地区案件词云画像不一致,特别是中等社会危险性值的案件特征中,能够看出社会危险性评估标准的多元化,决策权限分散于个案并且办案人员的主观裁量空间较大。

虽然必须承认,实践中以罪行危险性因素为主的主观心证式社会危险性判断模式,是现实条件局限下的权宜之计,但此种模式亦暴露出如下局限性:人身危险性因素相对匮乏,办案人员无意也无力拓展有关人身危险性的信息,由此导致依赖侦查案卷记载的罪行危险性因素进行判断。这些汇总起来,使得社会危险性错判的风险增加。

[39] 第三步可能并非仅有这两小步,从当前社会危险性因素的子类型看,还有“逃跑”“妨害作证”等其他类型,并且可能是就特定罪名特别审查某一种社会危险性类型。比如,对于(帮助)偷越国(边)境类案件,“逃跑”的风险或许比其他风险要高。其他罪名案件的犯罪嫌疑人,亦有各自的社会危险性特点。

[40] 侦查案卷主要记载与定罪量刑有关的情节,就社会危险性评估而言,这些信息难言完整和充分。

（三）围绕规范细化和要件重构的改革

提升评估能力的方法无外乎两个，要么从制度规范方面考虑如何重塑审查判断过程，要么引入更加可取的评估工具。很显然，主流改革思路是前一种。前文的分析已经说明，司法实践把社会危险性评估置于强制措施判断的最后一个环节，并且以案件基本事实和罪行危险性因素为主要判断依据。自2012年修改刑事诉讼法以来，^{〔41〕}理论界和实务界主要围绕规范细化提出改革方案，以期强化对社会危险性要件的审查。这主要表现为，立法和司法机关不断出台法律规定和规范性文件，细化社会危险性的判断要素和标准。具体来说，自2012年开始的一系列规范，主要采用列举具体情形的方式推动改革进程，包括2012年刑事诉讼法明确了社会危险性的要件和情形；2015年最高人民检察院、公安部《关于逮捕社会危险性条件若干问题的规定（试行）》细化了社会危险性的五种具体内涵；2018年刑事诉讼法进一步明确了犯罪的性质、情节、认罪认罚情况、对所居住社区的影响等要素，作为判断社会危险性的依据；2023年《人民检察院、公安机关羁押必要性审查、评估工作规定》进一步明确，对可能判处3年有期徒刑以下刑罚的犯罪嫌疑人等情形，应当开展羁押必要性审查，并对羁押必要性审查、评估的内容和方式作了细化。

在学术研究方面，表现为加强对逮捕三要件的阐述和重构，包括各要件的功能定位、要件之间的关系、适用顺序等，以突出社会危险性要件的重要性。就逮捕的实体要件，有学者主张，应当强调逮捕三要件之一的社会危险性的审查，阶层化地重构逮捕三要件，把证据要件作为基础要件、刑罚要件作为否定性要件，将社会危险性要件作为“核心要件”；^{〔42〕}还有学者认为，社会危险性要件是抑制羁押率的重点。^{〔43〕}

以上表明，改革的着力点主要在于：一是不断细化作为“大前提”的规范，二是突出社会危险性要件的重要性。改革的最主要成果是使办案人员有了更具可操作性、可理解性的规范依据。但是，这类指引只是解决了羁押判断中的一小部分案件，即本研究所谓的“简单”案件。甚至这种审查模式还可能把部分原本属于“复杂”案件的，归类为“简单”案件，^{〔44〕}从而并未化解那些需要结合多种因素综合评估社会危险性的判断难题。因而，此种路径的效果，笔者是十分怀疑的。

四、量化方法与规范路径的兼容

（一）规范改革路径的主要局限

理论上重新构造逮捕三要件的分工和关系，可能产生一种新的审查犯罪嫌疑人是否满足羁押条件的方法指引，特别是引导基层办案人员重视社会危险性要件对于控制羁押率的宝贵价

〔41〕 本研究将2012年作为加强社会危险性评估工作的起始年份，主要是因为2012年刑事诉讼法构造了逮捕审查新的要件体系，并细化了社会危险性判断的五种情形和径行逮捕的法定情形。之后，2015年最高人民检察院、公安部《关于逮捕社会危险性条件若干问题的规定（试行）》、2016年《人民检察院办理羁押必要性审查案件的规定（试行）》等规范性文件，实则遵循了同一制度逻辑。

〔42〕 参见前引〔2〕，李训虎文，第114页。

〔43〕 参见前引〔2〕，杨依文，第130页。

〔44〕 笔者曾经在实务工作中发现，法条规定的“径行逮捕”情形在实践中被无一例外地执行。其实，如此处理存在粗疏化的问题。有学者提出了类似的质疑，比如“10年有期徒刑以上”原本是一种预估，其不能代替社会危险性要件的既有判断程序。参见前引〔2〕，杨依文，第130页。

值。然而，单纯依靠规范细化和要件重构的改革路径，无法化解社会危险性判断不足的问题。这一路径存在一些明显的短板：

其一，规范列举的情形需要结合犯罪嫌疑人的具体情况作进一步判断。在规范列举的具有社会危险性情形中，“逃跑”“再犯罪”等都不是实际已经发生的危害，而只是面向未来的、未知的风险。办案中审查的关键在于通过案件及其犯罪嫌疑人的已知信息，从中提炼有价值的因素来预测风险的大小。要作出社会危险性大小的结论，当前主要基于办案人员的主观经验进行笼统判断，未来或许可以采用更为精准、科学的方法。

其二，规范改革路径无助于增加办案人员对社会危险性事实层面的认知。不论是细化规则还是列举具体情形，并不能给予办案人员有关案件和犯罪嫌疑人更多事实层面的信息，即规范改革路径集中于优化法律适用的“大前提”，而与事实层面的“小前提”无涉。已有研究注意到了这一局限性，比如从证据法的角度认为，社会危险性评价不足的根源是相关事实依据不足。^[45]

其三，沿袭既有改革思路，终归无法通过成文规范事先事无巨细地规定清楚每一类“复杂”案件情形的社会危险性判断方法。本研究认为，对于那些需要结合多种因素综合判断社会危险性的案件类型，是无法通过规范细化的路径实现判断准确性方面的改观的。“复杂”案件的特点决定了，即使成文规范具体列举了社会危险性的表现类型和判断方法，也依然需要审查者综合案件的细节对社会危险性作出判断，即不存在直接表征“复杂”案件社会危险性程度的单一指标。

综合以上分析，似乎既有的规范改革路径并不能彻底解决当下的实践问题。细化规则和重构要件的改革思路，只是有助于解决“简单”案件中的社会危险性评估问题。由于缺乏犯罪嫌疑人人身危险性方面的信息增益，并采取更为科学的方法归纳社会危险性的各种类型及其发生机制，当前的规范改革路径似乎无力解决“复杂”案件中的社会危险性判断难题。

（二）量化方法与规范路径的互补

实际上，量化方法与规范路径不是对立的关系，而是完全有可能实现有效互补。社会危险性的量化评估方法，其实质是将社会危险性因素首先予以具体化和类型化；其次研究各类型社会危险性的发生机制，借助罪行危险性和人身危险性的指标体系刻画这一机制；再次是运用大数据训练并赋予各指标以不同权重的方式，实现对社会危险性的科学赋值。

量化方法与规范改革路径各自的侧重点和互补性体现为：

其一，规范路径是在法律适用的“大前提”上做文章，而量化方法是在事实层面的“小前提”上下功夫。充实“小前提”，能够克服当前社会危险性审查个人主观色彩浓重、判断缺乏统一性的问题。^[46]

其二，量化评估方法能够拓展案件信息的边界。如果说规范路径的目标是更为准确地理解和把握规范的含义，量化方法则是通过大数据采集、人工智能算法，在人力资源、办案期限、机关分工等既定不变的情况下，尝试冲破传统办案模式的信息瓶颈，得到更多有关犯罪嫌疑人社会危险性的信息。同时，量化评估方法还将提升评估过程的可解释性，让社会危险性评估得

[45] 参见张琳：《逮捕羁押审查中社会危险性的证明问题研究》，《中国刑事法杂志》2023年第5期，第125页。

[46] 这一方面的不足，参见王子毅：《降低审前羁押率的影响因素分析与对策研究》，《中国刑事法杂志》2021年第4期，第101页。

到过程性开示，促使逮捕审查判断工作朝向科学化、可追溯、能事后说明的方向发展。

其三，量化方法也需要规范路径的指引。一方面，规范层面明文规定了具有社会危险性和不具有社会危险性的情形，这可以设计成量化评估的直接判断规则或者更高权重的指标。量化评估的改进重点，应当放在审查那些有多个参考指标且指标间关系复杂的案件。当案件的社会危险性值介于中间状态（比如在设置的阈值 0.5 附近）时，说明在司法实践中，办案人员欠缺预估社会危险性的能力，依靠规范指引也依然无法作出清晰判断。这是未来引入量化评估方法时应当着力处理的案件类型。

评估犯罪嫌疑人的社会危险性，主要依靠审查者综合各类社会危险性信息作出准确预测的能力。规范细化和要件重构，可能只是明确是与否的大致界限，而无助于改变社会危险性评估信息匮乏、未能精细化运用的现状。相较而言，对于需要更多案件和犯罪嫌疑人信息、更加需要科学确定各类信息权重的“复杂”案件，量化评估方法更能发挥其专长。

（三）引入量化工具的可行路径

在规范改革路径基本保持不变的情况下，笔者认为，加强引入量化评估工具的可行性在于：一方面，提升办案机关采集社会危险性信息的能力；另一方面，加强对社会危险性发生机制的认知度。

1. 提升采集社会危险性信息的能力

其一，扩大可用于构造量化评估公式的备选数据项类型。基于本文的实证研究发现，我国当前判断社会危险性值主要依赖罪行危险性因素，应通过量表等方法加强对人身危险性信息的收集，从而弥补判断数据项的结构性短缺。办案机关应当尽快完成量表设计，在案件立案后、逮捕审查决定作出前的全过程中，均可适当穿插针对犯罪嫌疑人、被害人和其他诉讼参与人的问卷调查，加快积累一批可以用于训练社会危险性评估模型的大数据集。

其二，利用数字化手段扩大对人身危险性信息的收集范围。应当借助这一轮司法数字化浪潮，尽快提升智能设备的引入程度，采集时间序列更长、收集过程更加客观的犯罪嫌疑人、服刑期满人员的相关社会活动数据。比如，非羁押码、电子手环、电子脚铐、车载物联网等设备，都可以记录被释放的犯罪嫌疑人的行踪。^[47] 诸如此类工作，能够有效收集非羁押状态下犯罪嫌疑人的社会危险性数据，从而丰富社会危险性评估模型的训练数据标签。彼时，本研究所采取的利用不批准逮捕理由说明书测量社会危险性的方法，就会被更科学的方法所取代。

2. 加强对社会危险性发生机制的认知度

其一，组织专家学者开展深入且系统的社会危险性发生机制的理论研究。如前所述，量化方法的数据建模，其主要功能是统合各类事实信息材料，通过数据建模的方法更准确地量化犯罪嫌疑人的社会危险性。量化公式是否科学、精准，关键取决于对社会危险性发生机制的把握是否到位。因此，改进量化评估方法，需要组织有经验的司法人员、专家学者，深入挖掘社会危险性与案件各要素之间的机制性联系。

[47] 据报道，“非羁押码”的操作流程为“下载安装手机 App—注册认证—完成注册流程—开始每日打卡”，具有外出提醒、违规预警、定时打卡、不定时抽检等多项功能。办案人员可同步安装“非羁押码”，“码”上对非羁押人员进行实时管控、动态监督和自动记录。“从‘非羁押码’App 在杭州部分地区试点到全面推行以来，截至 2023 年 10 月 27 日，全市运用‘非羁押码’有效监管非羁押人员 1607 人。”参见范跃红、方芳、方利利：《浙江杭州：应用“非羁押码”降低审前羁押率提升监管效能》，https://www.spp.gov.cn/zdgz/202011/t20201109_484305.shtml，2024 年 2 月 20 日最后访问。

其二,可以采取分类型、分地区的模型建构思路。实务中,可以区分社会危险性的具体类型、不同罪名、不同地区实现建模。^[48]这一思路有本文实证研究的相关数据予以支持,比如盗窃案件和故意伤害案件的社会危险性判断就存在很大差异。应当承认,由于公开样本量的限制,本研究对相关机制的挖掘还很不够。最适合继续深入探索建模方法的,是掌握了全库案件信息、具备更强资源调动和整合能力的司法实务部门。

以上两方面工作应当同步进行。社会危险性发生机制的研究为进一步的数据采集指明了方向,扩大的数据样本和数据项则有助于检验相关机制是否具备实证支持。两类成果将在具体构建社会危险性量化公式时共同发挥作用,有关机制的研究结论将主要决定公式的数据项体系和类型,海量的社会危险性数据库将通过机器学习得出各个数据项的权重。

Abstract: Through machine learning on more than 130,000 legal documents since 2020, this study reveals that when making detention decisions, Chinese judicial personnels assess the social dangerousness of criminal suspects, but such assessments focus on elements of “criminal dangerousness” while overlooking those of “personal dangerousness”. The introduction of the quantitative assessment method is mainly beneficial for the handling of “complex” cases in which it is difficult to accurately determine whether to detain the suspect relying solely on basic case informations. It is argued that the quantitative method has the potential to significantly reduce the detention rate in such cases. In the current practice of detention decisions, case handlers mainly rely on subjective assessment based on the elements of “criminal dangerousness”. Traditional reform path focuses on rule refinement and element reconstruction, but falls short of addressing the insufficient assessment of social dangerousness. Incorporating big data modeling methods can help to systematically integrate scattered detention decision information. In addition, we can also expand the training data of the model by using methods such as questionnaires, scales, and digital devices, and further improve the accuracy of quantitative tools by deepening the understanding of the mechanisms underlying social dangerousness and distinguishing between causes and types of danger to construct multiple sub-models.

Key Words: arrest, social dangerousness, quantitative assessment, necessity review of detention

[48] 近来这一方面的工作正在加强,比如专门针对非羁押人员的逃跑风险进行研究,有助于围绕该类风险开发子模型。并且,笔者十分同意“运用综合评估方法,将不同类型的审前风险杂糅在一起进行评估,会降低评估结果的准确性”这一判断。区分不同社会危险性的表现,分别研究相关机制并设计模型进行预测,是将该领域的量化工具开发引向深入的重要趋势。相关研究,参见前引[3],张吉喜文,第281页。