

“人工智能+司法”与公共安全治理

——来自刑事案件的实证证据

刘嘉琪 周黎安 贺强*

摘要: 本文考察人工智能(AI)对刑事犯罪治理的影响。研究表明,公安机关应用 AI 显著降低了犯罪数量。异质性分析发现,该效应在轻度案件、可见案件中更为明显,且对较低学历、单独实施犯罪、初次实施犯罪,以及在异地实施犯罪的犯罪者效果更强。机制分析表明,AI 应用有效缩短了破案时间,并经由信息传播形成威慑,促使潜在犯罪者更新对被捕概率的预期,进而抑制其犯罪动机。本文为 AI 提升公共安全治理效能提供了实证证据。

关键词: 人工智能;公共安全治理;犯罪威慑

DOI: 10.13821/j.cnki.ceq.2026.03.13

一、引言

近年来,人工智能(Artificial Intelligence,以下简称 AI)进入技术加速突破与场景深度融合的新阶段,为赋能经济发展、提升治理效能与改善民生福利带来重大机遇。2025 年 8 月,国务院印发《关于深入实施“人工智能+”行动的意见》,明确提出加强人工智能在公共安全预警、社会治安管理等领域的应用,提升运用人工智能维护和塑造国家安全的能力。^① AI 技术在治理领域的深度应用,将成为推进国家治理能力现代化的重要技术路径。与此同时,随着我国社会主要矛盾的转化,公众对安全与稳定的需求持续升级,推动治理体系向精细化、智能化方向转型。在此背景下,AI 技术赋能公共安全治理,构成了对社会需求的技术回应。公共安全治理事关社会稳定与人民福祉,AI 技术应用对该领域所产生的实际影响及其作用机制,亟待深入考察。

* 刘嘉琪,清华大学五道口金融学院;周黎安、贺强,北京大学光华管理学院。通信作者及地址:贺强,北京市海淀区颐和园路 5 号北京大学光华管理学院 2 号楼,100871;电话:18034998237;E-mail:heqiangnf@163.com。本文得到国家自然科学基金重大项目(72192844)的资助。作者感谢席天扬、张丹丹以及“人工智能的经济影响:前沿研究与政策启示”专题研讨会参会者提出的有益建议,感谢主编和匿名审稿专家的宝贵建议。文责自负。

^① 详见《关于深入实施“人工智能+”行动的意见》(国发〔2025〕11号),中国政府网,https://www.gov.cn/zhengce/content/202508/content_7037861.htm,访问时间:2025 年 10 月 28 日。

犯罪给社会运行与经济发展带来巨大的直接成本与间接负担(Anderson, 1999; Chalfin, 2015; Ranasinghe and Restuccia, 2018; Cohen, 2020),犯罪治理效能是衡量政府公共安全治理水平的关键维度。我国公安机关兼具刑事司法与行政执法双重职能,是犯罪防治体系的核心力量,但长期面临基层警力不足、技术手段滞后及侦查效率低下等问题(熊一新,2014;解源源和史全增,2014)。为疏解上述困境,公安系统近年来持续推进信息化与智能化转型,先后实施“金盾工程”“雪亮工程”等基础设施建设项目,并向“智慧公安”战略升级。这些改革举措推动AI技术在公安侦查领域逐步应用,为评估AI赋能公共安全治理的实际效能与作用机制提供了准自然实验场景。

本文基于2010—2022年全国公开的检察文书、刑事裁判文书数据及2015—2022年的政府采购数据,利用各城市公安机关AI采购在时间上的差异性构造实验组和控制组,采用渐进双重差分法考查AI技术应用对刑事司法治理的因果效应。本文研究发现,公安机关引入AI技术显著降低了所在地区的刑事犯罪数量,该结论在控制传统监控设备采购的干扰之后依然稳健。异质性分析显示,上述效应主要由轻度与可见案件驱动,且对较低学历、单独实施犯罪、初次实施犯罪及在异地(非户籍地)实施犯罪的犯罪者具有更强的效果。机制分析表明,AI有效缩短了案件侦破时间,并经由媒体等信息传播形成犯罪威慑,促使潜在犯罪者更新对被捕概率的预期,进而抑制其犯罪动机、减少犯罪数量。

本文在三个方面拓展了现有研究。首先,本文丰富了数字技术赋能公共治理的相关认知。近年来,数字技术迅猛发展并在公共治理领域广泛应用,催生了大量评估技术应用对公共治理影响的文献(Jin and Lee, 2014; Muralidharan et al., 2016; Doleac, 2017; Guriev et al., 2021; Dzansi et al., 2022; Greenstone et al., 2022; Banerjee et al., 2023; Assunção et al., 2023; Cohen, 2024),特别是近期兴起一批文献开始关注人工智能技术在司法治理中的作用(Nissan, 2017; Kleinberg et al., 2018; Ludwig and Mullainathan, 2024)。与本文背景相似,Mastrobuoni(2020)发现警察部门引入信息技术优化巡逻策略显著提升了警务效率;Jabri(2021)利用美国单一城市辖区的数据验证了算法驱动预测性警务对严重财产犯罪与暴力犯罪的抑制作用;Ma et al.(2026)则发现城市监控摄像头的大规模部署显著降低了犯罪率。本文以AI技术在公安机关的应用为研究对象,为理解AI技术在犯罪治理领域的实际效能提供了系统全面的实证证据。

其次,本文推进了关于犯罪威慑效应的研究。既有文献主要关注客观制裁与个体风险感知的关联(Kessler and Levitt, 1999; Pogarsky et al., 2004)、惩罚经历对风险感知的重塑效应(Matsueda et al., 2006; Maccoun and Reuter,

2009; Anwar and Loughran, 2011),以及潜在犯罪者对即时环境线索的反应机制及其对犯罪决策的影响(Sjoquist, 1973; Bachman et al., 1992; Gardner and Steinberg, 2005)等。区别于既有研究主要依赖问卷调查、深度访谈或描述性统计的分析范式,本文采用因果识别策略,在验证威慑效应存在性的基础上系统剖析其作用机制,并进一步考察受教育程度、犯罪经历、犯罪人数及跨地域作案等因素对该效应的异质性影响。

最后,本文整合检察文书与刑事裁判文书,构建了覆盖全国、全案由、案件级的刑事诉讼数据库,并基于该数据构造了测度公安机关刑事犯罪侦破效率的直接指标。犯罪经济学关心各种可能影响犯罪的因素,其中研究最丰富、因结论不一致而最具争议的是警务系统对犯罪的影响(Corman and Mocan, 2000; Di Tella and Schargrotsky, 2004; Levitt, 2004; Blanes i Vidal and Kirchmaier, 2018; Mastrobuoni, 2019)。鉴于刑事案件侦查阶段的信息披露极为有限,既有研究通常只能依赖加总统计数据,或局限于特定辖区与案由展开分析(Tauchen et al., 1994; Levitt, 1997; Mastrobuoni, 2020)。本文所构建的数据库有效克服了过往研究在数据精细度不够、样本代表性不足等方面的局限,为后续在犯罪经济学领域的深入研究提供了坚实的数据基础。

二、制度背景与研究假设

(一) 制度背景

1. 公安侦查技术的发展

我国公安刑事侦查技术的演进与信息化改革政策紧密结合,大致可划分为三个发展阶段。第一阶段(1949—1978年)为基础建设期:该阶段以手印鉴定、法医学检验、文件检验等传统技术手段为主,技术基础较为薄弱,相关机构尚处于逐步完善之中。第二阶段(1978—2015年)为信息化起步与普及期:随着计算机技术被引入犯罪信息管理,公安系统依托“金盾工程”的持续推进,建成了覆盖全国的公安三级信息网络,搭建了DNA与指纹数据库,完成了视频监控网络(基于传统摄像头)的全域覆盖,实现了全警种信息共享,显著提升了刑事侦查的技术能力。第三阶段(2015年至今)为智能信息化深化期:公安部部署实施“雪亮工程”,强力推进“智慧公安”战略布局,通过公安大数据平台深度整合既有数据资源,结合AI算法,在人脸识别与视频侦查、电子取证、犯罪预测三大核心领域取得突破性进展;同时依托无人机集群、区块链存证等新兴技术,推动警务模式由人力密集型向科技主导型转型,实现了从传统经验驱动向数据智能驱动的根本性转变。

传统侦查模式以人工现场勘查为主导,信息采集周期较长,且易受侦查人

员主观认知偏差、生理疲劳及注意力波动等因素干扰,增加破案线索遗漏与破案进程延误的风险。视频监控技术的引入一定程度上缓解了上述问题,实现了对嫌疑人行动轨迹与作案过程的客观记录。但受制于存储成本与数据管理标准,视频数据难以长期归档,且线索挖掘仍依赖人工逐帧审阅,检索效率低下并存在漏检风险,关键信息仍可能淹没于海量数据之中。AI技术的引入推动了侦查方式的系统性革新:在视频监控领域,基于深度学习的计算机视觉技术可智能解析大规模视频,自动识别嫌疑人的特征与行为模式,挖掘潜在关联线索;结合DNA检测技术,可显著提升身份比对精确度和犯罪记录检索效率;在网络空间治理方面,可完成网络溯源、物理定位与虚拟身份关联,对网络账号进行实时监控检测与行为数据分析,实现犯罪风险预警。当前,AI技术在刑事侦查领域的应用已展现出显著成效,大量刑事案件破获案例表明其在提升侦查效率和破案能力方面具有重要作用。^①

2. 刑事案件办理流程

本文的研究对象为触犯《中华人民共和国刑法》的刑事案件。完整的刑事案件办理流程,一般始于公安机关接到报案或发现线索,通常包括侦查、起诉与审判三个阶段。在侦查阶段,公安机关综合运用各类侦查技术手段调查案件事实、收集证据;当证据初步证明犯罪嫌疑人实施了犯罪行为时,可依法对其采取逮捕等强制措施(下文统称逮捕)^②。公安机关实施逮捕后,需在法定期限内将嫌疑人移交检察机关,期间继续开展侦查活动,包括补充收集证据、讯问犯罪嫌疑人等。侦查终结后,公安机关认为犯罪事实清楚、证据确实充分的,将案件移交检察机关审查起诉。在起诉阶段,检察院经审查认为符合起诉条件的,依法向法院提起公诉。在审判阶段,由人民法院根据查明的事实对被告人作出是否定罪量刑的判决。在上述流程中,侦查环节是公安机关查明案情、收集证据的核心阶段,直接影响刑事案件的侦破效率。高效的刑事司法体系不仅能对已发生的犯罪行为实施惩罚,更能通过提升破案概率形成威慑效应,抑制潜在犯罪者的犯罪动机,从而从源头上减少犯罪。刑事案件办理流程及威慑效应的逻辑框架参见图1。

^① “重庆南岸:‘山城夜警’AI赋能智慧守护城市‘烟火气’”,重庆市公安局南岸区分局官网, https://www.cqna.gov.cn/bm/qgafj/sy/dt_59966/202411/t20241108_13781810.html, 访问时间:2025年10月28日;“昆山警方利用‘AI警察’快速破获网络诈骗案”,中国新闻网, <https://baijiahao.baidu.com/s?id=1814145297868518224&wfr=spider&for=pc>, 访问时间:2025年10月28日。

^② 根据《刑事诉讼法》规定,为确保诉讼程序顺利进行,公安机关可根据侦查进展程度,依法采取逮捕等不同强度的强制措施限制犯罪嫌疑人的人身自由,防止嫌疑人逃跑、毁灭证据、继续犯罪、妨碍调查,或者进行其他对案件正常审理产生不利影响的行为。

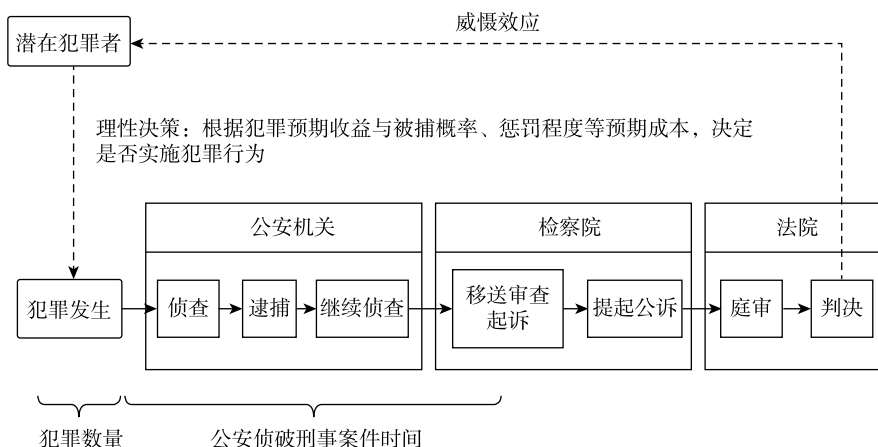


图1 刑事案件办理流程与犯罪威慑效应

(二) 理论分析与研究假设

犯罪经济学理论的奠基者 Becker(1968)首次将经济学分析框架引入犯罪行为研究,提出犯罪是一种理性选择行为:潜在犯罪者通过权衡预期收益与预期成本决定是否实施犯罪,其中,预期收益涵盖物质利益与心理满足等,预期成本包括被捕概率与惩罚强度等。Becker 进一步指出,惩罚的确定性与严厉性是威慑犯罪的两大核心变量,而公共治理部门的最优执法策略旨在最小化犯罪带来的社会总成本,包括直接损失、执法成本、惩罚成本等。Stigler(1970)在 Becker 框架基础上引入执法资源约束,即警力与经费总是有限的,并提出在预算约束下的执法资源最优配置理论,强调通过提升执法效率而非简单增加投入来预防犯罪,为后续关于技术赋能警务的研究奠定了理论基础。

基于前述理论,本文从潜在犯罪者与公安机关的策略互动出发,刻画双方的动态决策逻辑。潜在犯罪者权衡犯罪的预期收益与预期成本(含被捕概率),以决定是否实施犯罪;公安机关则在预算约束下通过最优资源配置选择目标破案率,以最小化社会总成本。双方动态博弈最终达成犯罪率与破案率的均衡。AI 技术在公安侦查中的引入,首先通过提升公安机关的客观破案效率形成威慑信号,该信号经由媒体报道或社会宣传等信息渠道扩散,提升潜在犯罪者对破案效率的主观预期,进而抬升其犯罪预期成本,抑制犯罪动机,最终表现为犯罪数量的减少^①。本文预期,相较于未采购 AI 的地区,采购了 AI 的地区刑事犯罪数量显著降低,据此提出本文的核心假设:

假设 1 AI 应用显著减少刑事犯罪数量。

根据理论分析, AI 技术通过影响潜在犯罪者的主观决策实现威慑效应,故

^① 在常住人口年度变化不大且回归中控制了城市年度固定效应的情况下,可将其理解为犯罪率。

其作用强度应随技术部署特征、案件特征和犯罪者特征的不同存在异质性。在技术部署维度,采购合同金额越高,通常意味着技术越先进、部署密度或覆盖范围越高,对公安破案效率的提升效果越强,故技术威慑作用越大。从案件特征来看,轻罪行为的预期收益相对较低,犯罪者大多不是出于严重的主观恶意,风险偏好通常较弱,更易被新技术应用所威慑;同时,前期传统监控设备的大规模普及为公众所熟知,潜在犯罪者容易预测 AI 技术首要依托视频采集与图像分析发挥作用,故 AI 应用对高可见性案件(即易被监控设备捕捉的案件类型)的威慑效应应当更为显著。从犯罪者特征角度,犯罪者的受教育程度(学历高低)、犯罪组织形式(单独或共同作案)、犯罪经历(初犯或累犯)及犯罪地域属性(本地或异地犯罪)等因素,均会影响其信息获取能力、风险感知水平及对技术规避可能性,从而体现出犯罪数量减少程度上的异质性。据此得出:

假设 2 AI 应用对刑事犯罪的抑制效应根据采购合同特征、案件特征和犯罪者特征的不同存在异质性。

从 AI 技术的作用机制分析,其对潜在犯罪者的威慑效应首先经由技术路径实现:AI 技术通过提升公安侦查效率、降低边际执法成本,推动均衡破案率上升,进而影响潜在犯罪者对被捕概率的主观预期。具体而言,AI 技术的引入一方面通过算法辅助(如人像识别、线索分析)提升公安人员的侦查效率与质量,另一方面通过自动化数据处理等替代传统人工劳动,降低边际执法成本。上述双重机制的共同作用,带来刑事案件侦破周期的缩短。因此,AI 应用对犯罪数量的抑制效应首先依托于其提升客观侦查效率的技术能力,本文将这一渠道界定为“技术效应”,并提出:

假设 3 AI 应用显著缩短刑事案件的侦破时间,即技术效应。

在前述技术效应基础上,破案时间的缩短与破案效率的提升形成威慑信号,潜在犯罪者经由公共媒体与社会网络等渠道感知该信号,并据此更新其对被捕概率的主观预期,进而调整犯罪决策。因此,AI 应用对犯罪数量的抑制效应不仅取决于技术效应本身,还取决于该威慑信号作为模因^①被传播的广度和深度,这直接影响潜在犯罪者信念更新的时效性与强度。本文将这一基于信息传播放大威慑信号的渠道界定为“传播效应”,并提出:

假设 4 AI 技术资讯的传播显著增强其对犯罪数量的抑制效应,即传播效应。

^① “模因”为传播学概念,指文化信息传播的基本单位,类似于生物学中的基因,通过复制、变异和选择的过程在社会中扩散,常用于分析网络语言和流行趋势,帮助理解谣言、宗教等现象的扩散机制等(道金斯,2018)。

三、研究设计与数据

(一) 变量

1. AI 采购^①

本文借鉴 Beraja et al.(2023)的供应商主营业务识别法,通过匹配政府采购合同与人工智能企业名录,识别各地级市公安机关的 AI 采购行为。具体识别步骤包括:首先,爬取 2015 年以来所有中国政府采购网公布的全部政府采购合同,根据采购人名称从中筛选出公安机关的采购,并将采购人地区精确到地级市;其次,利用企查查数据库的企业查询功能,依据企业经营范围是否包含 AI 相关的关键词^②,识别企业是否为人工智能类企业,整理形成人工智能企业名录;最后,通过匹配公安采购合同中的供应商名称与人工智能企业名称,识别公安的 AI 采购合同。考虑到 AI 技术从采购到实际部署存在时滞,本文将合同日期精确到季度层面,以平滑 AI 技术从采购到应用的时间差异。

本文将汇总的采购数据绘制成折线图,汇报于附图 A1^③,左右两个子图分别统计月度采购次数和月度累计采购次数。从图中可以看出,近年来各地 AI 相关的采购迅速增加,且具有明显的季节性因素,这与 AI 技术的发展进程和影响采购的客观因素相符。

2. 被解释变量

本文的被解释变量包括两类:一是刑事犯罪数量,二是公安破案时间。鉴于真实的刑事犯罪数量无法直接观测,本文以各地法院刑事裁判文书数量作为其代理变量。需要说明的是,裁判文书仅反映已进入司法程序且公开的刑事案件,可能低估实际犯罪水平。但由于双重差分方法识别的是 AI 采购对犯罪数量的相对变化(处理效应),该测量误差主要影响截距项,对估计量的偏误影响有限。而且针对 2013 年以前裁判文书公开程度系统性不足的问题,本文在稳健性检验中剔除了该时段样本,结果依然稳健(详见后文)。数据来源方面,总体及分样本的刑事案件数量来自检察文书和刑事裁判文书,分样本的案件特征和犯罪者特征通过大语言模型及文本分析方法从上述法律文书中提取得到。公安破案时间定义为犯罪发生日期至公安机关将案件移交检察院审查起诉日期之间的天数。其中,犯罪发生日期提取自刑事裁判文书,移送审查起诉日期

^① 本文采用 AI 采购作为 AI 应用的代理变量。限于数据可得性,地区采购 AI 后的实际使用强度与地域覆盖范围难以直接观测,但本文认为,在市级层面一旦完成采购, AI 技术至少会在部分案件或部分地区投入使用,其处理效应可通过市级总体层面的犯罪数据加以识别。该识别策略与既有文献中利用技术引入或政策启动时点评估治理效果的做法一致,不会对因果推断的可靠性构成系统性威胁。

^② 关键词包括芯片、图像识别、计算机视觉、语音识别、传感器等。

^③ 限于篇幅,附录未在正文列示,感兴趣的读者可在《经济学》(季刊)官网(<https://ceq.ccer.pku.edu.cn>)下载。

提取自检察文书。回归中对上述变量均取自然对数处理。

3. 调节变量

本文的调节变量包括合同采购金额和 AI 侦查技术资讯指数。其中,合同采购金额提取自前述政府采购合同数据,回归中取对数处理;AI 侦查技术资讯指数源自百度资讯指数^①(无单位),汇总至地级市-季度层面,时间范围是 2017—2022 年。

4. 变量描述

本文将上述数据匹配汇总至地级市-季度层面,形成基准回归和异质性分析的主要样本。主要变量的描述性统计汇报于表 1。

表 1 变量描述性统计

变量	观测值	均值	标准差
	(1)	(2)	(3)
刑事犯罪数量(单位:件)			
总体	16 644	56.9521	215.8419
可见	16 644	47.1702	179.8535
不可见	16 644	29.0062	109.6615
轻度	16 644	49.8274	194.1554
中度	16 644	6.3229	20.7502
重度	16 644	1.7633	6.7883
初犯	16 644	50.4954	185.4081
累犯	16 644	7.3202	36.3645
本地	16 644	29.4815	111.9511
异地	16 644	6.6790	40.5206
单犯	16 644	22.5496	114.0806
共犯	16 644	34.8696	146.6607
大学及以上	16 644	0.0439	0.3845
中学及以下	16 644	21.0222	82.3344
公安侦破时间(单位:天)			
总体	4 529	348.1789	1 294.0928
百度资讯指数(无单位)			
AI 侦查技术	10 360	821.1322	2 337.8843

注:本文将涉及公共安全、人身安全、交通安全等可被监控设备捕捉到画面的犯罪类型界定为“可见案件”,反之为“不可见案件”;按照刑期小于 3 年、3 到 10 年以及大于 10 年将犯罪分为轻度、中度和重度三类。

^① 百度作为国内用户规模领先的信息获取平台之一,其资讯指数的数据来源覆盖多元化的信息生产主体,既包括人民日报、新华社等传统权威媒体,也涵盖政府机构、企业以及大量个人创作者,这些内容通过多个渠道分发给网络用户,基于用户的阅读、评论、转发、点赞等行为数据计算得出百度资讯指数,能够客观反映公众对各类信息的被动触达与关注程度。

(二) 识别策略与模型设定

本文运用渐进双重差分法识别 AI 采购对刑事犯罪数量的影响,并构建如下模型:

$$Y_{c yt} = \beta_0 + \beta_1 Post_{c yt} + \mu_c + K'_{c yt} \Gamma + \rho_t + \delta_{cy} + \epsilon_{c pt}, \quad (1)$$

其中, c 表示公安机关所在地级市, y 为年份, t 为年份-季度。具体而言, $Y_{c yt}$ 为被解释变量,表示城市 c 在 y 年 t 季度的刑事犯罪总量。 $Post_{c yt}$ 为核心解释变量,若城市 c 的公安机关在 y 年 t 季度采购了 AI,则当季度及其后各季度取值为 1,否则为 0。为规避多次采购的累积效应,本文仅保留各地首次采购前后的样本,具体处理方法为将各地样本截断至第二次采购之前。 β_1 是本文最关心的系数,反映 AI 采购的处理效应。 μ_c 为城市固定效应,控制地区不随时间变化的因素; ρ_t 为年份-季度时间固定效应,控制宏观时间趋势; δ_{cy} 是城市-年份交互固定效应,控制随城市与年度变化的因素,如地区经济发展水平、公检法机关特征等可能同时影响 AI 采购决策与犯罪数量的变量。 $K_{c yt}$ 为控制变量,包含各城市公安机关累计采购传统(非 AI)摄像头的次数与金额,以控制传统监控设备采购的干扰。

鉴于各地区公安机关 AI 采购时点存在差异,本文采用渐进双重差分模型进行估计。然而,当处理效应存在异质性时,传统双向固定效应(Two-Way Fixed Effects, TWFE)估计量可能产生偏误(De Chaisemartin and D'Haultfoeuille, 2020)。为克服这一问题,本文采用 Callaway and Sant'Anna(2021)提出的异质性稳健事件研究法,以刻画 AI 采购的动态处理效应。

进一步地,本文采用多种方法对基准回归进行稳健性检验,并缓解潜在的内生性担忧。同时,本文从采购合同金额、刑事案件特征与犯罪者特征等多维度展开异质性分析以验证假设 2,并对技术效应与传播效应两个机制进行检验以验证假设 3 和假设 4。

四、实证结果分析

(一) 基准结果

本文首先汇报基准结果,以刑事犯罪数量的对数作为被解释变量,估计结果如表 2 所示。列(1)显示,双重差分项的估计系数为-0.1384,且在 1%水平上显著,鉴于被解释变量取对数形式,该结果可理解为,引入 AI 使当地刑事犯罪数量下降约 14%。列(2)展示在控制传统监控摄像头累计采购次数与累计采购金额后的回归结果,AI 采购的估计系数变化不大,且仍在 1%水平上显著。上述结果表明,本文的假设 1 得到验证,即 AI 应用对刑事犯罪数量具有显著的

抑制作用。需要指出,表 2 中的估计结果可能存在系统性低估(详见后文机制分析),真实效应或更高。

表 2 公安机关 AI 采购对刑事犯罪数量的影响:基准回归

被解释变量	刑事犯罪数量(件,对数)	
	(1)	(2)
AI 采购	-0.1384*** (0.0466)	-0.1180*** (0.0387)
传统摄像头累计采购次数(对数)		是
传统摄像头累计采购金额(对数)		是
城市固定效应	是	是
时间(年份-季度)固定效应	是	是
城市×年份固定效应	是	是
观测值	16 626	16 626
R ²	0.9789	0.9789

注:括号内为聚类到城市层面的稳健标准误,*、**和*** 分别代表在 10%、5%和 1%的水平上显著,下同。

(二) 事前趋势检验与稳健性检验

1. 事前趋势检验

图 2 展示了基准回归的事件研究法结果,横坐标为距离 AI 采购的相对季度,纵坐标为各期估计系数及其 95%置信区间。结果表明:其一, AI 采购前各期估计系数与零值无显著差异,未拒绝事前趋势平行的假设;其二, AI 采购当期,实验组地区的刑事犯罪数量显著下降,表明 AI 应用在短时间内即产生犯罪抑制效应;其三,第四期(约一年后)起标准误明显增大,95%置信区间逐渐拓宽并包含零值,长期效应的统计显著性有所减弱。这主要源于本文仅保留首次采购样本,后期样本量损失导致估计精度下降。尽管如此,各期点估计持续为负,且绝对值呈扩大趋势,表明处理效应的方向稳定,其实际影响程度可能随时间推移而增强。

2. 稳健性检验

为确保基准回归结果的可靠性,本文从多个维度展开稳健性检验。

(1) 样本选择偏差。鉴于 2013 年以前裁判文书的公开程度较低且存在地区差异,可能引发样本选择问题。为此,本文剔除 2013 年以前的样本重新进行估计,结果汇报于附表 A1。剔除早期样本后,核心解释变量的系数符号与显著性均未发生实质性改变,表明基准结果在样本选择层面具有稳健性。

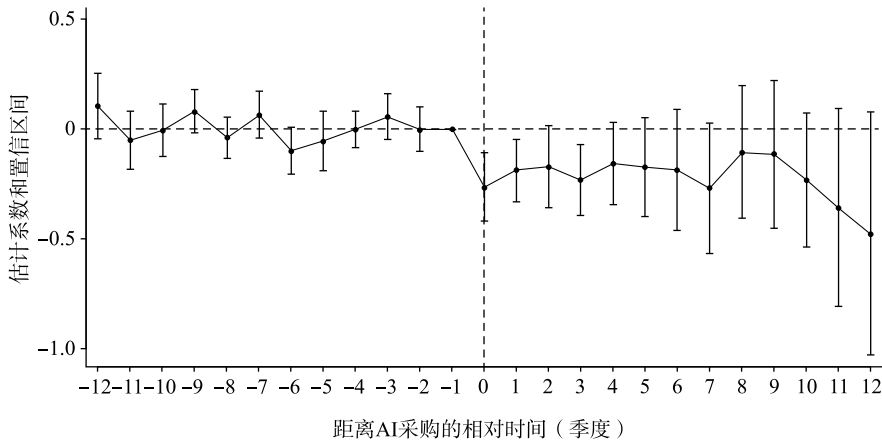


图2 基于异质性稳健估计量(Callaway and Sant'Anna, 2021)的事件研究估计结果

(2) 同期政策干扰。与AI部署同期实施的其他重大公共安全政策可能带来遗漏变量偏误。为此,本文对2010—2022年间中央层面在公共安全、公安治理及科技强警领域实施的重大政策与专项行动进行系统性梳理。结果显示,上述政策均由中央统一部署、在全国范围内同步推行,不存在区域差异化安排,其影响可被时间固定效应吸收,不构成本文的混淆因素。

(3) 识别策略调整。基准回归以供应商是否为人工智能企业作为AI采购的识别依据,该做法可能将部分非AI相关采购纳入处理组。为进一步提高识别精度,本文在供应商主营业务识别法的基础上,辅以关键词匹配法筛选采购标的是否确属AI产品或服务^①。鉴于关键词筛选法可能漏识部分表述模糊的真实AI采购,本文剔除未通过关键词匹配的样本重新估计,结果汇报于附表A2,基准结论依然稳健。

(4) 反向因果问题。AI采购前的犯罪数量可能影响当地采购决策(如犯罪高发地区可能更早引入AI技术),从而引发反向因果。借鉴Beck et al.(2010)的处理思路,本文考察了各地AI采购前犯罪数量均值与首次AI采购时间的关系,结果显示二者不存在显著相关性(详见附图A2),可排除采购前犯罪数量对采购决策的干扰。

(5) 竞争性解释。本文进一步考察公安局长晋升激励的异质性影响。结果表明,晋升激励对基准回归结果无显著作用,这意味着AI应用降低犯罪数量的效应主要通过影响潜在犯罪者的主观感知实现,而非由公安机关内部激励驱动,具体结果汇报于附表A3。

^① 关键词基于国家知识产权局《关键数字技术专利分类体系(2023)》中包含的所有人工智能技术的分类与关键词去重整理得到,共266个。

（三）安慰剂检验

传统(非 AI)监控设备在 AI 应用之前已在全国范围内广泛部署,其功能特性与作用边界对潜在犯罪者而言属于已知信息,难以触发其主观预期的更新,故具备作为安慰剂的理想特征。此外,由于传统监控设备采购亦可能通过提升破案效率而影响犯罪数量,有必要将两类技术置于同一分析框架下对其效应加以区分。为此,本文将 AI 采购与传统摄像头采购同时纳入双重差分回归模型,此时两项的估计系数分别对应排除另一技术干扰后的“净效应”。表 3 汇报安慰剂检验的结果。列(1)以刑事犯罪数量的对数为被解释变量,结果显示,传统监控设备采购项系数在统计上不显著,表明其难以形成犯罪威慑,这与理论预期相符。相比之下, AI 采购项的系数为-0.1389 且在 1%水平上显著,验证了基准结果的稳健性。

表 3 安慰剂检验

被解释变量	刑事犯罪数量(件,对数)
	(1)
AI 采购	-0.1389*** (0.0419)
传统监控设备采购	0.0314 (0.0673)
城市固定效应	是
时间(年份-季度)固定效应	是
城市×年份固定效应	是
观测值	16 531
R ²	0.9790

（四）异质性分析

1. 基于合同特征的异质性分析

本文首先考察采购合同金额的调节效应。通常认为,合同金额越大,意味着采购标的的规模或技术先进程度越高,其对犯罪数量的抑制效应可能更强。为此,本文将合同总金额取对数后与双重差分项进行交乘,以获得采购金额对处理效应的边际影响。回归结果汇报于附表 A4。结果显示,三重差分项显著为负,表明采购金额的增加强化了 AI 技术对犯罪数量的削减作用。

2. 基于案件特征的异质性分析

其次,本文考察刑事案件特征的调节作用。不同类型犯罪受 AI 技术的影

响程度可能存在差异,具体包括:①犯罪严重程度。轻罪的预期收益相对更低,其潜在犯罪者更易因受到威慑而减少犯罪活动,故预期AI采购后轻罪数量的减少程度更高。②是否为可见犯罪。基于前文分析,能够被监控画面捕捉到的犯罪应为AI技术发挥作用的主要场景,故预期AI对可见案件数量的减少作用更强。

本文按上述特征分组统计案件数量,并分别进行式(1)所示的回归。表4汇报基于案件特征的异质性分析结果。列(1)–(3)显示,轻度、中度与重度案件的双重差分项系数分别为 -0.0784 、 -0.0479 和 -0.0372 ,绝对值与显著性依次递减。这表明AI采购带来的犯罪数量减少主要集中于轻度案件,对中度案件的影响明显衰减,而对重度案件的影响在统计意义上已不显著(尽管系数为负且经济意义仍可观)。该结果印证了理论假说:轻罪的潜在犯罪者会因预期被捕概率提升而调整犯罪决策,而重罪实施者对技术威慑的敏感则相对较低。列(4)与列(5)分别报告可见案件与不可见案件的估计结果。AI采购项系数依次为 -0.0903 (在5%水平上显著)与 -0.0583 (在10%水平上显著),表明AI应用对两类案件均有抑制作用,但对可见案件的影响在经济意义和统计意义上均更加显著。这不仅支持威慑效应主要通过犯罪者心理渠道发挥作用,同时也从侧面说明AI技术发挥犯罪抑制作用的直接机制来源于技术效应。

表4 基于案件特征的异质性分析

被解释变量	刑事犯罪数量(件,对数)				
	案件严重程度			案件是否可见	
	轻度	中度	重度	可见	不可见
	(1)	(2)	(3)	(4)	(5)
AI采购	-0.0784^{**} (0.0327)	-0.0479^* (0.0270)	-0.0372 (0.0286)	-0.0903^{**} (0.0351)	-0.0583^* (0.0321)
传统摄像头累计采购次数(对数)	是	是	是	是	是
传统摄像头累计采购金额(对数)	是	是	是	是	是
城市固定效应	是	是	是	是	是
时间(年份-季度)固定效应	是	是	是	是	是
城市×年份固定效应	是	是	是	是	是
观测值	16 626	16 626	16 626	16 626	16 626
R ²	0.9793	0.9682	0.9295	0.9791	0.9786

3. 基于犯罪者特征的异质性分析

最后,本文考察犯罪者特征的异质性。根据理论分析,AI技术对犯罪数量的抑制效应来源于潜在犯罪者信念更新之下的自主决策调整,因此,潜在犯罪者的信息获取能力与风险认知水平均会影响该效应的强弱。本文运用大语言模型及文本分析方法,从刑事法律文书中提取刑事案件主要犯罪者的特征信息,对其认知水平与风险感知状况等加以刻画。

具体而言,本文识别的犯罪者特征包括:①受教育水平,划分为中学及以下(包括高中、初中、小学及文盲)与大学及以上两组。预期较低学历犯罪者因信息渠道有限而对AI技术认知不足,其主观被捕概率预期提升更为显著,故AI威慑效应应当更强。②犯罪者人数结构,区分单独犯罪与共同犯罪。共同犯罪通常具有更强的预谋性与组织化特征,被威慑而中止犯罪的概率较低,且案件侦破难度更大,故预期AI应用对单独犯罪的抑制效应更为显著。③犯罪经历,区分初犯与累犯。初犯面临的犯罪成本与心理门槛更高,对被捕风险更为敏感,而累犯通常具有更高的风险偏好,对AI威慑的敏感度相对较低。④犯罪地域属性,区分本地犯罪与异地犯罪。本地犯罪者通常掌握更多属地信息与社会支持网络,对新技术的感知相对更弱;而异地犯罪者因信息劣势,对被捕概率的预期更高,故威慑效应可能更强。

表5汇报基于犯罪者特征的分组回归结果,被解释变量为刑事犯罪数量的对数。列(1)显示,高学历犯罪者的AI采购项系数不具有统计显著性;列(2)显示低学历犯罪者的该系数显著为负,表明威慑效应仅对低学历犯罪者显著。列(3)与列(4)对比显示,共同犯罪的系数不显著,而单独犯罪的系数显著为负。列(5)与列(6)对比显示,累犯的系数显著性较弱,而初犯的AI采购项系数在经济意义和统计意义上均更为显著。列(7)与列(8)对比显示,本地犯罪者的系数显著性弱于异地犯罪者。上述结果表明,低学历、单独实施犯罪、初次实施犯罪及在异地实施犯罪的犯罪者受AI威慑更为显著,与预期相符。综合来看,AI技术的威慑强度(AI应用对犯罪数量的减少程度)与犯罪者的风险感知能力呈负相关关系,即信息劣势越明显的犯罪群体,其受AI威慑而减少犯罪的程度越高。

根据以上分析,本文假设2得到验证:AI应用对犯罪数量的抑制效应根据采购合同特征、案件特征和犯罪者特征的不同而存在显著的异质性,具体表现为,在采购金额更高的地区,对犯罪程度更轻的、可见性更强的案件,以及对学历更低、单独作案、初次作案、异地作案的犯罪者,AI技术的威慑效应更为显著。

表 5 基于犯罪者特征的异质性分析

被解释变量	刑事犯罪数量(件,对数)							
	犯罪者学历		犯罪者人数		犯罪次数		犯罪者户籍	
	大学及以上	中学及以下	共犯	单犯	累犯	初犯	本地	异地
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AI 采购	0.0049 (0.0090)	-0.0832** (0.0346)	-0.0267 (0.0397)	-0.1543*** (0.0384)	-0.0465* (0.0267)	-0.0751** (0.0325)	-0.0649* (0.0342)	-0.0732** (0.0312)
传统摄像头累计采购次数(对数)	是	是	是	是	是	是	是	是
传统摄像头累计采购金额(对数)	是	是	是	是	是	是	是	是
城市固定效应	是	是	是	是	是	是	是	是
时间(年份-季度)固定效应	是	是	是	是	是	是	是	是
城市×年份固定效应	是	是	是	是	是	是	是	是
观测值	16 626	16 626	16 626	16 626	16 626	16 626	16 626	16 626
R ²	0.7432	0.9786	0.9802	0.9668	0.9694	0.9793	0.9783	0.9585

(五) 机制分析

前文理论分析指出, AI 应用通过技术效应与传播效应抑制刑事犯罪数量, 二者是核心假设成立的前提条件。为验证技术效应, 本文将被解释变量替换为刑事案件的侦破时间(取对数), 以识别 AI 应用对公安破案效率的影响, 其余变量设置与基准回归保持一致。为验证传播效应, 本文基于国家知识产权局发布的《关键数字技术专利分类体系(2023)》, 选取公众认知度较高且已广泛应用于公安侦查的人工智能技术关键词^①, 将其百度资讯指数进行加总, 构建“AI 侦查技术资讯指数”(AI_index), 作为影响潜在犯罪者被捕概率预期的信息传播度的代理变量。^② 本文将该指数与 Post_{cyt} 项交乘, 进行式(2)所示的回归, 以检验相关资讯的媒体传播力度对犯罪数量减少的加强作用。

$$Y_{cyl} = \beta_0 + \beta_1 Post_{cyl} \times AI_index_{cyl} + \beta_2 Post_{cyl} + \beta_3 AI_index_{cyl} + \mu_c + K'_{cyl} \Gamma + \rho_t + \delta_{cy} + \epsilon_{cyl}. \quad (2)$$

① 这些关键词包括: 人脸识别、指纹识别、DNA 检测、虹膜识别、智能摄像头、智能监控。

② 理论上, 最能直接影响潜在犯罪者对于被捕概率预期的信息应当为“AI 技术显著提升破案效率”类新闻报道; 但受限于数据可得性(无法获取网络公开的所有新闻资讯并筛选 AI 赋能破案相关内容)及公共安全问题敏感性(主流舆情平台未收录“AI 破案”等关键词)等, 难以直接度量。尽管本文构建的“AI 侦查技术资讯指数”不仅包含“AI 助力破案”这类特定信号, 亦涵盖部分仅描述技术本身而未明示破案效果的报道, 但公众可基于常识推断, 此类技术一旦部署于公安侦查, 将显著增强侦查能力并提高破案效率, 进而更新其被捕概率预期。

表6汇报技术效应和传播效应的估计结果。列(1)以刑事案件侦破时间的对数为被解释变量,结果显示AI采购项的系数为-0.6459,且在1%水平上显著,表明AI应用使当地刑事案件的侦破时间平均缩短约65%,公安破案效率得到显著提升,假设3(技术效应)得证。事件研究法结果汇报于附图A3,与事前趋势平行的假设一致,且该效应随时间推移持续显现并不断增强。列(2)汇报AI采购与资讯指数交乘的三重差分回归结果,被解释变量为刑事犯罪数量的对数。三重差分项系数为-0.5033,且在5%水平上显著,表明传播效应存在。鉴于AI侦查技术资讯指数包含部分与破案应用无关的噪声信息,可能导致估计结果存在向下偏误,真实效应或更为显著。

此外,结合基准回归结果可知,若犯罪总量在一定时期内保持稳定且公安侦查资源处于饱和状态,仅考虑技术效应,破案效率提升将带来侦破案件数量的增加。然而,威慑效应促使潜在犯罪者放弃犯罪行为,表现为刑事犯罪数量的减少。综合而言,观测到的总破案数量减少表明,AI应用对犯罪数量的抑制效应确实通过潜在犯罪者的心理威慑机制实现,且鉴于技术效应本身具有提升案件侦破数量的正向作用,观测到的刑事犯罪数量减少效应实际上存在低估。

表6 机制分析:技术效应与传播效应

被解释变量	刑事案件侦破时间(天,对数)	刑事犯罪数量(件,对数)
	(1)	(2)
AI采购	-0.6459*** (0.0683)	-0.1326** (0.0577)
AI采购 × AI侦查技术资讯指数		-0.5033** (0.2385)
传统摄像头累计采购次数(对数)	是	是
传统摄像头累计采购金额(对数)	是	是
城市固定效应	是	是
时间(年份-季度)固定效应	是	是
城市 × 年份固定效应	是	是
观测值	4 194	5 214
R ²	0.6286	0.9845

五、政策建议与研究展望

本研究为深化AI在公共安全治理领域的应用、推进国家治理体系现代化提供了以下政策参考:第一,优化AI技术的空间布局与应用激励。建议优先在

交通枢纽、商业中心及城市广场等犯罪高发的公共空间重点部署,并建立技术应用效能与警务绩效考核的联动机制,将 AI 使用效果纳入基层警务工作评价体系,以提升一线部门的技术采纳动力。第二,完善数据治理与制度保障。建议建立完善数据采集与使用的规范框架,在提升治安效能的同时保障公民隐私,平衡技术应用的效率与伦理边界。第三,强化技术融合与信息传播。建议推动 AI 技术与传统警务模式的深度融合,构建多层次治安防控网络;通过适度公开 AI 监控覆盖范围、发布典型破案案例等方式强化信息传播,放大技术应用的犯罪威慑效应;建立动态技术更新机制,及时引入新一代 AI 技术以保证犯罪威慑效应的持续性。

本文聚焦 AI 采购对刑事犯罪治理效能的影响,但存在以下局限:其一,政府在 AI 项目上的财政投入可能对其他公共安全或社会服务支出产生挤出效应。限于数据可得性,本文尚未充分评估 AI 应用的社会净福利效应与成本收益比。其二,本文虽识别了 AI 应用的整体威慑效应,但未深入探讨不同类型 AI 技术(如通用型、计算机视觉型、大数据分析型等)的差异化作用机制。未来研究可在获取更细颗粒度数据的基础上,从社会福利视角系统评估 AI 应用的综合效应,并进一步比较不同技术路径的治理效能差异。

参考文献

- [1] Anderson, D. A., "The Aggregate Burden of Crime", *The Journal of Law and Economics*, 1999, 42(2), 611-642.
- [2] Anwar, S. and T. Loughran, "Testing a Bayesian Learning Theory of Deterrence Among Serious Juvenile Offenders", *Criminology*, 2011, 49(3), 667-698.
- [3] Assunção, J., C. Gandour, and R. Rocha, "Deterring Deforestation in the Amazon: Environmental Monitoring and Law Enforcement", *American Economic Journal: Applied Economics*, 2023, 15(2), 125-156.
- [4] Bachman, R., R. Paternoster, and D. Ward, "The Rationality of Sexual Offending: Testing a Deterrence/Rational Choice Conception of Sexual Assault", *Law and Society Review*, 1992, 26(2), 343-372.
- [5] Banerjee, A., R. Hanna, B. A. Olken, E. Satriawan, and S. Sumarto, "Electronic Food Vouchers: Evidence from an At-Scale Experiment in Indonesia", *American Economic Review*, 2023, 113(2), 514-547.
- [6] Beck, T., R. Levine, and A. Levkov, "Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States", *The Journal of Finance*, 2010, 65(5), 1637-1667.
- [7] Becker, G. S., "Crime and Punishment: An Economic Approach", *Journal of Political Economy*, 1968, 76(2), 169-217.
- [8] Beraja, M., D. Y. Yang, and N. Yuchtman, "Data-intensive Innovation and the State: Evidence from AI Firms in China", *The Review of Economic Studies*, 2023, 90(4), 1701-1723.

- [9] Blanes i Vidal, J., and T. Kirchmaier, “The Effect of Police Response Time on Crime Clearance Rates”, *The Review of Economic Studies*, 2018, 85(2), 855-891.
- [10] Callaway, B., and P. H. Sant’Anna, “Difference-in-Differences with Multiple Time Periods”, *Journal of Econometrics*, 2021, 225(2), 200-230.
- [11] Cameron, L., X. Meng, and D. Zhang, “China’s Sex Ratio and Crime: Behavioural Change or Financial Necessity?”, *Economic Journal*, 2019, 129(618), 790-820.
- [12] Chalfin, A., “Economic Costs of Crime”, *The Encyclopedia of Crime and Punishment*, 2015, 1-12.
- [13] Cohen, I., “Technology and the State: Building Capacity to Tax via Text”, *Journal of Public Economics*, 2024, 236, 105154.
- [14] Cohen, M. A., *The Costs of Crime and Justice*. Abingdon;New York;Routledge, 2020.
- [15] Corman, H., and H. N. Mocan, “A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City”, *American Economic Review*, 2000, 90(3), 584-604.
- [16] De Chaisemartin, C., and X. D’Haultfoeuille, “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects”, *American Economic Review*, 2020, 110 (9), 2964-2996.
- [17] Di Tella, R., and E. Schargrodsky, “Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack”, *American Economic Review*, 2004, 94(1), 115-133.
- [18] Doleac, J. L., “The Effects of DNA Databases on Crime”, *American Economic Journal: Applied Economics*, 2017, 9(1), 165-201.
- [19] Dzansi, J., A. Jensen, D. Lagakos, and H. Telli, “Technology and Local State Capacity: Evidence from Ghana”, 2022, *National Bureau of Economic Research*.
- [20] Edlund, L., H. Li, J. Yi, and J. Zhang, “Sex Ratios and Crime: Evidence from China”, *Review of Economics and Statistics*, 2013, 95(5), 1520-1534.
- [21] Gardner, M., and L. Steinberg, “Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: An Experimental Study”, *Developmental Psychology*, 2005, 41(4), 625-635.
- [22] Greenstone, M., G. He, R. Jia, and T. Liu, “Can Technology Solve the Principal-Agent Problem? Evidence from China’s War on Air Pollution”, *American Economic Review: Insights*, 2022, 4(1), 54-70.
- [23] Guriev, S., N. Melnikov, and E. Zhuravskaya, “3G Internet and Confidence in Government”, *The Quarterly Journal of Economics*, 2021, 136(4), 2533-2613.
- [24] He, G., and W. Peng, “Guns and Roses: Police Complicity in Organized Prostitution”, *Journal of Public Economics*, 2022, 207(1), 104599.
- [25] Jabri, R., “Algorithmic Policing”, *SSRN Electronic Journal*, 2021.
- [26] Jin, G. Z., and J. Lee, “Inspection Technology, Detection, and Compliance: Evidence from Florida Restaurant Inspections”, *The Rand Journal of Economics*, 2014, 45(4), 885-917.
- [27] Kessler, D. and S. Levitt, “Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation”, *Journal of Law and Economics*, 1999, 42(1), 343-363.
- [28] Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan, “Human Decisions and Machine Predictions”, *The Quarterly Journal of Economics*, 2018, 133(1), 237-293.
- [29] Levitt, S. D., “Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime”,

- American Economic Review*, 1997, 87(3), 270-290.
- [30] Levitt, S. D., "Understanding Why Crime Fell in the 1990s: Four Factors That Explain the Decline and Six That Do Not", *Journal of Economic Perspectives*, 2004, 18(1), 163-190.
- [31] [英]理查德·道金斯,《自私的基因》,卢允中等译。北京:中信出版集团,2018年。
- [32] Ludwig, J., and S. Mullainathan, "Machine Learning as a Tool for Hypothesis Generation", *The Quarterly Journal of Economics*, 2024, 139(2), 751-827.
- [33] Ma, H., M. Xu, W. You, and J. Feng, "Keeping an Eye on the Villain: Assessing the Impact of Surveillance Cameras on Crime", *Journal of Development Economics*, 2026, 178, 103557.
- [34] Maccoun, R., and P. Reuter, "Drug War Heresies: Learning from Other Vices, Times, and Places", *Cambridge University Press*, 2009, 1(1), 1-300.
- [35] Mastrobuoni, G., "Police Disruption and Performance: Evidence from Recurrent Redeployments Within a City", *Journal of Public Economics*, 2019, 176, 18-31.
- [36] Mastrobuoni, G., "Crime Is Terribly Revealing: Information Technology and Police Productivity", *The Review of Economic Studies*, 2020, 87(6), 2727-2753.
- [37] Matsueda, R., D. Kreager, and D. Huizinga, "Deterring Delinquents: A Rational Choice Model of Theft and Violence", *American Sociological Review*, 2006, 71(1), 95-122.
- [38] Muralidharan, K., P. Niehaus, and S. Sukhtankar, "Building State Capacity: Evidence from Biometric Smartcards in India", *American Economic Review*, 2016, 106(10), 2895-2929.
- [39] Nissan, E., "Digital Technologies and Artificial Intelligence's Present and Foreseeable Impact on Lawyering, Judging, Policing and Law Enforcement", *AI & Society*, 2017, 32, 441-464.
- [40] Pogarsky, G., A. Piquero, and R. Paternoster, "Modeling Change in Perceptions About Sanction Threats: The Neglected Linkage in Deterrence Theory", *Journal of Quantitative Criminology*, 2004, 20(4), 343-369.
- [41] Ranasinghe, A., and D. Restuccia, "Financial Frictions and the Rule of Law", *Journal of Development Economics*, 2018, 1-24.
- [42] Sjoquist, D., "Property Crime and Economic Behavior: Some Empirical Results", *American Economic Review*, 1973, 63(3), 439-446.
- [43] Stigler, G. J., "The Optimum Enforcement of Laws", *Journal of Political Economy*, 1970, 78(3), 526-536.
- [44] Tauchen, H., A. D. Witte, and H. Griesinger, "Criminal Deterrence: Revisiting the Issue with a Birth Cohort", *The Review of Economics and Statistics*, 1994, 76(3), 399-412.
- [45] 解源源、史全增,“基层公安机关警力不足的类型化分析及改革路径”,《中国人民公安大学学报(社会科学版)》,2014年第4期,第37—43页。
- [46] 熊一新,“警务改革背景下我国警务辅助力量建设——以英国和我国香港特别行政区辅警制度及警务改革为视角”,《中国人民公安大学学报(社会科学版)》,2014年第4期,第1—16页。

AI Application and Public Security Governance: Evidence from Criminal Cases

LIU Jiaqi

(Tsinghua University)

ZHOU Li'an HE Qiang*

(Peking University)

Abstract: This study examines the impact of artificial intelligence (AI) on crime governance. Results show that AI application by public security agencies significantly reduces crime. Heterogeneity analysis reveals stronger effects for minor, visible cases, and among offenders with lower education levels, those acting alone, first-time offenders, and those committing crimes outside their usual jurisdiction. Mechanism analysis indicates that AI accelerates investigations and raises offenders' perceived probabilities of arrest through media channels. Our findings provide empirical evidence on how AI enhances public safety governance.

Keywords: artificial intelligence; public security governance; crime deterrence

JEL Classification: K42, O33, H40

* Corresponding Author: HE Qiang, Guanghua School of Management, Peking University, No. 5 Yiheyuan Road, Haidian District, Beijing 100871, China; Tel: 86-18034998237; E-mail: heqiangnf@163.com.