

统筹发展与安全的数据治理 路径研究*

续 继 杨卿栩 侯嘉奕

摘要:以高效能数据治理促进数字经济高质量发展与高水平数据安全良性互动具有迫切性与必要性。然而,不少国际研究认为数据治理的安全成效往往以牺牲发展为代价,建立统筹发展与安全的数据治理路径并非易事。为评价中国数据治理实践,本文应用APP产品安全检测大数据和双重差分模型,对《关于引导规范教育移动互联网应用有序健康发展的意见》进行政策效应分析。研究发现:第一,该政策显著降低了教育APP的安全漏洞数和数据违规行为数,有效提升了企业数据安全风险防范能力和规范管理水平;第二,该政策有助于维护市场秩序,倒逼高风险APP加速安全转型,缩小了与低风险APP的数据安全差距;第三,该政策对更具市场地位的企业和更受市场关注的产品实现了更有效的规制;第四,该政策的推行提升了应用市场规模,助力资本实力相对薄弱和用户资源相对欠缺的APP进入市场。研究表明,以备案制度降低准入门槛、以协同联动打破条块分割、以共建共治共享提升管治水平的治理路径,能够统筹发展与安全,实现数字经济市场活力与秩序的平衡、经济效率与社会公平的兼顾。

关键词:数据治理 统筹发展与安全 数字经济高质量发展 高水平数据安全

一、引言

伴随着数字经济的蓬勃发展,数据流通与使用日益频繁,网络安全风险和安全隐患逐渐增多,数据治理愈发关系着国家的长治久安和人民的福祉增进,在贯彻总体国家安全观的前提下强化数据治理实践对于推进国家安全体系和能力现代化具有重要意义。党的二十届三中全会提出“加快构建促进数字经济发展体制机制”和“提升数据安全治理监管能力”,为完善数据治理体系指明了方向。高效能数据治理不仅要为数据风险防控提供托底保障,维护国家安全、公共利益和个人权益,确保高水平的数据安全,还应促进数据的高效利用和流通,激发数据要素创新活力,推动数字经济的高质量发展。在数据资源开放与共享成为数字经济运行核心的当下(陈晓红等,2022),数字经济的高质量发展与高水平数据安全休戚相关。数据治理推进数据更高效、更安全流通,可以促进数据要素创造更大的社会价值(黄璜,2020),带动数字经济的发展;而数字经济繁荣又可以反哺数据要素的价值发挥(陈雨露,2023),为数据安全提供技术保障。探索以推进数字经济高质量发展和高水平数据安全良性互动为目标的数据治理路径具有较高的可行性和必要性。

国际上对数据治理的实践路径展开了广泛探讨,提出了提升安全治理效能的相关建议,其中包括落实主体责任和构建奖惩机制(戈登等,2015;阿杰里德等,2016)、增强数据安全隐患通报和声誉制裁(李等,2018;唐、温斯顿,2020)、建立统一的数据安全法规体系和执行指南(亚伯拉罕等,2019)以及建立与市场发展趋势相匹配的监管机制(宾斯、别蒂,2020)。这些研究为防范数据安全风险和规范数据管理指明了方向,对于矫正企业数据安全投入不足和数据滥用所产生的负向外部性具有助力。然而,建设统筹发展与安全的数据治理体系并非易事,不少国际实证研究和准实验研究发现,数据治理所产生的安全成效是以牺牲发展为代价的。数据安全合规增加了企业的运营成本,导致了企业进入数字经济领域的意愿下降(詹森等,2022),并且相关成本被不成比例增加到了小型企业和初创企业上(坎贝尔等,2015),为大企业建立“围墙花园”进行数据壅

收稿时间:2025-1-6;反馈外审意见时间:2025-4-11、2025-5-19;拟录用时间:2025-12-3。

*本研究得到国家社会科学基金重大项目(基金号:2024MZD026)的资助。侯嘉奕为本文通讯作者。

重大选题征文

断提供了借口,进一步加剧了垄断、阻碍了创新(布雷尔等,2020)。而繁杂的数据合规条例亦会反向激励用户与高安全风险的企业达成私下共识,无视监管规定(亚伯拉罕等,2019),这不仅使得用户承担了更多的安全风险,也打击了企业遵守规则的积极性,相关治理模式难以为继。由此,处理好发展与安全的关系仍是完善数据治理方案所面临的重要且棘手的议题。面对这一棘手议题,中国作为数字经济大国,具有巨大的数据要素市场发展潜力,在全球数据治理中扮演了重要角色(范柏乃、盛中华,2024),也可为世界贡献中国数据治理的实践经验。

近年来,中国移动互联网快速发展,引领全球数字通信产业变革。2020年末,中国手机网民规模已接近10亿人,占整体网民比例接近100%,移动互联网的普及为在线教育蓬勃发展创造了契机,手机在线教育用户规模达3.4亿,教育移动互联网应用(以下简称“教育APP”)显著改善了学习的灵活性和可及性,使终身学习和碎片化学习成为可能。然而,移动互联网市场的快速壮大也使数据治理面临更严峻的挑战。一方面,APP安全漏洞可能引发黑客攻击、恶意代码植入等安全事件;另一方面,部分APP存在超范围收集用户信息、违规使用个人数据等行为,威胁用户隐私安全。这些数据安全隐患若不能及时解决,将对社会安全和人民生活带来较大威胁。值得欣喜的是,伴随我国数据治理实践从点到面逐步深化,数据安全风险得以较好控制。其中,2019年8月由教育部、中央网信办、工业和信息化部、公安部等八部门联合发布实施的《关于引导规范教育移动互联网应用有序健康发展的意见》(以下简称《规范教育APP意见》)作为首个全国性教育APP治理政策,在实践中取得了显著成效(如图1所示)。

《规范教育APP意见》旨在引导和规范教育APP健康有序发展,该政策将统筹发展与安全作为重要指引,针对企业存在的数据安全风险防范不足和数据管理不规范问题,提出了系统性的引导与纠正举措。在治理理念上,该政策以正确处理政府与市场、管理与服务、安全与发展的关系为基本原则,部署了构建一体化治理框架、完善奖惩机制、强化风险预警和推动精准治理等举措,与国际文献倡导的治理理念相吻合。在实施路径上,该政策进行了备案制度建立、综合协同体系搭建、监督渠道拓展等创新探索,积累了推动有效市场和有为政府更好结合、不搞“一刀切”的宝贵治理经验。这一先行政策实践使得教育APP的数据安全性有较大幅度提升。

此后,中国逐渐从重点行业、重点领域的监管探索转向全局性的治理布局。2022年12月,《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》(简称“数据二十条”)明确提出要贯彻总体国家安全观,将“形成政府监管与市场自律、法治与行业自治协同、国内与国际统筹的数据要素治理结构”作为构建数据基础制度的工作原则。2023年2月,《工业和信息化部关于进一步提升移动互联网应用服务能力的通知》将强化APP运行管理、提高APP风险预警能力等作为筑牢终端安全防线的抓手,将准确登记信息作为夯实接入企业责任的保障举措。2023年7月,《工业和信息化部关于开展移动互联网应用程序备案工作的通知》将维护网络安全和公共利益、促进互联网行业规范健康发展作为总体要求,在全行业APP中推广备案工作。党的二十届三中全会以来,中国统筹发展和安全的数据治理路径更为清晰。《中共中央办公厅 国务院办公厅关于加快公共数据资源开发利用的意见》于2024年9月发布,提出“厘清部门和管理边界,逐步形成权责清晰、条块协同的公共数据资源开发利用格局”。《网络安全管理条例》于2025年1月正式实施,将统筹发展和安全作为总体思路,进一步细化“协同配合、信息沟通”的数据安全监督检查机制和畅通

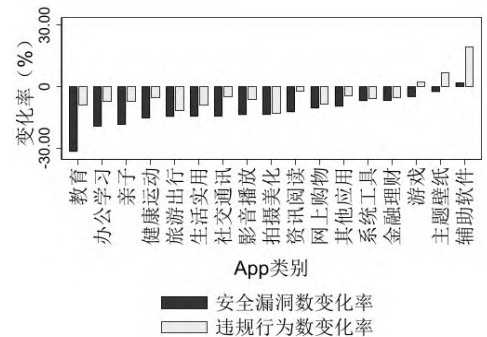
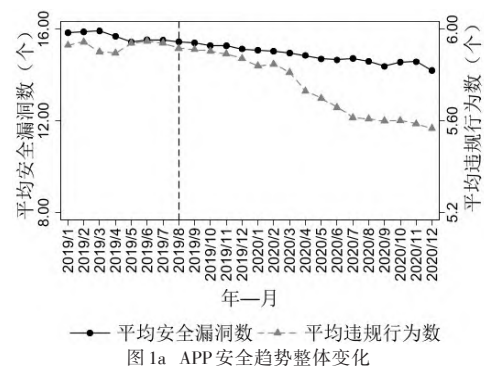


图1 2019~2020年数据安全变化

注:数据来源为APP产品安全检测大数据,图1b安全漏洞数和违规行为数的变化率分别表示2020年1月平均安全漏洞数和平均违规行为数相较于2019年1月的同比变化率。

网络数据安全投诉、举报渠道。《规范教育APP意见》所实践的备案制改革、条块协同和共建共治共享亦在上述治理中得以深化。由此,《规范教育APP意见》是在践行全球治理倡议背景下进行的中国特色数据治理路径探索,其经验做法具有较高代表性。

为分析我国数据治理实践,本文基于2019年1月~2020年12月APP产品安全检测大数据,应用双重差分模型(DID)对《规范教育APP意见》政策效应展开分析。本文研究发现《规范教育APP意见》推行强化了数据安全监管力度,教育APP相较于非教育类APP在安全漏洞数和违规行为数上均有显著下降,且这一结果相对稳健。同时,《规范教育APP意见》较好地统筹了发展与安全,对市场秩序维护、市场公平保障和市场业态培育产生积极作用,可以倒逼高风险APP加速安全转型,还使得资本实力较弱、用户资源较少的APP面对更包容的监管环境,市场进入率得以提升。由此可以说明,以备案制度降低准入门槛、以协同联动打破条块分割、以共建共治共享提升管治水平的治理路径可以统筹发展与安全,做到平衡数字经济市场活力与秩序、兼顾经济效率与社会公平。

本文的创新点和主要贡献如下。第一,本文应用覆盖面广、代表性强的APP产品安全检测大数据,细致分析了APP应用市场的发展态势和安全动态,为掌握我国数字经济发展和数据安全现状与趋势提供了量化支撑。第二,本文对具有较高政策代表性且仍较少被学界关注的《规范教育APP意见》进行了细致分析,该政策是中国长久以来治理智慧的缩影,深入研究该政策有助于对中国特色的数据治理路径进行理论提炼。同时,鉴于国内数据治理领域的定量研究相对匮乏,本文的实证分析可为相关文献提供有益补充。第三,本文从统筹发展与安全的视角,综合且全面进行了政策效应评估。以往多数国际文献在探讨数据治理时仅关注安全或发展单一维度,且多对治理路径难以平衡发展与安全进行批判,而本文实证研究则印证了中国在统筹发展与安全上做出的贡献,为构建中国视角下的数据治理理论研究提供思考。

本文后续安排如下:第二部分为文献综述;第三部分为政策背景,对《规范教育APP意见》施策重点和治理经验进行介绍;第四部分为数据说明和计量模型,介绍本文所使用的APP产品安全检测大数据、变量构造方法和研究识别策略等;第五部分为实证分析,对《规范教育APP意见》所产生的数据安全治理成效进行细致分析;第六部分为进一步分析,对《规范教育APP意见》如何影响数字经济发展的市场秩序、市场公平和市场培育进行阐述;第七部分为结论与政策启示。

二、文献综述

企业作为数据的关键处理者和数据流通的重要参与者,其数据安全保护能力的构建直接关系到国家的安全和发展(朱雪忠、代志在,2020)。从安全视角来看,若企业缺少安全风险防范能力建设,则易遭遇网络系统攻击、干扰和破坏,网络数据的完整性、保密性和可用性难以保证(刘金瑞,2021),极大影响社会安全。同时,若企业存在数据管理失范,则会造成用户隐私泄露,增加用户的心理困扰和安全风险,还可能导致与用户有类似特征群体隐私特征被推测(阿奎斯蒂等,2016),使更广泛的社会群体面临安全威胁。从发展视角来看,企业的安全风险防范不足和数据管理不规范,会造成用户和企业的信任危机,这种危机不仅会使得企业面临信誉受损和法律责任(克什特里,2014),还会产生负向溢出效应,影响到用户对同类企业的信任程度(马丁等,2017),进而影响数字经济业态繁荣与高质量发展。

企业数据安全保护能力建设事关重大,但鉴于企业通常倾向于短期逐利,仅依靠企业自律难以达到符合社会共同利益的建设水平。一是公地悲剧导致企业缺乏增加数据安全投资与部署的动力。由于网络外部性,数据安全部署的收益与用户数量密切相关,安全防护成效亦取决于行业合作伙伴的防护投入,当市场没有足够多企业进行数据安全投入时,则会出现投资效率低下和成本大于收益的困境(安德森、摩尔,2006)。二是机会主义导致企业具有主观违规使用数据的动机。滥用数据、过度采集数据等有助于企业扩大市场优势,迎合了企业短期盈利目标(戈德法布、阙,2023)。三是激励不相容导致企业缺乏数据安全管理与防范的自驱力。用户相对于企业处于信息不对称的弱势地位,难以核实自身数据是否被不当使用(科克

斯,2017),企业因此缺乏加强数据安全保护和预防措施的紧迫性。此外,即便企业发生了危害数据安全的相关事件,亦有经验采取较为系统的策略规避对自身的负面影响,例如,企业可根据新闻媒体的繁忙程度调整事件公告发布的战略时机,以削弱公众关注和因此引发的信誉危机(福德勒、舒茨,2022)。由此,仅依靠企业自律难以提升其数据安全风险防范水平和数据规范管理,增强政府数据治理力度是规避企业引发数据安全隐患的重要机制。

尽管有国际研究对于数据安全治理持保留态度(科尔尼希等,2021),但更多研究对数据治理可产生的安全效益持乐观态度,总结出了提升数据安全治理效能的政策路径。在防范数据安全隐患上,一是强化主体责任和建立权责清晰的奖惩制度,推动企业落实数据保护细则。为落实企业数据管理职责,可根据行业特点和数据特征细化数据安全监管细则,推动企业建立数据安全保护体系,设计数据保护目标和风险评估流程,缓解数据安全外部性难题(戈登等,2015)。通过激励相容监管机制设计,助力企业将数据安全作为自身竞争优势之一,从而加强数据安全治理力度,实现新技术在数据管理领域的推广(阿杰里德等,2016;戈登等,2018)。二是加强数据安全隐患通报,倒逼企业强化数据安全保障投入。详尽的数据安全披露指南和规范的数据安全事件通报流程,可以增加社会关注与引发社会舆情,企业为避免声誉制裁,会主动防范数据安全漏洞和增强数据安全投入(罗曼诺斯基等,2011;李等,2018;唐、温斯顿,2020)。在强化数据规范治理上,一是建立相对统一的数据保护法律框架和操作规则,减少企业利用政策不一致性而钻漏洞的行为。设计清晰且明确的数据安全政策、标准和方针,确保数据治理的实践程序与数据监管规则相吻合(亚伯拉罕等,2019)。搭建相对一致的数据监管框架,探索一致性机制、一站式服务体系与联合监管等,以保障数据安全领域的合作与协调(巴纳德-威尔斯等,2016)。二是加强精准治理布局,主动适应数据市场发展态势,建立与新技术、新模式和新产业相匹配的治理框架。数据治理变革不应仅局限于渐进性改革,而是应创造出更具法律前瞻性和创新性的治理框架,以协调日益变化的数据安全与数据流通需求(布里、谢尔,2016)。数据治理布局还应在时效性和持续性之间平衡,以及时甄别并应对市场变化和新技术涌现所产生的新型监管挑战(宾斯、别蒂,2020)。

虽然上述研究较多表达了对数据治理提升数字安全水平的强烈期待,但国际学者亦谨慎指出数据治理对于数字经济发展的潜在负面影响。一是增加市场秩序维护成本。以规则复杂性和限制性增强为特征的“过度治理”可能会限制以数据为主导的创新,诱使用户绕过监管与企业达成私下共识,承担不必要的数据安全风险(亚伯拉罕等,2019);这会导致遵守数据治理规则的企业承担更多的合规成本,而无法享受到市场扩张或用户积累的额外收益。二是阻碍市场公平。数据治理可能导致反竞争性,跨业经营、服务范围广泛的大型企业可以快速适应流程变化,交易成本被不成比例地施加给了专注于利基市场的小企业和新进入者上(坎贝尔等,2015),小企业和新兴企业的投资环境和生存环境也会进一步恶化(贾等,2021)。三是市场活跃度受到威胁。数据安全规则趋严增加了企业采集、存储、处置与转售数据的复杂性,影响了数字经济领域投资,导致了相关初创企业的估值下滑(卓等,2021)。以欧盟的通用数据保护条例(简称GDPR)政策为例,在GDPR生效后,谷歌应用市场的APP进入率下降了近一半(詹森等,2022)。此外,大企业不仅可以通过数据合规建立比较优势,还可以以保护用户数据为借口建立“围墙花园”进行数据垄断,在相对封闭的平台采集和利用用户数据,并拒绝将数据共享给小型企业和新进入者(布雷尔等,2020),或进一步影响中小企业的市场活跃度。综上,分析数据治理的有效性不仅应从安全视角进行评估,还应充分考虑其对数字经济发展的影响。探索政策层面如何平衡发展与安全正成为国际上的研究热点,而国内相关定量研究仍相对缺失,因此将成为本文重点研究方向。

三、政策背景

如何兼顾数字经济发展和数据安全保障,已成为世界各国面临的共同治理难题。中国作为全球最大的数字经济体之一,同样面临严峻的数据安全挑战,早期APP违规收集使用个人信息、泄露隐私数据等问题较为突

出。随着近年来我国数据治理体系的不断完善和数据安全监管力度的持续加强,数据安全保障水平得以有效提升。深入总结中国数据治理的实践经验,可以为全球数据治理提供经验参考。

《规范教育APP意见》于2019年8月实施,是首个全国层面全面规范教育APP的政策文件,该政策进行了中国特色治理路径创新探索,其治理策略也在中国后续数据治理中得以应用和进一步完善。该政策将“正确处理政府与市场、管理与服务、安全与发展的关系”作为基本原则,针对企业数据安全风险防范不足和数据管理不规范问题,提出了相应规范与约束。在防范安全风险上,《规范教育APP意见》落实网络安全主体责任和统一的网络安全等级保护要求,以强化网络攻击防范和保障系统平稳安全运行;推动网络安全审核、认证与检测机制建设,全面提高网络安全保障水平。在规范数据管理上,《规范教育APP意见》提出建立收集、储存、传输、使用等覆盖数据流通全流程的数据保障机制;严格禁止变相强迫用户授权、收集超服务范围个人信息、与用户违法约定以及泄露用户隐私等行为。

与国际文献相对照,《规范教育APP意见》的实施举措与国际倡议的治理理念相契合。一是构建一体化治理框架。在全国层面推行一致性的规范办法,避免监管标准不统一造成地区间监管套利。二是完善奖惩机制。健全教育APP管理责任体系,搭建教育APP选用退出机制、负面清单机制和黑名单机制。搭建全国信用信息共享平台整合惩戒信息,确保治理信息及时互通共享,增加联合惩戒成效。三是强化风险预警。建立常态化的监测预警通报机制,进一步加强安全威胁整改工作。建立信用监管机制,根据教育APP的投诉举报和隐患修复相关情况进行信用记录,再依据信用记录的优良减少或增加检查次数。四是推动精准治理。在广泛调研基础上划出风险防范和管理规范等方面的红线,根据行业动态及时展开技术和安全层面抽查,以提升监管的精准度。

同时,《规范教育APP意见》围绕推动有效市场和有为政府的更好结合落实举措,在统筹发展与安全上进行了以往国际研究中较少被提及的开拓性探索。一是以备案制替代审批许可制,营造鼓励创新的市场氛围。取消设置准入许可,通过备案及时摸底教育APP供需双方情况,为事中事后监管提供了有力支撑。从事前的审查批准转向事中的过程监督和事后的结果审核,是对数字经济新业态采取包容审慎的监管模式,对于营造公平竞争的治理环境和加快向服务型治理转型具有积极意义。同时,教育部对于备案主体、备案时间、备案内容和备案流程进行严格把控,依托国家教育资源公共服务平台为各省工作开展提供信息化支撑和数据共享服务,各省则负责执行备案信息登记具体事项,依据全国统一标准组织本地区的教育APP提供者进行备案。“一省备案,全国有效”的模式,不仅极大简化了企业备案流程,降低了企业进入新业态的门槛,还有助于全国统一大市场建设,促进了技术和数据跨区域流通共享。二是建立多部门协同联动,破除条块分割困局。《规范教育APP意见》要求,教育行政部门与网信部门、电信主管部门、公安部门进行协同联动和联合治理,开展网络安全等级保护定级备案和等级测评,规定教育APP提供者在发布和上线新应用时,需提供和更新网络安全等级保护定级备案的证明、等级测评报告等备案材料,推动教育APP开发商增强安全技术防护措施,遵守安全管理和履行安全义务。这一协同治理的模式可充分发挥各部门专业优势,确保无盲区、无死角的高效治理。同时,“谁主管谁负责、谁开发谁负责、谁选用谁负责”的全国统一管理下,以省为单位探索优质教育APP的推荐机制、选用退出机制、负面清单和黑名单制度,则可以保障央地协同,充分调动地方积极性,将地方可复制经验快速在全国范围推广。三是强化共建共治共享,确保治理能效发挥和管治水平提升。建设政府、市场、社会参与的多主体参与治理体系,积极引入专业机构、社会组织、学生家长等社会力量进行监督,有助于增强公众参与治理的“主人翁”意识。搭建信用评价体系和推行服务评议制度,可以及时获得市场反馈,减少对企业经营功能的干涉,以行业自律为抓手促进数字经济的高质量发展。拓展监督渠道,推动意见反馈和投诉举报的扁平化,可以及时回应公众关切,化解社会矛盾,提高治理的公信力。

将《规范教育APP意见》作为本文的研究对象具有较强的学术价值和政策意义。一方面,该政策强调问题导向,避免“一刀切”,与国际倡议的多项举措相契合,且该政策坚持有效市场和有为政府相结合的中国特色治

理路径,使得全球治理理念得以更好发挥。研究这一施策对象明确的政策,不仅可以较为准确地识别治理举措对于提升数据安全的作用,还可以较好地分享中国数据治理的实践经验。另一方面,该政策以备案制度降低准入门槛、以协同联动打破条块分割、以共建共治共享提升管治水平的治理路径是中国特色数据治理理念的缩影,也是践行高效能治理的有益探索,研究该政策的路径创新,可以为以高效能治理促进高质量发展和高水平安全良性互动提供对策建议。

四、数据说明与计量模型

为全面分析如何在数据治理中统筹发展与安全,本文应用APP产品安全检测大数据进行实证分析,识别《规范教育APP意见》推行的政策效应。

(一)数据介绍

本文主要使用的数据是APP产品安全检测大数据,该数据来自国内一家大型网络安全企业(下文称为“A企业”)。A企业通过对国内全部安卓类应用商店上架的APP进行高专业度的安全检测,建立起了APP产品安全检测数据库。A企业进行APP安全检测的具体步骤如下。首先,对国内全部共计1000余家的安卓类应用商店列表进行数据抓取,对各个安卓类应用商店的APP上架、下架、版本更新等情况进行数据采集。其次,在采集安卓应用商店中各个版本APP的安装包数据后,A企业利用智能化筛查与人工辅检相结合的方式对APP进行安全检测,检测重点包含数据安全漏洞和数据违规行为两个方面。其中,数据安全漏洞主要指APP开发商在代码逻辑设计、编写等环节上产生的错误和缺陷,这类安全漏洞会导致黑客利用漏洞植入恶意代码或手机病毒、攻击应用程序、造成用户隐私数据泄露等。A企业参考《国家信息安全漏洞库(CNNVD)》技术标准,划分了数据库注入漏洞、日志数据泄露风险、截屏攻击风险等122种安全漏洞指标,并依据此对APP进行检测和打分。数据违规行为主要指APP开发商未经用户同意违法违规收集使用个人信息行为。A企业依据《APP违法违规收集使用个人信息行为认定方法》,可以将APP数据违规行为划分为六大类,即违规收集个人信息,超范围收集个人信息,违规使用个人信息,强制用户使用定向推送功能,APP强制、频繁、过度索取权限以及APP频繁自启动和关联启动等。进一步地,在上述六大类的基础上进一步细分出30小类违规行为。需要说明的是,由于隐私违规检测技术复杂且成本较高,A公司仅对全样本APP进行安全漏洞的全覆盖检测,随后选取全样本中约20%的APP进行违规行为的随机抽样检测。

本文获得了2019年1月~2020年12月期间A企业进行数据安全检测的全样本,检测范围为安卓类应用商店总下载量排名前5%的APP,共涉及超过10万个APP时间跨度24个月的观测值。使用该样本研究数据治理问题具有较高的代表性,原因包括以下几个方面。第一,安卓类应用市场的市场集中度较高,排名前5%的APP下载量占全部APP总下载量的比重超过80%。第二,根据应用商店常见的推送规则,下载量排名低于前5%的APP往往不会被应用商店推荐,因此排名前5%的APP是大众用户在浏览应用商店时所接触到的主要APP。第三,本文选取研究样本包含教育、社交通讯、游戏、网上购物、金融理财等17个行业类别的APP^①,已经近乎完全覆盖生活场景下用户常用的所有APP产品品类。并且,由于手机内存限制和使用习惯的考量,用户通常安装手机APP的数量不超过100个^②,使用总下载量前5%的检测样本已经可以较为全面和准确地反映大众用户日常接触到的APP应用市场生态。需要注意的是,本文数据仅涵盖安卓类应用商店的APP,未包含iOS等其他操作系统的商店数据,但考虑到我国安卓系统智能终端的市场份额占比近八成,本文所使用的安卓APP数据在反映我国移动应用市场整体生态方面仍具有很强的代表性。同时,本文谨慎地将结论限定在研究样本范围内,避免研究结论的过度推广。

应用APP产品安全检测大数据分析中国数据安全政策的相关效应有如下优势。一方面,保障实证分析的可靠性。本文数据广泛覆盖了安卓市场常用APP,且类别划分明确,可以较为清晰地识别《规范教育APP意见》对于教育APP和非教育类APP所产生的不同影响。相较于使用调查问卷和准实验开展的小样本研究,本文使用的样本量达百万级,可以更为精准地、更具代表性地进行数据治理的政策分析,避免了调查实验中易出

现的回答偏误(本多夫、诺曼,2018)。另一方面,确保了研究分析的代表性和全面性。本文不仅获得APP层面安全漏洞数和违规行为数,用以“对症下药”分析数据安全风险防范和数据规范管理的安全治理议题,还获取了APP下载量、版本信息、行业类别、开发商名称、开发商所在地区等数据,以更全面掌握APP应用市场发展态势。此外,本研究额外获取了样本APP在2016年1月~2018年12月期间出现与否等相关情况,识别了APP市场进入情况,并将工商注册数据中的企业名称与APP开发商名称进行匹配,以分析不同规模APP开发商所面临的影响异质性,从而更为细致全面分析统筹发展与安全的数据治理路径。

(二)指标构造与数据描述

首先,本文根据APP的行业类别将APP划分为教育类和非教育类两类,其中教育APP被判定为受到政策影响的处理组,其他非教育类APP则作为对照组。随后,本文使用APP产品安全检测大数据,按照每个APP版本更新的具体日期,将数据转化成一个月度层面的面板数据。A企业采用“版本触发检测”方案以确保安全检测及时高效,即假设APP在同一个版本内具有相同的安全特征,仅在APP进行了版本更新后对此APP进行新一轮安全检测。本文遵照上述假设,设定APP在未进行版本更新的月份中维持了与之前月份相同的安全漏洞数和违规行为数,计算了各APP在不同月度的安全漏洞数和违规行为数,由此获得了全部的APP—月份维度的面板数据,共计约233万个观测值。

表1展示了描述性统计的结果。APP安全漏洞数和违规行为数的平均值分别为15.1个和5.8个,仅0.95%的APP观测不存在安全漏洞和仅19.54%的APP观测不存在违规行为,APP应用市场的数据安全风险防范和数据规范管理仍有待增强。APP的平均总下载量超过2650万次,APP开发商的平均注册资本为1117万元^③,鉴于平台经济具有“赢者通吃”的特征,而本文所使用的观测样本均为安卓类应用商店中的主流APP,其开发商的注册资本规模相对较大,与平台经济发展特征相符,侧面印证了本文使用数据的代表性和可信度相对较高。

在对本文所使用的数据进行了整体描述后,本文进一步对比教育APP和非教育类APP在数据安全上的表现。图2a和图2b分别展示了2019年1月~2020年12月期间,教育APP和非教育类APP的安全漏洞数和违规行为数的时间趋势图。竖线表示政策发生的时间点,即2019年8月《规范教育APP意见》出台。根据图2a,教育APP从整体来看在数据安全风险防范上相较于非教育类APP具有更佳表现,即便在《规范教育APP意见》发布前,教育APP的安全漏洞数也低于非教育类APP。这或许与中国国情相关,教育被视为国之根本,教育APP通常需要接受更为严格的政府监督和社会关注,在安全风险防范这一先行行为上具有更多约束。根据图2b可以发现,在《规范教育APP意见》发布前,教育APP和非教育类APP的违规行为数均相对较高,这表明APP开发商为达到经营目的而主观违背数据管理规范的行为普遍存在,强化数据治理具有较高紧迫性。从时间变化来看,教育APP和非教育类APP的安全漏洞数和违规行为数均持续减少,这或由于近年来数据安全愈发受到社会各界重视,APP开发商意识到数据安全与企业的

表1 描述性统计

变量	观测值	平均值	标准差	最小值	最大值
版块A: APP—月度层面					
安全漏洞数(个)	2325668	15.087	8.212	0	36
违规行为数(个)	381201	5.771	4.038	0	20
版块B: APP层面					
是否是教育APP (是=1;否=0)	139108	0.021	0.143	0	1
总下载量(百万次)	139108	26.504	631.005	0.010	69015.080
版块C: 开发商层面					
开发商注册资本 (万元)	45641	1117.305	2045.971	10	12800

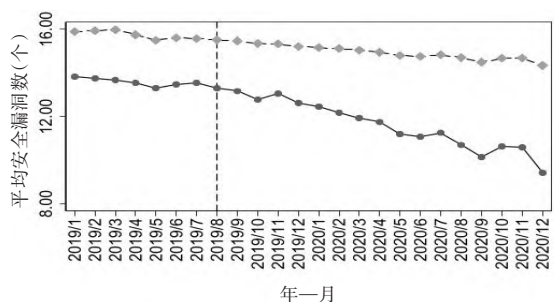


图2a 平均安全漏洞数

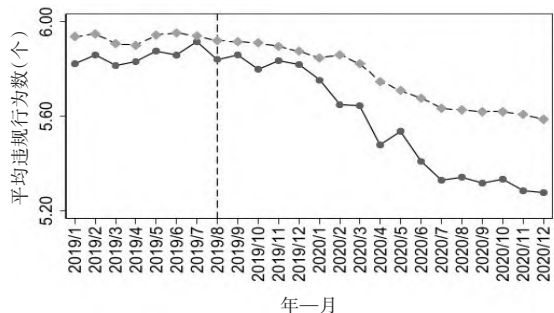


图2b 平均违规行为数

图2 教育APP和非教育类APP安全性的月度趋势

重大选题征文

可持续发展密切相关,因此加强了数据安全治理,数据安全隐患和侵犯用户隐私的行为均有所减少,APP总体安全性不断提高。进一步分时间阶段来看,在《规范教育APP意见》推行之前,两类APP在安全漏洞数和违规行为数上保持相对平行的变化趋势,但在该政策推行后,教育APP的安全漏洞数和违规行为数具有明显的下降趋势,这与本文的预期假设相一致,即《规范教育APP意见》对教育APP的数据安全提升具有积极的推动作用。

(三)识别策略

本文在标准双重差分模型基础上,进一步控制了APP层面和时间层面的双重固定效应,从而排除了每个APP特有的不随时间变化的特征,以及全部APP共同面临的时间趋势对结果变量的影响。本文建立的双向固定效应模型如下:

$$Y_{ict} = \alpha + \beta Treat_i \times Post_t + \delta_i + \gamma_t + \theta_c \times \gamma_t + \varepsilon_{ict} \quad (1)$$

其中, Y_{ict} 表示APP数据安全性的代理指标,包括安全漏洞数和违规行为数; i, t, c 分别表示APP序号、APP安全检测具体年月时间和APP开发商所处城市, Y_{ict} 即位于城市 c 、序号 i 的APP在时间 t 的安全漏洞数或违规行为数。 $Treat_i$ 代表是否处理组,当该APP属于教育APP时,则判定为处理组,取值1,该APP属于非教育类APP,取值0; $Post_t$ 表示观测时点是否为《规范教育APP意见》推行后,取值1表示政策推行后;取值0表示政策推行前; δ_i 为APP的个体固定效应,控制个体固定效应可以排除APP固有特征对APP安全性造成的影响, γ_t 表示时间固定效应,即控制年份与月份的联合固定效应,从而排除随时间变化的宏观经济环境对因变量所产生的影响干扰。考虑到不同APP所在城市的经济发展和政策落实存在差异,为了避免城市维度随时间变化的潜在遗漏变量对于估计结果的干扰,本文在模型中引入了 $\theta_c \times \gamma_t$, 即控制了城市和时间的联合固定效应。回归中使用在APP层面进行聚类的稳健标准误。系数 β 是本文最核心的估计量,反映了《规范教育APP意见》影响APP数据安全表现的因果效应。

双重差分模型建立在平行趋势假设的前提下,即处理组与对照组的目标变量在政策推行之前应保持平行趋势。为了检验平行趋势假设是否成立,本文利用事件研究模型进行估计,即检验处理组的个体在未受到政策干预的情况下,是否与对照组保持相同的变化趋势。本文构造的事件研究模型如下:

$$Y_{ict} = \alpha + \sum_{k=-7}^{17} \beta_k Treat_i \times Period_k + \delta_i + \gamma_t + \theta_c \times \gamma_t + \varepsilon_{ict} \quad (2)$$

为使平行趋势假设检验覆盖全样本集、全观测期,本文把事件研究模型的时间窗定义为政策推行前7个月到政策推行后17个月。其中, $Period_k$ 是一系列观测期的虚拟变量, k 用来表示观测期与政策推行时间点之间的时间距离。为了防止多重共线性,本文把政策推行前一期作为基准期,即不在事件研究模型中加入 $Period_{-1}$ 相关虚拟变量。本文关注 β_k 系列系数, β_k 反映了不同观测期相对于基准期的相对处理效应。当 $k < 0$ 时, β_k 不拒绝零假设,对照组和处理组具有相同变化趋势,符合平行趋势假设。模型(2)的其他变量定义与模型(1)一致。

五、实证分析

《规范教育APP意见》施策方案不仅与国际倡议理念相契合,还进行了中国特色的创新探索。下文将检验《规范教育APP意见》所取得的实际安全监管效果,为分享中国数据安全治理经验提供实证依据。

(一)基准结果

本节应用模型(1)进行基准回归,结果见表2。其中,第(1)~(2)列因变量为安全漏洞数,第(3)~(4)列因变量为违规行为数。第(1)列和第(3)列控制了APP固定效应和城市固定效应,第(2)列和第(4)列则进一步控制了城市和时间的联合固定效应。根据结果可知,第(1)列中是否处理组与政策是否推行交叉项的系数为负且显

表2 基准回归结果

	(1)	(2)	(3)	(4)
	安全漏洞数		违规行为数	
$Treat_i \times Post_t$	-1.365*** (0.106)	-1.326*** (0.107)	-0.130** (0.052)	-0.131** (0.053)
APP固定效应	是	是	是	是
时间固定效应	是	是	是	是
时间×城市固定效应	否	是	否	是
观测值	2325668	2325668	381201	381201
R ²	0.834	0.836	0.938	0.939

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

著为负,第(2)列该系数仍然显著为负,并且与第(1)列对应系数相差较小,说明结果相对稳健。平均而言,在其他条件不变时,《规范教育APP意见》的推行使得APP的安全漏洞数减少了1.3~1.4个,相当于整体样本均值的9%。第(3)列中是否处理组与政策是否推行交叉项的系数显著为负,第(4)列该项系数依然显著,且与第(3)列结果相差较小。平均而言,在其他条件不变时,《规范教育APP意见》的推行使得APP的违规行为数平均减少了0.13个,相当于整体样本均值的2.3%。综上,《规范教育APP意见》推行在数据安全风险防范和数据行为规范两个方面均取得了显著成效,对数据安全提升具有积极意义。

(二)稳健性检验

上文已初步研究发现《规范教育APP意见》对于数据安全提升的显著作用,下文将进一步通过平行趋势检验、倾向得分匹配法、考虑APP版本更新、排除同期政策干扰、考虑政策溢出效应等方式,验证研究结果的稳健性^④。

1. 平行趋势检验

本文应用事件研究模型(2)进行平行趋势检验,验证在《规范教育APP意见》推行前处理组是否与对照组保持平行的变化趋势,满足双重差分模型的应用前提。图3为事件研究的结果。其中,横坐标为具体年月,政策推行的前一期被定义为基准期,在图中均用垂直虚线标注,图3a和图3b的纵坐标分别表示处理组与对照组在安全漏洞数和违规行为数上的差值相对于基准期差值的变化。

由图3可以看出,相较于《规范教育APP意见》推行前一期,政策推行前的其他观测期,其对照组和处理组在安全漏洞数和违规行为数上的差异不存在显著变化,且置信区间相对较小,结果相对稳健。即《规范教育APP意见》推行前,对照组和处理组具有相同变化趋势,上述结果说明处理组和对照组在事前具有平行的变化趋势,证明平行趋势假设成立。同时,相较于《规范教育APP意见》推行前一期,政策推行后的观测期,处理组与对照组在安全漏洞数和违规行为数上的相对差距显著增加,具体表现为处理组相较于对照组在安全漏洞数和违规行为数上的下降趋势更为显著,这表明该政策推行对处理组的数据安全提升具有积极的正向作用。且从时间趋势来看,随着时间推移,处理组在安全漏洞数和违规行为数上的相对下降幅度得以保持,甚至下降幅度有所增加,这表明该政策不仅在短期助力数据安全保护,还可以在较长时期维持政策效果,产生可持续的数据安全风险防范和数据规范管理效能。值得注意的是,虽然两项指标均通过了平行趋势检验,但违规行为数在政策前存在一定波动,置信区间相对较宽。基于结果的稳健性考虑,本文以安全漏洞数的分析结果为主要结论,将违规行为数的结果作为补充证据。因此本文对违规行为数相关研究结果持更为审慎的态度,将之视为边缘支持证据,而将安全漏洞数相关结论视为主要结论。

2. 倾向得分匹配法

考虑到作为处理组的教育APP与对照组APP在样本量上存在较大差异,为提高研究结论的可靠性,本节采用倾向得分匹配(PSM)方法构建特征更为相似的对照组,以缓解潜在的样本选择偏误问题。在匹配过程中,本文选取了APP在政策推行前的关键特征作为匹配变量,包括下载量、更新频率、索取权限数量、安装包大小以及APP年龄等。这些特征可以比较全面地反映APP的市场影响力、技术复杂度和发展成熟度,是决定APP安全性能和监管响应能力的重要因素。基于这些关键特征,本文采用1:1的最邻近匹配方法,为每个教育APP匹配一个特征最为相似的非教育类APP,从而构建了样本分布更为均衡

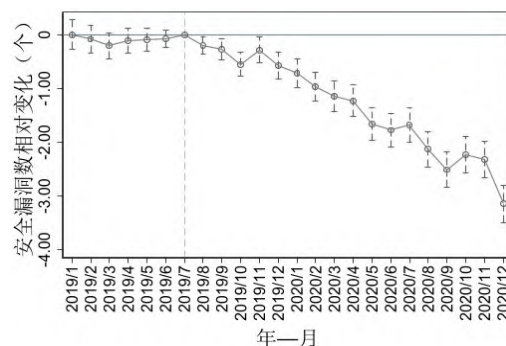


图3a 安全漏洞数事件研究

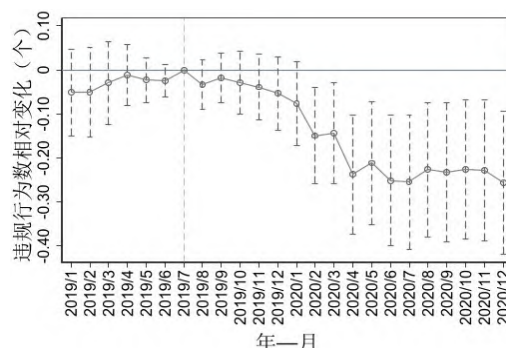


图3b 违规行为数事件研究

图3 平行趋势检验

的子样本数据集^⑤。

基于PSM匹配后的子样本,本文重新估计了双重差分模型,相关结果见表3。研究结果表明,在控制APP固定效应、时间固定效应以及城市—时间联合固定效应后,教育APP相较于匹配后的对照组,安全漏洞数显著减少了1.2个,违规行为数显著降低了0.19个。这些结果与基准回归结果高度一致,表明《规范教育APP意见》提升教育APP数据安全的积极影响具有较强的稳健性,不受样本构成变化的显著影响。

3. 考虑APP版本更新

基准回归中假设APP在未进行版本更新的月份中维持与之前月份相同的安全漏洞数和违规行为数。为检验这一假设对于本文结果稳健性的影响,本文以APP版本为观测单位的面板数据,将APP未进行版本更新的月份设为观测值缺失,即考虑每个版本首次发布时的安全检测结果,完全避免了对未更新月份的假设处理。表4第(1)~(2)列分别汇报了新构造数据下安全漏洞数和违规行为数的回归结果。结果显示,估计系数的大小和显著性与基准结果相比没有明显变化,说明本文的主要结果是稳健的。此外,本文还在基准模型中控制了版本更新与否的哑变量*Update_i*,从而控制版本更新对APP安全性的可能影响,见表4的第(3)~(4)列。相关结果显示,控制版本更新后的回归系数与基准结果高度一致,进一步验证了本文结论的可靠性。

4. 排除同期政策干扰

在本文研究的窗口期,除《规范教育APP意见》推行外,还存在着其他监管政策,为此下文将进一步排除同期其他政策的干扰影响。其中,《儿童个人信息网络保护规定》(简称《儿童保护规定》)于2019年8月颁发,自2019年10月起实施,与《规范教育APP意见》推行时间相接近。不过两者施策针对性有所不同,《儿童保护规定》重点在于保护不满十四周岁的未成年人(简称儿童)的个人信息,《规范教育APP意见》则使保障教育APP健康有序发展,保护的群体既包括儿童,也包括其他使用教育APP的非儿童用户。考虑到《儿童保护规定》的保护对象为儿童,对APP监管也主要集中于儿童使用的相关APP,为排除《儿童保护规定》所产生的政策干扰,本文根据APP名称和简介,判别出受众为儿童用户的APP,将教育APP进一步划分为面向儿童群体的APP和面向非儿童群体的APP,并将面向儿童群体的APP从样本中剔除,重新应用模型(2)进行回归分析。表5的第(1)列和第(2)列分别展示了去掉面向儿童群体的APP后,《规范教育APP意见》推行对于APP安全漏洞数和违规行为数的影响。结果显示,该影响系数依然显著为负,且回归系数变化与表2基准结果相差较小,这表明主要结果依然是稳健的,《儿童保护规定》实施对于本文主要结果干扰作用较小。

在对教育APP监管加强的同期,游戏类APP、金融类APP的治理环境也发生了变化。一方面,伴随

在对教育APP监管加强的同期,游戏类APP、金融类APP的治理环境也发生了变化。一方面,伴随

表3 倾向得分匹配法

	(1) 安全漏洞数	(2) 违规行为数
<i>Treat</i> × <i>Post</i> _{<i>i</i>}	-1.185*** (0.250)	-0.190** (0.075)
APP固定效应	是	是
时间固定效应	是	是
时间×城市固定效应	是	是
观测值	44705	33485
R ²	0.681	0.919

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

表4 考虑APP版本更新

	(1)	(2)	(3)	(4)
	使用“版本—月份”维度数据 安全漏洞数	违规行为数	控制APP版本更新 安全漏洞数	违规行为数
<i>Treat</i> × <i>Post</i> _{<i>i</i>}	-1.459*** (0.139)	-0.148* (0.086)	-1.301*** (0.107)	-0.130** (0.053)
<i>Update_i</i>	—	—	-0.149*** (0.006)	-0.041*** (0.004)
APP固定效应	是	是	是	是
时间固定效应	是	是	是	是
时间×城市固定效应	是	是	是	是
观测值	532605	92145	2325668	381201
R ²	0.763	0.919	0.836	0.939

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

表5 排除其他同期政策的干扰

	(1)	(2)	(3)	(4)	(5)	(6)
	排除《规定》干扰 安全漏洞数	违规行为数	排除游戏类APP监管影响 安全漏洞数	违规行为数	排除金融类APP监管影响 安全漏洞数	违规行为数
<i>Treat</i> × <i>Post</i> _{<i>i</i>}	-1.290*** (0.109)	-0.139** (0.055)	-1.269*** (0.109)	-0.107** (0.053)	-1.287*** (0.107)	-0.128** (0.053)
APP固定效应	是	是	是	是	是	是
时间固定效应	是	是	是	是	是	是
时间×城市固定效应	是	是	是	是	是	是
观测值	2321203	375857	1761222	329416	2168188	357270
R ²	0.836	0.939	0.818	0.934	0.834	0.939

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

着网络游戏内容审核与监管的增强,游戏类APP治理备受重视,2018年3月,原国家新闻出版广电总局发布《游戏申报审批重要事项通知》,称由于机构改革,所有游戏版号的发放全面暂停,2018年12月底,原新闻出版广电总局公布审核恢复后的首批游戏版号。2019年10月,《国家新闻出版署关于防止未成年人沉迷网络游戏的通知》以强化实名制验证方式加强对游戏类APP治理。另一方面,伴随着数字金融监管加强,金融类APP的管理规范正逐步推广和细化。2019年9月,中央人民银行印发《移动金融客户端应用软件安全管理规范》,对移动金融客户端应用软件提出了提升安全防护能力、加强个人金融信息保护、提高风险监测能力、健全投诉处理机制、强化行业自律管理等五项规范管理方向。2020年2月,中国人民银行正式发布《个人金融信息保护技术规范》从安全技术和安全管理上对个人金融信息保护提出规范性要求。

本文考虑到,游戏类APP和金融类APP监管生态发生的变化,也可能会导致本文回归结果的解释偏误,为此进一步排除上述潜在因素对本文政策分析产生的影响。本文首先在主样本中剔除了游戏类APP相关样本,以排除游戏类APP治理态势变化所产生的政策干扰,重新应用模型(1)进行回归分析,表5的第(3)和第(4)列为剔除游戏类APP后的回归结果,《规范教育APP意见》影响依然显著为负且稳健。同样地,为了排除金融类APP的监管举措变化所产生的政策干扰,本文在回归样本中剔除掉了金融类APP相关样本,回归结果见表5的第(5)和第(6)列,相关系数依然稳健为负,且与表2的基准回归系数基本吻合。上述结果表明,《规范教育APP意见》推行对数据安全治理产生了积极作用,而非其他同期政策干扰造成的混淆偏差^⑥。

5. 考虑政策溢出效应

本文使用的双重差分模型的有效性建立在政策仅对处理组产生影响、对对照组无直接作用的假设基础上。然而,《规范教育APP意见》作为数据治理领域的重要政策,可能通过多种渠道对非教育类APP产生溢出效应,进而影响本文识别策略的有效性。为确保研究结论的可靠性,本文从两个维度检验可能的政策溢出渠道。

首先,从APP功能相似性角度考虑溢出效应。办公学习类APP与教育APP在功能定位和用户群体方面具有较高相似性,可能受到该政策的间接影响。为检验这一可能性,本文构建了办公学习类APP的虚拟变量 $StudyOffice_i$,并与政策虚拟变量进行交互,检验政策是否对办公学习类APP产生显著影响,回归结果见表6第(1)和第(2)列。研究结果显示,办公学习类APP交互项系数均不显著,表明《规范教育APP意见》对办公学习类APP的安全性未产生显著影响。同时,主变量系数依然显著为负,且与基准结果高度一致,进一步验证了本文识别策略的有效性。此外,本文在表6第(3)和第(4)列中直接将办公学习类APP从样本中剔除,重新估计政策效应。回归系数与基准结果高度相似,证明在排除功能相似APP溢出效应的潜在干扰后,本文主要结果仍然是稳健的。

其次,从开发商层面也有可能存在溢出效应,由于同一开发商可能同时开发教育APP和非教育类APP,当其教育APP受到政策约束时,可能会对该开发商开发的其他类别APP产生影响。为检验这一渠道,本文构建了开发商层面的虚拟变量 $SameDeveloper_i$,用以标识是否为开发过教育APP的开发商所开发的非教育类APP,

表6 排除潜在的溢出效应

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	控制办公学习类APP		排除办公学习类APP		控制相同开发商		排除相同开发商APP	
	安全漏洞数	违规行为数	安全漏洞数	违规行为数	安全漏洞数	违规行为数	安全漏洞数	违规行为数
$Treat_i \times Post_t$	-1.297*** (0.107)	-0.125** (0.053)	-1.306*** (0.107)	-0.128** (0.054)	-1.300*** (0.107)	-0.128** (0.053)	-1.298*** (0.107)	-0.127** (0.053)
$StudyOffice_i \times Post_t$	-0.002 (0.038)	0.020 (0.025)	—	—	—	—	—	—
$SameDeveloper_i \times Post_t$	—	—	—	—	-0.071 (0.057)	0.016 (0.045)	—	—
APP固定效应	是	是	是	是	是	是	是	是
时间固定效应	是	是	是	是	是	是	是	是
时间×城市固定效应	是	是	是	是	是	是	是	是
观测值	2325668	381201	1968900	310798	2325668	381201	2213012	357463
R ²	0.835	0.939	0.839	0.937	0.835	0.939	0.836	0.939

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

重大选题征文

并与政策虚拟变量进行交互。第(5)和第(6)列的结果显示,交互项系数同样不显著,表明政策未通过开发商渠道对非教育类APP产生显著溢出效应。第(7)和第(8)列进一步剔除了开发过教育APP的开发商所开发的其他非教育类APP,主要政策效应依然稳健显著。

综上所述,无论是从APP功能相似性还是开发商层面,《规范教育APP意见》均未对非教育类APP产生显著的溢出效应,本文的双重差分识别策略是有效的,研究结论具有较强的可信度。

六、进一步分析

如何在数据治理中平衡安全与发展是国际文献关注的热门议题。现有文献表明,数据治理在提升安全性的同时,往往伴随着扭曲市场秩序、强化反竞争效应、抑制市场活力等“副作用”(亚伯拉罕等,2019;布雷尔等,2020;詹森等,2022)。这主要源于信息不对称条件下数据治理引致的制度性交易成本差异,加剧了市场集中与创新抑制,使得数据治理难以在安全与效率之间实现有效平衡。与此不同,中国数据治理实践坚持系统观念和综合治理理念,为统筹发展与安全提供了新的路径选择,有望实现数字经济高质量发展与高水平数据安全的良性互动。基于此,本文进一步探究《规范教育APP意见》对数字经济发展的影响,以检验该政策在提升数据安全水平的同时,是以牺牲发展为代价,还是实现了以安全促发展的目标。

(一)数据治理与市场秩序

国际研究认为,数据治理在带来安全性提升的同时并不必然维护市场秩序,用户可能会选择与企业合谋,绕过监管以规避繁琐的合规流程,承担不必要的安全风险(亚伯拉罕等,2019),导致高风险APP安全性能难以改善,挤占低风险APP的市场空间,出现“劣币驱逐良币”现象。为探索《规范教育APP意见》对于市场秩序维护的有效性和可持续性,本文将判别各APP在政策推行前的风险水平,并对不同风险水平APP的安全治理成效进行对比。倘若《规范教育APP意见》推行后,APP的安全性提升与APP初始风险水平成反比,则意味着高风险APP与低风险APP的数据安全差距会进一步扩大,长期以往数据安全的两极分化现象会加剧,且高风险APP不必像低风险APP一样付出高昂的安全保障成本就可以在应用市场立足,这不仅会破坏数据治理的奖惩机制,降低APP开发商提升数据安全保障的能动性,也会导致用户难以甄别自身安全是否受到威胁,从而阻碍可信数据空间培育。反之,若该政策推行后,APP的安全性提升与APP初始风险水平成正比,则意味着APP应用市场的数据安全分布相对均衡,高风险APP与低风险APP的数据安全差距会持续缩小,各APP开发商均有动机增强数据安全保障以立足市场,从而助力减少用户甄别APP安全性的搜索成本,推动健康有序的可信数据空间培育。

为此,本文构造事前APP风险水平的衡量指标 $Risky_i$,将样本APP在政策发生前的最近一次安全检验中所测得的安全漏洞数和违规行为数,分别作为事前APP风险水平的衡量指标。在模型(1)的基础上,本文增加事前风险水平、是否处理组与政策是否推行的三重交叉项,回归结果见表7。其中,第(1)列因变量为安全漏洞数,事前风险水平根据事前安全漏洞数判定,第(2)列因变量为违规行为数,事前风险水平根据事前违规行为数判定。根据结果可知,第(1)和第(2)列的三重交互项系数均显著为负,在其他条件不变时,当《规范教育APP意见》推行后,初始风险水平越高的APP其安全漏洞改善效果和违规行为减少成效越显著。由此可知,《规范教育APP意见》所进行的数据治理过程,没有出现“劣币驱逐良币”的市场失序现象。这或得益于《规范教育APP意见》构建的条块协同治理机制,打破了部门间信息壁垒,压缩了监管套利空间,缓解了治理盲区导致的逆向选择问题。在激励相容的制度设计下,高风险APP面临更强的转型约束,而低风险APP的竞争优势得以显现,有效避免了“劣币驱逐良币”现象,推动了数字应用市场的优胜劣汰机制形成。

表7 数据治理与市场秩序

	(1)	(2)
	安全漏洞数	违规行为数
$Treat_i \times Post_t \times Risky_i$	-0.129*** (0.014)	-0.036** (0.018)
$Treat_i \times Post_t$	0.350* (0.196)	0.082 (0.093)
APP固定效应	是	是
时间固定效应	是	是
时间×城市固定效应	是	是
观测值	2190188	309131
R ²	0.838	0.939

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

(二)数据治理与市场公平

上文的研究发现,《规范教育 APP 意见》推行对于维护市场秩序上的积极作用,下文将进一步检验该项政策对于市场公平竞争的影响,从而更加综合地评价该政策在统筹发展与安全上的考量。保障中小微企业公平参与市场竞争是打破行业垄断和激发创新活力的必由之路。然而,国际研究表明,数据治理存在挤压中小微企业生存空间的隐患,这是由于发展初期的中小微企业提升安全性往往需要付出较多的机会成本,经营发展亦会随之受到影响,相比之下实力雄厚的大型企业则有更多实力应对监管要求(布雷尔等,2020)。由此,倘若对实力雄厚的 APP 和实力薄弱的 APP 应用相同的监管约束,或会增加中小微开发者的合规负担,影响应用市场的公平与效率,不利于数字经济高质量发展。

为了检验《规范教育 APP 意见》对于不同实力 APP 是否采用了相同监管的约束,本文首先根据 APP 开发商注册资本和 APP 下载量,从 APP 开发商资历和 APP 用户规模两个维度衡量 APP 的市场实力。同表 7 回归模型相类似,本文构造连续变量 $\ln_Size_Developer_i$ 和 $\ln_Download_App_i$ 衡量 APP 开发商资本实力和 APP 用户规模, $\ln_Size_Developer_i$ 为开发商注册资本对数值, $\ln_Download_App_i$ 为 APP 下载量对数值。随后,本文分别将上述两个衡量 APP 实力的变量与政策虚拟变量、时间虚拟变量进行三重差分,并引入基础模型中进行回归分析,回归结果见表 8。

表 8 的异质性分析显示,具有较强市场实力的 APP 会面临更严格监管约束,具体来看,《规范教育 APP 意见》推行后,当 APP 开发商注册资本越大、用户资源更丰富时,所取得安全风险防范和数据规范管理提升效果更明显^⑤。这表明《规范教育 APP 意见》的推行不仅提升了 APP 的安全性,也更好地规制了更具市场地位的企业和更受市场关注的产品。这或由于备案制度的实施增强了监管部门的信息获取能力,为差异化监管提供了基础。大型企业无法利用市场地位进行监管套利,小型企业制度性交易成本则相对下降,公平竞争的市场环境得以形成。

(三)数据治理与市场培育

已有研究认为,过度严格的数据监管会增加企业维持运营成本,导致相关市场的进入率下降(詹森等,2022)。特别是具有市场优势的大型企业往往会左右数据安全监管实施进程,掌握规则优势将中小微企业排除在市场之外(布雷尔等,2020),进一步阻碍了市场培育。而《规范教育 APP 意见》深化“放管服”改革背景下发布的或充分尊重市场规律,对于教育 APP 的市场进入产生较小负面冲击。并且,上文研究表明,更具市场地位的企业和更受市场关注的产品得以更好规制,则无法应用规则优势阻碍相对弱势的企业和产品发展,实力薄弱的 APP 在公平有序的市场环境下,得以持续涌入市场,应用市场整体活跃度有可能进一步增加。

为了检验上述猜测,本文应用 APP 进入率衡量应用市场活跃度。首先,本文根据 APP 首次出现时间判定市场进入时间。其次,本文构造市场进入相关虚拟变量,当观测期在进入时间之前时,则判定未进入市场,赋值 0,反之,则判定已进入市场,赋值 1。最后,本文根据 APP 行业类别对应用市场进行划分,用各类别 APP 各期的新进入数和存量总数之比,求得各类别 APP 应用市场各期的进入率。本文应用模型(1),将因变量替换为市场进入率,应用 APP 类别一月份层面数据进行回归,结果见表 9 第(1)列。进一步地,本文分析不同市场地位 APP 在进入率上的异质性。本文对所有 APP 开发商的注册资本进行排序,将开发商注册资本超过中位数的 APP 定义为大型开发商 APP,反之定义为小型开发商 APP。类似地,本文对政策发生前 APP 下载量在全部 APP 中进行排序,超过中位数的 APP 定义为大型 APP,反之定义为小型 APP。随后,本文根据 APP 开发商资历和 APP 用户规模区分 APP 市场实力,分别以市场进

表 8 数据治理与市场公平

	(1)	(2)	(3)	(4)
	安全漏洞数	违规行为数	安全漏洞数	违规行为数
$Treat_i \times Post_i \times \ln_Size_Developer_i$	-0.204*** (0.059)	-0.078** (0.036)	—	—
$Treat_i \times Post_i \times \ln_Download_App_i$	—	—	-0.175*** (0.055)	-0.120*** (0.029)
$Treat_i \times Post_i$	1.672* (0.899)	0.435 (0.560)	1.328* (0.796)	1.722*** (0.425)
APP 固定效应	是	是	是	是
时间固定效应	是	是	是	是
时间×城市固定效应	是	是	是	是
观测值	1563974	299057	2324041	376991
R ²	0.824	0.938	0.836	0.939

注:括号中的数值为在 APP 层面聚类的稳健标准误;***、**和*分别表示在 1%、5%和 10%水平下显著。

重大选题征文

入率为因变量进行分组回归,结果见表9第(2)~(5)列。

表9第(1)列的结果显示,《规范教育APP意见》推行后,教育APP市场活跃度显著上升,新进入市场的APP数量显著升高,教育APP应用市场进入率增加了30个百分点,这意味着应用市场规模的持续壮大,新业态、新模式持续涌现。且不同于国外所诟病的过度数据治理阻碍市场活跃,《规范教育APP意见》推行不仅减少了数据安全隐患,还进一步盘活了APP应用市场,激发了数字经济发展活力。表9第(2)~(5)列的结果显示,《规范教育APP意见》推行后,资本实力更强、用户资源更丰富的APP未出现更高的进入率。这表明,具有市场优势的企业并没有左右《规范教育APP意见》的推行,亦没有利用合规优势建立起更封闭和垄断的数字经济生态。同时,初始用户规模较小和资本实力相对薄弱的APP在政策推行后,进入率反而增加,避免了应用市场集中度的上升。这或由于该政策采用了共建共治共享的治理路径,明确共同责任和利益,更具市场地位的APP因声誉机制率先进行合规示范,实力相对薄弱的APP则因此减少了试错成本,从而降低了合规难度,增加了市场活跃度。

七、结论与政策启示

本文研究发现,《规范教育APP意见》有效降低了教育APP的安全漏洞数和违规行为数,对于提升企业的数据安全风险防范能力和数据规范管理水平具有积极意义。并且,《规范教育APP意见》的推行有助于维护市场秩序、保障市场公平和促进市场培育。在市场秩序方面,该政策倒逼了高风险APP加速安全转型,缩小与低风险APP的数据安全差距。在市场公平方面,该政策对于资本实力更强、用户资源更丰富的APP具有更强的监管约束,更好地规制了更具市场地位的企业和更受市场关注的产品。在市场培育方面,该政策助力资本实力相对薄弱和用户资源相对欠缺的APP进入市场,提升了应用市场的整体活跃度。

《规范教育APP意见》以备案制度降低准入门槛、以协同联动打破条块分割、以共建共治共享提升管治水平的治理路径符合中国治理的一贯理念,也使得国际倡导构建一体化治理框架、完善奖惩机制、强化风险预警和推动精准治理相关理念更好落地,为以高效能治理促进高质量发展和高水平安全良性互动提供了有益探索。基于对政策效应分析以及对我国数据治理实践的总结,本文获得政策启示如下。

第一,打通条块信息壁垒,以系统观念构建数据综合治理体系。构建全国一体化数据治理平台,加强跨部门、跨地区、跨行业的数据互联互通,打破数据孤岛,推进各部门协调配合。建立央地协同治理问题动态反馈机制,确保权责清晰和奖惩有据,及时理清数据安全保障全过程的重点环节和薄弱环节,甄别治理盲区和治理死角。建立数据治理责任认领机制和问责机制,保障监管规则清晰、治理过程透明和合规流程便捷,避免企业耗费大量精力应对繁琐的合规流程而错失发展机遇。第二,完善监测和督导机制,以服务型治理保障创新活力。进一步完善和规范备案制度,破除各类市场准入壁垒。建立数据治理风险预警和督察制度,做好安全风险和数据违规隐患的摸排工作,推动包容审慎的治理,以预防性警示减少企业处罚负担。推广负面清单和免罚清单,建立信用惩戒缓冲机制,完善减少企业准入限制创新容错机制,对于违规企业设置观察期,在可控范围内实行容错纠错。持续发布和更新数据合规指引,组织相关宣讲和细则解答,定期开展座谈会和问卷调查,畅通企业建议反馈渠道,加快数据治理模式从“被动响应”向“主动服务”转型,营造鼓励创新的市场生态。第三,打造社会治理共同体,建设共建共治共享的数据治理生态。搭建数据治理共同体,推进治理评价机制和服务评议机制建设,为公众反馈、评价治理能效和服务水平提供渠道,推动社会组织、专业机构、行业协会、企业、居民参与到政府数据治理中。强化数据治理通报机制,对数据违规行为及时进行声誉制裁,助力公众及时掌握数据治理动态和有效行使监督权。推广“以大带小”帮扶机制,大企业为中小企业提供搭建数据合规管理体系的咨询和培

表9 数据治理与市场培育

	(1)	(2)	(3)	(4)	(5)
	全样本	大开发商	小开发商	大型APP	小型APP
<i>Treat</i> × <i>Post</i>	0.296** (0.120)	0.150 (0.100)	0.497*** (0.133)	0.046 (0.110)	0.503*** (0.132)
类别固定效应	是	是	是	是	是
时间固定效应	是	是	是	是	是
观测值	408	408	408	408	408
R ²	0.927	0.854	0.908	0.915	0.891

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

训,缓解中小微企业数据合规负担。创新共建共治奖励机制,为精准发现问题线索的公众提供奖励,通过财税政策激励大企业开展帮扶工作。

需要说明的是,受限于数据来源,本文结论主要适用于安卓应用市场中下载量较高的APP群体。在数据可得情况下,未来将进一步拓展样本范围,更系统、全面地探究数据治理生态的影响机制,以增强所研究结论的外推性。此外,进一步深入挖掘政策效应的具体传导机制,以开展与国际数据治理实践的对比研究,也将有助于更好地总结和推广中国数据治理经验。从以高效能治理促进高质量发展和高水平安全良性互动视角出发,对我国数据治理体系进行更为全面和深入的分析,将是作者未来的拓展方向^⑧。

(作者单位:续继,中国社会科学院经济研究所;杨卿栩、侯嘉奕,香港大学商学院)

注释

①行业类别包含教育、主题壁纸、健康运动、办公学习、影音播放、拍摄美化、旅游出行、游戏、生活实用、社交通讯、系统工具、网上购物、亲子、资讯阅读、辅助软件、金融理财、其他应用共17大类。

②根据2019年8月30日由中国互联网络信息中心发布的《中国互联网络发展状况统计报告》,各年龄段网民人均手机安装APP数量在32~66个之间。

③本文对开发商注册资本进行了2%的缩尾处理。

④简化固定效应设定、控制不同APP类别的时间趋势和安慰剂检验分别见《管理世界》网络发行版附录一、附录二和附录三。

⑤PSM平衡性检验在《管理世界》网络发行版附录四中呈现。附表四还呈现了使用同一类别APP的数量做倒数加权的回归结果。

⑥本文进一步排除了疫情因素干扰,见《管理世界》网络发行版附录五。

⑦本文进一步对表7和表8的相关结果进行稳健性检验,见《管理世界》网络发行版附录六。

⑧中外文人名(机构名)对照:戈登(Gordon);阿杰里德(Adjerid);李(Li);唐(Tang);温斯顿(Whinston);亚伯拉罕(Abraham);宾斯(Binns);别蒂(Biitti);詹森(Janssen);坎贝尔(Campbell);布雷尔(Bleier);阿奎斯蒂(Acquisti);克什特里(Kshetri);马丁(Martin);安德森(Anderson);摩尔(Moore);戈德法布(Goldfarb);阙(Que);科克斯(Kox);福德勒(Foerderer);舒茨(Schuetz);科尔尼希(Kollnig);罗曼诺斯基(Romanosky);巴纳德-威尔斯(Barnard-Wills);布里(Burri);谢尔(Schär);贾(Jia);卓(Zhuo);本多夫(Benndorf);诺曼(Normann)。

参考文献

- (1)陈晓红、李扬扬、宋丽洁、汪阳洁:《数字经济理论体系与研究展望》,《管理世界》,2022年第2期。
- (2)陈雨露:《数字经济与实体经济融合发展的理论探索》,《经济研究》,2023年第9期。
- (3)范柏乃、盛中华:《数字风险治理:研究脉络、理论框架及未来展望》,《管理世界》,2024年第8期。
- (4)黄璜:《数字政府:政策、特征与概念》,《治理研究》,2020年第3期。
- (5)刘金瑞:《数据安全范式革新及其立法展开》,《环球法律评论》,2021年第1期。
- (6)朱雪忠、代志在:《总体国家安全观视域下〈数据安全法〉的价值与体系定位》,《电子政务》,2020年第8期。
- (7)Abraham, R., Schneider, J. and Vom Brocke, J., 2019, "Data Governance: A Conceptual Framework, Structured Review and Research Agenda", *International Journal of Information Management*, 49, pp.424~438.
- (8)Acquisti, A., Taylor, C. and Wagman, L., 2016, "The Economics of Privacy", *Journal of Economic Literature*, 54(2), pp.442~492.
- (9)Anderson, R. and Moore, T., 2006, "The Economics of Information Security", *Science*, 314, pp.610~613.
- (10)Adjerid, I., Acquisti, A., Telang, R., Padman, R. and Adler-Milstein, J., 2016, "The Impact of Privacy Regulation and Technology Incentives: The Case of Health Information Exchanges", *Management Science*, 62(4), pp.1042~1063.
- (11)Barnard-Wills, D., Chulvi, C. P. and De Hert, P., 2016, "Data Protection Authority Perspectives on the Impact of Data Protection Reform on Cooperation in the EU", *Computer Law & Security Review*, 32(4), pp.587~598.
- (12)Benndorf V. and Normann H. T., 2018, "The Willingness to Sell Personal Data", *Scandinavian Journal of Economics*, 120(4), pp.1260~1278.
- (13)Binns, R. and Biitti, E., 2020, "Dissolving Privacy, One Merger at a Time: Competition, Data and Third Party Tracking", *Computer Law & Security Review*, 36(4), No.105369.
- (14)Bleier, A., Goldfarb, A. and Tucker, C., 2020, "Consumer Privacy and the Future of Data-Based Innovation and Marketing", *International Journal of Research in Marketing*, 37(3), pp.466~480.
- (15)Burri, M. and Schär, R., 2016, "The Reform of the EU Data Protection Framework: Outlining Key Changes and Assessing Their Fitness For a Data-Driven Economy", *Journal of Information Policy*, 6, pp.479~511.
- (16)Campbell, J., Goldfarb, A. and Tucker, C., 2015, "Privacy Regulation and Market Structure", *Journal of Economics & Management Strategy*, 24(1), pp.47~73.
- (17)Foerderer, J. and Schuetz, S. W., 2022, "Data Breach Announcements and Stock Market Reactions: A Matter of Timing?", *Management Science*, 68(10), pp.7298~7322.
- (18)Goldfarb, A., Que, V. F., 2023, "The Economics of Digital Privacy", *Annual Review of Economics*, 15(1), pp.267~286.

- (19) Gordon, L. A., Loeb, M. P., Lucyshyn, W. and Zhou, L., 2015, "Increasing Cybersecurity Investments in Private Sector Firms", *Journal of Cybersecurity*, 1(1), pp.3~17.
- (20) Gordon, L. A., Loeb, M. P., Lucyshyn, W. and Zhou, L., 2018, "Empirical Evidence on the Determinants of Cybersecurity Investments in Private Sector Firms", *Journal of Information Security*, 9(2), 133~153.
- (21) Li, H., N, W. G. and Wang, T., 2018, "SEC's Cybersecurity Disclosure Guidance and Disclosed Cybersecurity Risk Factors", *International Journal of Accounting Information Systems*, 30, pp.40~55.
- (22) Martin, K. D., Borah, A. and Palmatier, R. W., 2017, "Data Privacy: Effects on Customer and Firm Performance", *Journal of Marketing*, 81(1), pp.36~58.
- (23) Janssen, R., Kesler, R., Kummer, M. E. and Waldfogel, J., 2022, "GDPR and the Lost Generation of Innovative APPs", NBER Working Paper, No.30028.
- (24) Jia, J., Jin, G. Z. and Wagman, L., 2021, "The Short-Run Effects of the General Data Protection Regulation on Technology Venture Investment", *Marketing Science*, 40(4), pp.661~684.
- (25) Kollnig, K. and Binns, R., Van Kleek, M., Lyngs, U., Zhao, J., Tinsman, C. and Shadbolt, N., 2021, "Before and After GDPR: Tracking in Mobile APPs", *Internet Policy Review*, 10(4), pp.1~30.
- (26) Kox, H., Straathof, B. and Zwart, G., 2017, "Targeted Advertising, Platform Competition and Privacy", *Journal of Economics & Management Strategy*, 26(3), pp.557~570.
- (27) Kshetri, N., 2014, "Big Data's Impact on Privacy, Security and Consumer Welfare", *Telecommunications Policy*, 38(11), pp.1134~1145.
- (28) Romanosky, S., Telang, R. and Acquisti, A., 2011, "Do Data Breach Disclosure Laws Reduce Identity Theft?", *Journal of Policy Analysis and Management*, 30(2), pp.256~286.
- (29) Tang, Q. and Whinston, A. B., 2020, "Do Reputational Sanctions Deter Negligence in Information Security Management? A Field Quasi-Experiment", *Production and Operations Management*, 29(2), pp.410~427.
- (30) Zhuo R., Huffaker B. and Greenstein S., 2021, "The Impact of the General Data Protection Regulation on Internet Interconnection", *Telecommunications Policy*, 45(2), No.102083.

Research on the Data Governance Pathway that Coordinates Development and Security

Xu Ji^a, Yang Qingxu^b and Hou Jiayi^b

(a. Institute of Economics, Chinese Academy of Social Sciences; b. Faculty of Business and Economics, The University of Hong Kong)

Abstract: Leveraging high-efficiency data governance to foster a virtuous cycle between high-quality digital economy development and high-level data security is both urgent and necessary. However, numerous international studies suggest that the security benefits of data governance often come at the expense of development, making it challenging to establish a governance pathway that coordinates development and security. To evaluate China's data governance practices, this paper employs big data from App security testing and a difference-in-differences model to analyze the policy effects of the Opinions on Guiding and Regulating the Orderly and Healthy Development of Educational Mobile Internet Applications. First, the research findings show that the policy significantly reduces security vulnerabilities and data violations in educational Apps, effectively enhancing enterprises' risk prevention capabilities and compliance management standards. Second, it helps maintain market order by compelling high-risk Apps to accelerate security transformation, thereby narrowing the data security gap with their low-risk counterparts. Third, the policy achieves more effective regulation of enterprises with greater market power and products with higher market visibility. Fourth, the policy implementation expands the application market scale and facilitates market entry for Apps with relatively limited capital and user resources. The research indicates that a governance pathway characterized by lowering entry barriers through registration-based systems, breaking institutional silos through coordinated collaboration, and enhancing governance capacity through co-construction, co-governance, and benefit-sharing can coordinate development and security, achieving a balance between market vitality and order in the digital economy while reconciling economic efficiency with social equity.

Keywords: data governance; coordinating development and security; high-quality development of digital economy; high-level data security

Research on the Data Governance Pathway that Coordinates Development and Security

Xu Ji^a, Yang Qingxu^b and Hou Jiayi^b

(a. Institute of Economics, Chinese Academy of Social Sciences; b. Faculty of Business and Economics, The University of Hong Kong)

Summary: Data governance is critical to both national security and social development. Effective data governance must ensure robust data security while promoting the high-quality digital economy development. However, establishing a governance pathway that balances development and security remains challenging, as numerous international studies suggest that security benefits often come at the expense of development. As a major digital economy, China has undertaken extensive explorations in data governance, and summarizing these practices can contribute valuable insights to global governance efforts. This paper examines China's approach to coordinating development and security through the lens of the Opinions on Guiding and Regulating the Orderly and Healthy Development of Educational Mobile Internet Applications. This policy represents China's first nationwide comprehensive regulation of educational apps, incorporating internationally advocated governance concepts such as integrated governance frameworks, improved incentive and penalty mechanisms, strengthened risk early-warning systems, and precision governance. It also introduces innovative measures including registration-based reform, coordinated collaborative regulation, and co-construction, co-governance, and benefit-sharing mechanisms, reflecting an inclusive and prudent regulatory approach consistent with China's longstanding data governance philosophy.

Using big data from app security testing and a difference-in-differences methodology, this paper analyzes the policy's effects and reveals four key findings. First, the policy enhances data security by significantly reducing security vulnerabilities and violations in educational apps, effectively strengthening enterprises' data security risk prevention capabilities and compliance management standards. Second, it maintains market order by compelling high-risk apps to accelerate security transformation, narrowing the security gap with low-risk counterparts and preventing "bad money driving out good" market dynamics. Third, it ensures market fairness by imposing stronger regulatory constraints on apps with greater capital strength and user resources, more effectively regulating enterprises with substantial market power and products with higher market visibility without adding compliance burdens to those in weaker market positions. Fourth, it fosters market cultivation by facilitating market entry for apps with limited capital and user bases, preventing dominant enterprises from leveraging compliance advantages to create closed, monopolistic digital ecosystems, thereby enhancing overall market vitality.

Based on these findings, this paper proposes three policy recommendations: dismantling information barriers across administrative divisions to build a comprehensive data governance system grounded in systems thinking; improving monitoring and supervision mechanisms to safeguard innovation through service-oriented governance; and establishing a social governance community to create an ecosystem characterized by co-construction, co-governance, and benefit-sharing. The main contributions of this paper include: employing big data with broad coverage and high representativeness to comprehensively examine data governance pathways across security and development dimensions; providing unique insights and quantitative analysis of China's current data governance landscape; offering empirical support and policy recommendations for relevant policy implementation and academic research.

Keywords: data governance; coordinating development and security; high-quality development of digital economy; high-level data security

JEL Classification: O38, O33

《统筹发展与安全的数据治理路径研究》附录

附录一 简化固定效应设定

附表1第(1)~(2)列展示了不含任何固定效应的基础结果,第(3)~(4)列不控制APP个体固定效应,但控制时间固定效应和APP类别固定效应。结果表明,即使在显著简化固定效应设定的情况下,政策效应仍然稳健显著,系数大小基本保持一致。

附表1 更改模型固定效应

	(1)	(2)	(3)	(4)
	安全漏洞数	违规行为数	安全漏洞数	违规行为数
$Treat_t \times Post_t$	-1.183*** (0.110)	-0.135* (0.074)	-1.194*** (0.110)	-0.134* (0.074)
$Treat_t$	-2.157*** (0.154)	-0.078 (0.138)	—	—
$Post_t$	-0.790*** (0.014)	-0.195*** (0.019)	—	—
APP固定效应	否	否	否	否
时间固定效应	否	否	是	是
时间×城市固定效应	否	否	否	否
APP类别固定效应	否	否	是	是
观测值	2325668	381201	2325668	381201
R ²	0.001	0.001	0.058	0.012

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

附录二 控制不同APP类别的时间趋势

在基准模型的基础上引入了APP类别与时间趋势的交互项,控制不同类别APP特有的时间趋势。结果显示,《规范教育APP意见》的政策效应仍然显著为负,本文的主要发现仍然是稳健的。

附表2 控制不同APP类别的时间趋势

	(1)	(2)
	安全漏洞数	违规行为数
$Treat_t \times Post_t$	-1.107*** (0.120)	-0.119** (0.050)
APP固定效应	是	是
时间固定效应	是	是
时间×城市固定效应	是	是
APP类别×时间趋势	是	是
观测值	2325668	381201
R ²	0.835	0.939

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

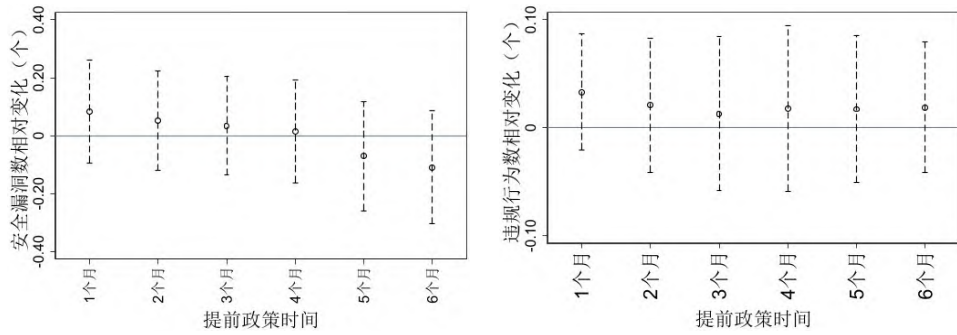
附录三 安慰剂分析

为验证本文识别的政策效应确实源自《规范教育APP意见》的实施,而非其他时间趋势或不可观测因素的干扰,本文实施了两种安慰剂检验策略。

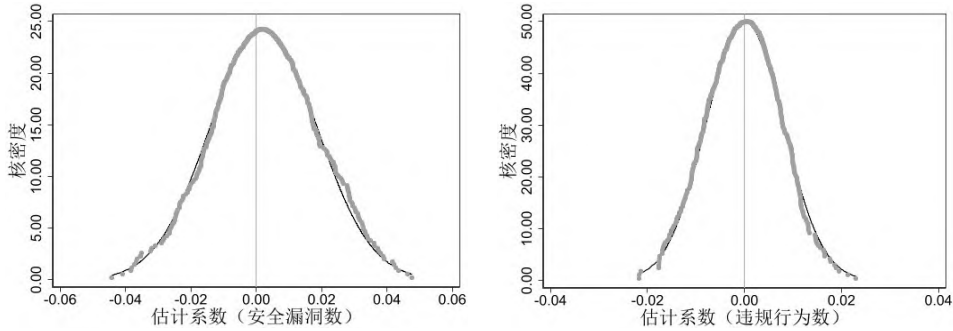
首先,本文进行了政策时间点置换检验。如附图1所示,分别将政策实施时间人为提前1~6个月,构造了6个虚拟政策时间点,并在这些“虚假”时间点上重新估计处理效应。若本文识别的效应真正来自政策实施,则在政策真实生效前的时间点不应观察到显著的处理效应。检验结果表明,无论是安全漏洞数还是违规行为数,所有虚拟时间点的处理效应估计值均不显著。这证实了在政策实际实施前,处理组与对照组之间并不存在系统性差异趋势,验证了前文的核心发现,而非其他因素引起的趋势差异。

其次,本文实施了处理组随机置换检验。如附图2所示,在保持总观测值不变的前提下,随机指定相同比例的APP为“处理组”,重复500次模拟实验并记录每次的处理效应估计值,绘制了系数分布直方图。结果显示,随机指定处理组后的系数估计呈现典型的钟形分布,中心位于0附近,与本文的实际估计值相差甚远。这意味着仅通过随机抽样几乎不可能得到与实际观测到的政策效应相当的估计结果,进一步排除了样本选择或模型设定可能导致的伪因果关系。

两种安慰剂检验的结果共同支持了本文因果识别策略的有效性,增强了本文对《规范教育APP意见》确实产生了实质性政策影响这一核心结论的可信性。政策效应估计结果既不受处理时间点选择的影响,也不可能通过随机分配处理状态而偶然获得,表明本文观察到的APP安全改善确实可归因于监管政策的实施。



附图1 安慰剂检验:政策提前1~6个月



附图2 安慰剂检验:随机500次选取处理组

附录四 倾向得分匹配法相关检验及补充图表

附录就正文倾向得分匹配法进一步进行PSM平衡性检验,证实了匹配的有效性,匹配后的处理组与对照组在各项关键特征上不存在显著差异。

附表3 PSM平衡性检验

变量名	匹配状况	均值(处理组)	均值(对照组)	组间差异t检验	p值
下载量(取对数)	匹配前	15.461	15.120	2.32	0.036
	匹配后	15.461	15.37	0.77	0.442
更新频率(次/月)	匹配前	1.043	1.024	0.54	0.520
	匹配后	1.043	1.033	0.28	0.793
APP索取权限数(个)	匹配前	26.584	26.850	-0.59	0.561
	匹配后	26.584	26.650	-0.12	0.902
安装包大小(GB)	匹配前	27.163	32.052	-3.06	0.004
	匹配后	27.163	26.483	0.35	0.733
APP年龄(月)	匹配前	36.671	37.909	-1.84	0.074
	匹配后	36.671	36.016	0.78	0.433

为缓解样本极度不平衡对估计结果的扭曲,除使用倾向得分匹配法外,附录进一步使用同一类别APP的数量做倒数加权回归,回归结果依然显著,且与基准结果相近。

附表4 使用同一类别APP的数量做倒数加权回归

	(1) 安全漏洞数	(2) 违规行为数
$Treat \times Post_t$	-1.327*** (0.110)	-0.128** (0.055)
APP固定效应	是	是
时间固定效应	是	是
时间×城市固定效应	是	是
观测值	2325668	381201
R ²	0.752	0.915

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。

附录五 排除疫情因素干扰

考虑到新冠疫情可能对 APP 市场产生冲击,特别是对某些特定类别 APP 的使用需求产生显著影响。2020 年疫情期间,游戏类、生活实用类(含外卖等)、社交通讯类(含远程会议等)、网上购物类等 APP 的用户增长较为明显,这些 APP 类别的快速发展可能会对本文的政策效应识别产生干扰。为排除新冠疫情所产生的潜在影响,本文在回归样本中剔除了上述疫情期间用户增长较多的 APP 类别,重新应用模型(1)进行回归分析。附表 5 的第(1)和第(2)列为剔除疫情相关 APP 后的回归结果。在控制了疫情期间特定 APP 类别快速增长等潜在干扰因素后,本文的回归结果依然显著。

附表 5 排除疫情因素的干扰

	(1)	(2)
	安全漏洞数	违规行为数
$Treat_t \times Post_t$	-1.367*** (0.111)	-0.132** (0.056)
APP 固定效应	是	是
时间固定效应	是	是
时间×城市固定效应	是	是
观测值	1271438	218053
R ²	0.840	0.935

注:括号中的数值为在 APP 层面聚类的稳健标准误;***、**和*分别表示在 1%、5%和 10%水平下显著。

附录六 异质性分析稳健性检验

在正文中,本文采用连续变量衡量 APP 事前安全风险水平和市场实力,构造了三重交叉项。附录中,本文进一步应用分类变量的方式构造相关变量,进行异质性分析的稳健性检验。

本文生成事前 APP 风险水平新变量——是否为高风险组 $Risky_t$,当安全漏洞数或违规行为数高于全市场同期 APP 的中位数时,定义为高风险组,赋值为 1,反之定义为低风险组,赋值为 0。其中,附表 6 第(1)列因变量为安全漏洞数,高风险组根据安全漏洞数高于中位数判定,第(2)列因变量为违规行为数,高风险组根据违规行为数高于中位数判定。附表 6 结果依然支持正文结论。

附表 6 使用分类变量:数据治理与市场秩序

	(1)	(2)
	安全漏洞数	违规行为数
$Treat_t \times Post_t \times Risky_t$	-1.920*** (0.226)	-0.193** (0.098)
$Treat_t \times Post_t$	-0.622*** (0.116)	-0.013 (0.061)
APP 固定效应	是	是
时间固定效应	是	是
时间×城市固定效应	是	是
观测值	2190188	309131
R ²	0.838	0.939

注:括号中的数值为在 APP 层面聚类的稳健标准误;***、**和*分别表示在 1%、5%和 10%水平下显著。

并且,本文构造 APP 开发商资历和 APP 用户规模的分类变量。构造衡量 APP 开发商资本实力的变量 $Large_Developer_t$,对 APP 开发商的注册资本在所有 APP 开发商中进行排序,将开发商注册资本超过中位数的 APP 定义为大型开发商 APP,对 $Large_Developer_t$ 变量赋值 1,反之,定义为小型开发商 APP,赋值 0。类似地,本文还构造 $Large_App_t$ 变量衡量 APP 用户规模,对政策发生前 APP 下载量在全部 APP 中进行排序,超过中位数的 APP 定义为大型 APP,对 $Large_App_t$ 变量赋值 1,反之,定义为小型 APP,赋值 0。附表 7 结果依然支持正文结论。

附表7 使用分类变量:数据治理与市场公平

	(1)	(2)	(3)	(4)
	安全漏洞数	违规行为数	安全漏洞数	违规行为数
<i>Treat</i> × <i>Post</i> × <i>Large_Developer_i</i>	-0.599*** (0.232)	-0.275*** (0.108)	—	—
<i>Treat</i> × <i>Post</i> × <i>Large_App_i</i>	—	—	-0.764*** (0.2084)	-0.333*** (0.0355)
<i>Treat</i> × <i>Post_i</i>	-1.011*** (0.151)	0.015 (0.074)	-0.824*** (0.153)	0.127*** (0.025)
APP固定效应	是	是	是	是
时间固定效应	是	是	是	是
时间×城市固定效应	是	是	是	是
观测值	1675163	299057	2324041	376991
R ²	0.821	0.938	0.835	0.939

注:括号中的数值为在APP层面聚类的稳健标准误;***、**和*分别表示在1%、5%和10%水平下显著。