

数据垄断问题研究进展*

沈坤荣 林剑威

摘要:在数字经济时代,数据要素在经济发展中的作用日益凸显,但是,数据垄断问题制约了数据要素价值的释放。本文首先探讨了数据垄断的内涵,然后从数据要素本身的特性出发分析数据垄断问题的成因,并从宏观经济效率和微观用户福利两个方面综述数据垄断带来的负面影响。之后,本文从数据共享、用户数据使用限制、数据可携带权与用户数据可迁移性三个方面,重点分析当前主要的数字垄断治理方式及政策。最后,本文总结并提出了数据垄断治理研究的未来可拓展方向,以期为我国加快构建数据基础制度、完善数据治理的相关研究和政策实践提供有益启示。

关键词:数据垄断 数据共享 数据权利 数据治理

加强平台反垄断和数据治理已成为全球共识。依靠数据制度安排来治理当前的平台垄断问题,以及通过数据制度的建设在“事前”规避垄断问题的出现,是当前全球各主要国家的反垄断与数据治理的重要内容之一。我国2022年12月出台的《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》也明确指出要打破数据垄断,促进公平竞争。

国内外学界对数据造成的垄断问题展开了一系列研究。在充分借鉴吸收国外相关成果的基础上,针对我国数字经济发展的具体国情开展深入研究,探索数据垄断治理的中国方案,有利于我国充分利用在数据资源上的大国优势、在数据应用场景上的先行优势,加快实现数字经济高质量发展。基于此,本文系统梳理和归纳了现有国外文献关于数据垄断问题的研究进展,重点总结和评述对于数据垄断问题的治理方法,并展望未来可能的研究方向,以期能够为学界深化该领域的研究和有关部门完善数据治理的具体实践提供有益参考。

一、数据垄断的内涵

一般而言,大企业或大平台在数据、资本、技术等方面均占有优势,但随着数据作为一种生产要素在经济活动中的作用逐渐凸显,数据在成为企业的重要资产的同时,也日益成为大企业或大平台的市场垄断势力的重要来源(Veldkamp, 2023)。已有一些文献讨论了数据与市场结构的关系,实证检验结果证实了数据在推动形成垄断性市场结构中的作用。Crouzet & Eberly(2019)研究发现,以数据、软件等为主的无形资本的增长是由行业领导者即大企业所推动的,行业集中度的提高和这些无形资本的上升有关。Tambe et al.(2020)研究发现,自20世纪90年代以来,数字资本在企业当中的积累十分不平衡,主要集中在小部分的“超级明星公司”(superstar firms)中,并且数字资本的集中度远远高于其他资产。Autor et al.(2020)基于美国1982年以来的经济普查微观数据研究发现,正是在数据和数据技术上的无形资本的优势催生了超级明星公司,从而造成了行业集中度的显著上升。

尽管数据已经成为市场垄断势力的重要来源,但大企业或大平台对于数据资源的独占或垄断是否就是“数据垄断”,在理论上存在一些争议。一方面,企业对于数据资源的独占可能是合理的。从

* 沈坤荣、林剑威(通讯作者),南京大学商学院,邮政编码:210093,电子邮箱:shenkr@nju.edu.cn, linjianwei@smail.nju.edu.cn。基金项目:国家社会科学基金重大项目“我国高质量发展的能力基础、能力结构与推进机制”(19ZDA049)。感谢匿名审稿人的宝贵修改建议,文责自负。

生产过程上看,与其他生产要素一样,企业必须先占有数据,才能使用数据,进而发挥数据的生产力价值。企业需要在数据的收集、记录、存储等方面进行投资才能获得数据,企业为数据的生成支付了对价,从而应该享有相应的数据占有与使用的权利(Veldkamp & Chung, 2019)。另一方面,企业对数据资源的独占可能是效率的体现和要求。Autor et al. (2020)研究发现,行业集中度上升的行业也是创新速度最快的行业。一些企业最初凭借其创新能力和效率,通过合法竞争而获得大量市场份额和数据资源,只是一旦获得主导地位,就会反过来利用数据资源建立各种壁垒。因此,垄断企业对数据资源的独占可能是其卓越能力的体现和结果。并且,由于数据开发利用的规模经济效应以及网络效应,大企业对于数据资源的占有可能更有利于发挥数据的价值。因此,大企业或大平台对数据资源的独占并不一定构成需要对其进行反垄断治理的充分条件。

对于数据垄断的认定,一些文献将重点放在平台如何利用数据达成或强化市场支配地位方面,当平台利用数据维护和强化其市场垄断时,即构成“数据垄断”。市场份额或市场集中度、超额利润常常被视为认定企业是否具有市场支配地位的重要因素。但是,从市场份额的角度来看,数字经济中存在明显的网络效应,这往往会造成“赢家通吃”的局面,使得数字经济中的市场集中度更高,数字平台也天然具有垄断特性,因此,较高的市场份额可能恰好是市场均衡的结果,而并不一定代表平台具有垄断势力(Tirole, 2023)。从利润的角度来看,当前的收入和利润可能难以捕捉平台在数字经济中的市场竞争优势。Veldkamp (2023)认为,数字经济催生了许多新的商业模式,使货币收入无法准确揭示企业正在产生和积累的价值,例如优步(Uber)和亚马逊(Amazon)一类的数据密集型企业都亏损多年,而这些企业积累的数据资产才是真正有价值的。这些设法积累数据并将其货币化的企业可以在后续获得市场主导地位,并利用这一地位提取垄断租金。因此,在平台经济当中,以市场份额和利润判断企业是否利用数据形成了市场垄断,需要更为慎重的具体考量。

沿袭双边市场理论的建立和发展,已有研究对于平台垄断市场势力的认定和分析不仅仅只考虑市场份额或垄断利润,对于企业垄断行为的分析也不仅仅只是局限于价格和产量,而是纳入更多的非价格效应(Rochet & Tirole, 2003; Weyl, 2010)。Khan (2017)分析了亚马逊的商业战略及其形成垄断地位的条件和影响,认为不能仅仅以价格和产量来衡量亚马逊的主导地位对市场竞争的损害,需要进一步关注平台会利用收集到的数据来削弱竞争对手的地位。Bamberger (2017)以打车平台优步为例直接提出了认定平台是否构成垄断需要考虑的几个问题,从数据的角度看,包括是否利用数据优势建立市场竞争壁垒,是否利用数据不公平地在其他市场建立主导地位,是否利用数据建立价格歧视、价格合谋、掠夺性定价,是否限制消费者的隐私权利等。Khan (2018)分析了平台市场势力的来源,认为平台的市场势力具备多种形式,包括平台基于对数据的占有对其他竞争企业进行限制、将主导地位拓展至其他市场、利用收集的用户数据强化消费者对平台的依赖等。Van Dijk (2019)认为,消费者剩余的概念不足以解释平台垄断势力对用户的影响,对于平台垄断势力的分析既需要考虑歧视性定价等对消费者剩余的短期影响,也需要考虑对用户隐私等方面的潜在的长期后果。可见,当前已有研究主要关注平台的垄断势力对市场竞争和用户福利的影响,对于“数据垄断”的认定也主要集中于垄断平台如何利用数据优势排斥市场竞争、降低用户福利。

综合已有文献相关研究,可以界定数据垄断的内涵:数据垄断不仅仅是“数据占有的垄断”,更多的是企业或平台“基于数据的垄断”,即利用数据资源占有或垄断的优势,以数据作为其维护和强化市场垄断地位、获取垄断利润的工具。当企业或平台的数据垄断由于排斥市场竞争造成经济效率下降、造成用户福利损失时,数据垄断就成为需要治理的数据垄断“问题”。此时,有关政府机构就具有对其进行相应的反垄断规制和治理的必要性及法理性。

学界的相关研究也已将考虑和分析企业或平台垄断问题的视角深入数据维度。诸多研究直接从数据本身及其特性出发分析数字企业或平台如何利用数据形成或强化市场垄断势力,从而导致经济效率下降和用户福利损失等问题;在此基础上,进一步分析对数据的共享、交易和权利分配等制度

性安排如何影响企业和用户决策,从而规制和避免数据垄断。后文将从数据本身的特性出发,具体综述数据垄断的成因及影响,并重点综述数据垄断治理手段的有效性和利弊。

二、数据垄断的成因与影响

(一)数据垄断的成因

数据如何导致市场垄断问题的出现或者加剧市场垄断问题?从最新的相关研究来看,主要可以从以下几个方面进行归纳。

1. 数据成为市场进入的壁垒。这一成因主要与数据要素的特性——内生生成性(endogenous generation)有关。Cong et al.(2021)指出,数据的内生生成特征是将其与劳动和资本区别开的关键特征之一。数据的内生生成性指的是数据属于经济活动的“副产品”(by-products),只要有经济活动,就有数据产生,经济活动越多,可产生的数据也就越多(Jones & Tonetti, 2020)。具体地,在将数据纳入数理模型分析的诸多文献当中即表现为,企业所拥有的数据量是其产量或者用户数量的增函数。这意味着,拥有较大市场份额的大企业基于其产量规模和用户规模可以生成和掌握更多的数据资源。因此,如果数据是企业的必要或重要的投入要素,新进入企业由于缺乏数据资源,将难以同在位企业竞争。Calvano & Polo(2021)指出,数据构成了垄断企业的“在位优势”,在位企业的数据基础成为其他企业进入市场所必须支付的高昂的“固定成本”,这可能会阻止其他企业进入市场。Tirole(2023)直接指出,正是因为谷歌(Google)和 Facebook 可以访问其他人无法访问的大量数据,这使得其在搜索和广告展示方面占据主导地位,诸多企业一开始就会放弃进入该市场进行竞争。

2. 数据强化垄断企业的优势。既有研究表明,企业掌握数据资源并基于此进行数据挖掘和分析,形成“数据驱动决策”(data-driven decision-making),可以优化企业的生产和管理流程,降低成本,提高生产率(Brynjolfsson & McElheran, 2016; Müller et al., 2018; Rammer et al., 2022)。特别是,数据愈发成为支撑创新活动的核心要素,推动创新活动加速数字化,可以显著促进企业创新能力和创新水平的提升(Gao & Hitt, 2012; Wu et al., 2020)。但是,一方面,数据的内生生成特征意味着原本占据较大市场份额的大企业拥有更多的数据资源;另一方面,大企业在数据分析能力、数字工具运用方面的能力都更强,相对于小企业,大企业将从数据中获得更大的好处(Bai et al., 2021)。

因此,数据的作用在大企业和小企业之间是不对称的,而这将进一步强化大企业的优势,扩大企业之间的差距,推动大企业形成更强的市场势力。Begenau et al.(2018)构建了一个企业博弈模型,考察了在金融市场中数据如何推动大企业发展壮大。大企业不仅比小企业更擅长利用金融市场中的数据来降低资本成本,而且在融资过程中,由于大企业有更多的经济活动和更长的历史,产生了更多的数据,这使得外部投资者能够基于数据分析更好地对大企业进行预测,从而使大企业的投资不确定性和投资成本相比小企业更低,即大企业和小企业之间的风险溢价差距扩大了,这对于大企业而言,意味着数据要素使得其融资成本下降更多,因此,大企业能够在金融市场中占据优势,大企业变得越来越大。Eeckhout & Veldkamp(2022)也在模型中刻画了企业对数据的使用如何创造市场势力,数据的规模经济效应使数据丰富的企业能够以更低的边际成本和更大的规模组织生产,大企业可以将生产配置到消费者想要的商品上,从而获得更高的价格加成。

更为重要的是,数据的开发利用具有自我强化的特征。企业对数据的开发利用可以改善用户体验或提高产品质量,企业有了更好的产品可以吸引更多的用户或现有用户更多使用,从而获取到更多的数据,这又可以反过来进一步优化产品,从而形成一个自我强化的循环,Hagiu & Wright(2021)将这个自我强化的循环过程称之为“数据支持学习”(data-enabled learning)的过程。Farboodi & Veldkamp(2022)在纳入数据要素的内生增长模型中刻画了数据的这一特征如何导致和加速企业规模的分化。拥有更多数据的大企业更具生产力,从而将有更多的生产和交易活动,而这又创造出更多的数据,并进一步提高大企业的生产力和数据生成,形成“数据反馈循环”(data feedback loop)。这种数据的递增收益会使那些数据拥有量低的企业、行业或国家长期局限于少量的生产和交易,从而

无法取得进一步发展。因此,数据将进一步放大企业之间的差距,使得原本的垄断企业和平台保持或进一步强化其市场支配地位。Prüfer & Schottmüller(2022)进一步在模型中考虑了多市场结构,并提出“关联市场”的概念,即数据所揭示的用户信息可能具有通用性,数据不仅可以降低企业在某个市场的创新成本,也可以降低与该市场相近或相似的“关联市场”的创新成本。因此,拥有大量关于用户偏好和特征数据的大企业在这些关联市场上也具有竞争优势,如果市场进入成本不高,能够找到“数据驱动的商业模式”的大企业,在长期内几乎可以主导任何市场。

3.数据成为垄断平台实施垄断行为的工具。垄断平台可能会依靠在数据资源和数据运用能力上的优势实施各种基于大数据的垄断行为,主要表现在两个方面:一方面,数据成为垄断平台获取超额垄断利润的工具。利用大数据技术,平台可以从数据中推断出关于消费者偏好的相关信息,从而进行价格歧视(Fainmesser & Galeotti, 2016; Fainmesser & Galeotti, 2020)。不仅如此,垄断平台还可以利用个人数据和人工智能技术操纵消费者支付更高的价格(Van Loo, 2017)。Van Loo & Aggarwal(2023)对亚马逊看似拥有垄断权力却向消费者收取低价的“悖论”进行了详细分析,揭示了该公司如何利用其复杂的算法和庞大的数据集来建立一个令消费者“误解”的市场。比如,亚马逊可能一直在以低廉、有竞争力的价格提供许多产品,但通过利用消费者的行为偏见,亚马逊可以让大量消费者永远找不到这些低价产品,消费者在亚马逊上购物可能会支付更高的价格。可见,利用数据和数字技术,垄断平台可以实施更为多样的攫取超额利润的垄断行为。

另一方面,数据成为垄断平台排斥和限制市场竞争的工具。比如,垄断平台基于垄断数据的优势,以搜索降权、流量限制、技术障碍等惩罚性措施强制第三方供应商接受交易,从而维持平台的市场优势地位。在这方面,已有的研究主要关注平台的跨市场特性(Khan, 2018)。平台利用数据算法,通过“自我优待”、强制性技术捆绑、搭售等行为,可以将垄断势力扩张到其他市场。对于既销售自己产品也销售其他供应商产品的垄断平台(如亚马逊),平台基于大数据技术在搜索排名、广告展示等方面给予自己的产品以优惠待遇或优先权的做法被称为“自我优待”(self-preferencing)(Reimers & Waldfoegel, 2023)。Zhu & Liu(2018)利用亚马逊的数据实证研究了平台进入供应商产品市场的策略模式,发现亚马逊可以从大数据中轻松检测和识别到热门的产品类别,然后会瞄准并进入这些热门的产品类别的市场,并在平台上通过“自我优待”等方式占据市场销售优势,最终,许多供应商将被挤出市场,而平台可以在这些市场形成支配地位。

(二)数据垄断的影响

数据垄断的负面影响可以从两方面归纳:一是从企业间竞争的角度,数据垄断导致市场竞争减少,经济效率下降;二是从企业与用户(消费者)之间的角度,企业依靠数据垄断攫取超额利润,侵占消费者剩余,最终损害用户福利,这其中还涉及消费者的隐私问题。

1.数据垄断与经济效率下降。随着数据成为大企业或大平台的市场垄断势力的重要来源,垄断性的市场结构将被加速推动形成,而这将导致经济整体创新减少、全要素生产率增速下降和经济增长率降低。Prüfer & Schottmüller(2022)构建了一个双寡头企业动态研发竞争模型,说明了数据垄断如何导致经济整体创新的减少。用户数据可以降低企业创新的边际成本、提高产品的质量。小企业知道大企业的数据优势使其能够以较低的边际创新成本匹配小企业的任何创新活动,从而会放弃创新;了解到这一点,大企业的最佳反应也是节省对创新的投资,并仍然获得垄断利润;一旦达到这一临界点,大企业就可以永远停止投资,并在所有剩余时期享受垄断利润。因此,企业间博弈的最终结果是,无论是大企业还是小企业,都没有激励投资于创新。

随着数据在经济活动中的作用日益凸显,以数据、软件等为主的无形资产也在企业中大量积累(Tambe et al., 2020),一些文献认为,企业间无形资产的差距可以解释行业集中度上升、经济整体创新效率下降和长期增长率下降。Crouzet & Eberly(2021)、Crouzet et al.(2022)讨论了无形资产的上升如何与近年来美国的行业集中度上升和全要素生产率增速下降的现象相关联。De Ridder(2024)构建了一个考虑企业在无形资产上具有异质性的熊彼特增长理论模型,具体分析了以数据为主的无

形资本如何导致具有高无形资本的大企业具有更高的市场势力,并降低经济整体的创新效率。在标准的垂直创新框架下,企业间只有所生产产品质量的异质性,通过创新开发出高质量产品的企业成为该产品的唯一生产商,但是,当考虑无形资本差异时,具有高无形资本的企业可以利用其在数据等方面的优势弥补产品质量的差距,从而削弱创新者,这使得大多数低无形资本的初创企业无法与具有高无形资本的大企业竞争,选择放弃创新和进入该产品市场。最终,积累了高无形资本的大企业始终具有更大的市场势力,经济的整体创新和经济增长率却下降。这些文献表明,由于数据垄断所导致或加剧的企业间无形资本的差距,将影响市场竞争并最终导致经济效率下降。

2. 数据垄断与用户福利损失。平台或企业通过所掌握的用户的购买历史和个人特征等数据,基于大数据技术,可以推断和掌握用户的个人偏好等方面的信息。一方面,这使得企业可以更好地进行产品推荐和提供个性化服务,从而使得用户受益。因此,从这个角度看,由于数据开发利用的规模经济效应,大平台或企业的数据垄断将有利于更好发挥数据的价值,用户将从中受益。但是,另一方面,这也使得平台可以从中得到各消费者对产品的保留价格,从而更好地进行价格歧视以攫取超额利润,从而侵占了消费者剩余,减少了用户福利。如果价格歧视导致的消费者剩余的减少超过了个性化带来的好处,最终导致用户福利损失,那么从用户福利的角度看,这也构成了需要治理的数据垄断问题。Bergemann et al.(2015)证明,当垄断企业具有关于消费者偏好的额外信息时,可以完全分割市场并进行价格歧视,并实现消费者剩余和生产者剩余的每一个组合。此时,尽管交易是公平的,但垄断企业可以从公平贸易中获取所有收益使得消费者剩余为零。Ichihashi(2020)构建的博弈论模型比较了个性化推荐与价格歧视对用户的影响。在这个模型中,企业可以从用户数据中了解用户的偏好,设定价格,并进行产品推荐,此时,用户会选择保留一部分个人数据以避免遭受价格歧视。考虑到这一点,企业倾向于预先为每件产品设定较高的价格。最终,尽管用户可以从个性化推荐中收益,但用户在高价格中遭受的损失超过了个性化推荐获得的收益,企业对数据垄断的利用仍然会降低用户福利。

网络效应是分析平台数据垄断对用户福利的影响所需要考虑的另一个重要因素。一方面,数字平台往往具有典型的网络效应,用户从商品或服务中获得的价值或效用取决于其用户数量,因而,用户聚集在一个占主导地位的平台中可以得到更高的效用,平台由此所形成的数据垄断似乎是自然且合理的。相应地,用户全部聚集在一个主导平台的损失是失去了对多个差异化平台的选择。如果只有单一的主导平台,用户可以得到更大的网络效应的好处,而如果有多个差异化的平台,用户也可以从多样化的选择中提高效用。Farronato et al.(2024)首先在模型中比较了网络效应和多个差异化平台对用户效用的影响,并利用美国两个最大的宠物托管服务平台的合并作为外生冲击进行实证检验发现,当用户具有不同的偏好时,网络效应的好处与失去多样化选择的损失相互抵消,用户使用一个主导平台并不比使用两个独立平台获得更高的效用。因此,数据垄断所引起或强化的垄断性市场结构也会通过减少用户的多样化选择而导致用户福利损失。

此外,从用户福利的角度看,数据垄断所导致的另一个负面影响是用户的隐私问题。Goldfarb & Que(2023)指出,用户的隐私具有工具性价值,平台将数据作为实施垄断行为工具的重要方式之一,就是从用户数据中分析和推断出用户偏好等各方面的隐私信息。而用户对涉及其隐私的数据做出知情决定的能力受到严重阻碍,因为用户往往对何时何地出于何种目的被收集数据以及会产生何种后果处于信息不完善或信息不对称的状态。平台对用户数据的过度使用由于隐私问题会给用户造成负效用(Acquisti et al.,2016)。

三、数据垄断的治理方式与政策

与数据垄断的负面影响相对应,数据垄断治理的目标可以分为两个方面:一是防止数据成为市场进入和市场竞争的壁垒,促进市场有序竞争;二是防止垄断企业依靠数据垄断造成用户福利损失。学界早期对于平台垄断规制的研究主要分析了征税、处罚、业务拆分等传统规制手段的作用(Kind

et al., 2008, 2010; Bourreau et al., 2018)。但 McIntosh(2019)指出,数据的性质可能导致传统规制手段在某些方面无效,数据方面的垄断是反垄断政策需要考虑的更为紧迫的问题。与以垄断企业或平台为规制对象有所不同,当前学界的诸多研究将反垄断的关注对象深入数据本身,从针对数据的制度性安排出发,分析讨论数据治理制度如何避免大企业或大平台对数据资源进行垄断,以及如何避免数据成为实施垄断行为的工具,从而促进市场竞争和提高用户福利。具体地,本部分从数据共享、用户数据使用限制、数据可携带权与数据可迁移性三个方面对此展开综述。

(一)数据共享

1. 数据共享的合理性。数据共享是治理数据垄断问题的直接手段。作为官方发布的全球首部针对平台经济的系统性反垄断指南,我国2021年颁布的《国务院反垄断委员会关于平台经济领域的反垄断指南》明确提出将开放网络、数据或平台等基础设施作为救济措施之一。由于直接拉平了企业间数据资源的差距,垄断企业失去实施市场垄断行为的数据优势基础,与其他垄断治理方式相比,强制垄断企业将数据开放共享或许是治理数据垄断问题效果最为明显的手段之一。并且,由于数据具有非竞争性、低成本复制性的特征,将数据开放共享给其他企业使用,理论上并不影响原有企业的使用。因此,数据共享既能缓解企业数据资源占有的不平等,缓解数据垄断问题,又能促进数据的充分使用、发挥数据价值。

数据开放共享的合理性与垄断治理中的“必要设施原则”(essential facilities doctrine)有关。该原则是指,如果垄断企业控制了竞争所必须使用的基础设施,且该基础设施不能被重新建设,那么垄断企业负有允许竞争者使用该基础设施的责任。在数字经济时代,随着数据成为企业发展所必需的生产要素,令垄断企业开放共享数据以满足市场竞争和经济发展的数据需求具有一定的法理性(Khan, 2017)。Chen et al.(2022)分析了数据共享如何减轻大企业技术并购所导致的负面效应,认为数据共享加强了数据应用领域的竞争,使企业垄断势力更难形成。Tirole(2023)认为,促进数据共享或许可以在不破坏数据的网络外部性好处的情况下促进竞争,相比于拆分垄断企业的政策可能更有效。甚至于,开放共享数据可能对企业本身有利。Benzell et al.(2023)发现,一种追求成为“反向企业”(inverted firm)的模式已经出现,与利用数据构筑进入壁垒和形成独特资产的战略不同,这类企业主动开放共享数据,通过向第三方开放数字服务并获取部分外部盈余成为核心平台,其利用开放应用程序编程接口(API)这一关键的技术来定量估计采用反向企业战略的效果。实证研究发现,开放API的企业具有更高的市值,因为这些企业通过开放数据成功地将自己放置于企业生态系统的中心位置,并获得巨大回报。

2. 数据共享带来的损失。数据共享有可能降低企业进行数据投资的积极性,引发新的效率损失。如果考虑到这种成本,那么数据共享政策即使缓解了数据垄断问题,也不一定提升了社会的净福利水平。已有文献进一步考虑到了数据共享政策的负面影响,并讨论了其净福利效应。Hagiu & Wright(2021)在一个Bertrand模型中模拟了不同数据政策下的企业竞争,企业从数据当中学习并提高产品质量,产品质量较高的企业成为“领先者”,产品质量较低的企业被淘汰,成为“失败者”。数据共享政策会抑制两家企业收集数据的动机,但是最终对消费者剩余的影响取决于两家企业的差距有多大。如果两家企业的实力(特指企业从数据中学习的速度和效率)势均力敌,在没有数据共享政策时,可能的“失败者”为了不被淘汰,会向消费者提供补贴以获取数据,这提高了消费者剩余;而在数据共享政策的情况下,势均力敌的两家企业都没有动机补贴消费者而倾向于“搭便车”获得数据,补贴的取消对消费者的负面效应超过了两家企业长期保持竞争的积极影响,因此,数据共享政策降低了消费者剩余。如果两家企业实力差距较大,两家企业都不会提供补贴,因为获取更多的消费者数据对其竞争力的作用不大,此时如果有数据共享政策,会使“失败者”单方面受益而对“领先者”没有影响,从而数据共享政策是有利的。因此,Hagiu & Wright(2021)认为,数据共享政策改善消费者剩余的一个关键条件是,从数据共享中受益的“失败者”远远落后于“领先者”,即企业间的差距足够大。

Cong & Mayer(2023)在一个两阶段的双寡头竞争模型中模拟了企业的数据竞争,数据由用户的分享决策和企业的数据投资共同生成,数据共享的潜在好处是,可以增加所有企业的数据存量,从而提高整体的产品或服务的质量,同时可以降低过度的市场势力。同样,数据共享导致了搭便车问题——企业为了收集数据需要在第一阶段降低产品价格以吸引更多的用户、补贴用户的隐私成本,并且支付数据投资的成本,但是其他企业却得到了同样的数据收益。这使得企业收集数据的动机降低,导致第一阶段产品价格上涨,数据投资水平下降,因此,数据共享政策的目标原本是最大化可用数据,但实际上却会减少社会的总体数据存量,进而导致产品或服务的质量下降。当企业之间的实力(特指企业从数据中提高产品和服务质量的能力)对称时,数据共享政策导致数据收集不足以及第一阶段的价格上涨,降低了用户福利。当企业之间的实力不对称时,在数据共享政策下和均衡状态时,只有实力强的企业进行数据收集,实力弱的企业不进行任何数据投资,通过数据共享而获得了数据并分走了一部分市场,从而降低了市场集中度,市场价格下降,提高了用户福利水平。但是,由于两家企业的数据投资相较于基准情况大幅下降,社会数据存量减少导致的用户福利损失仍然超过了垄断问题缓解而带来的用户福利提升。因此,与Hagiwara & Wright(2021)不同,Cong & Mayer(2023)认为,数据共享总是会降低用户的福利水平。同样地,Jin & Vasserman(2021)以美国一家汽车保险公司的自愿监测计划为例,讨论了要求企业公开监测数据的反事实政策效果,研究发现,要求汽车保险公司公开共享监测数据会导致更少的监测、更少的数据生成和更低的消费者福利。

3. 引入数据交易机制。数据共享不一定能够提升社会净福利水平的一个重要原因是,对于垄断企业而言,数据共享只有成本而没有收益,因而其会调整数据投资行为应对数据共享政策。那么,如果给予企业的数据共享一定的收益,是否能够减少对数据投资激励的打击?更进一步地,如果存在数据交易机制或数据交易市场,基于对数据交易利润的追求,垄断企业是否会选择出售一部分数据?如果垄断企业选择出售数据,那么其他企业借由数据交易机制可以从垄断企业或者其他企业获取数据资源,一定程度上弥补与垄断企业在数据资源占有量之间的不对等。因此,数据交易机制的存在或许能够缓解数据垄断问题。

一些文献进一步考虑了当存在数据交易机制时是否能够激励企业有偿“共享”数据,从而缩小企业之间数据资源占有的差距,缓解垄断问题。企业能购买到的数据取决于其他企业愿意出售的数据量,对于出售数据的企业而言,需要权衡出售数据获得的收益和因竞争对手获得数据而导致的自身市场份额受损失的成本(Jones & Tonetti, 2020)。Choe et al.(2024)研究发现,出于对竞争的担忧,企业会有选择地进行数据分享和交易,分享更偏好的那部分消费者的数据,但保留对自己最偏好的那部分消费者的数据。因此,如果竞争企业不能得到想要的的数据,数据交易就无法在实质上缓解数据垄断问题。

此外,数据交易本身就面临诸多的挑战。数据要素交易仍面临权属界定、估值定价、流通规则和流通技术等关键性瓶颈问题。比如,数据的非竞争性和非排他性使得数据在售出之后,在当前技术条件和产权规则下,原始所有者很难控制数据被再次出售。Duch-Brown et al.(2017)认为,数据的外部性、数据所有者的战略行为和数据交换中的交易成本将导致数据市场失灵。数据市场化配置的现实矛盾,各国在实践过程中尚未形成体系成熟的数据交易市场,当前数据交易市场的整体规模仍较小,数据交易流转量普遍较低。Elsaify & Hasan(2020)的经验研究表明,只有17%左右的数字企业参与数据市场,50%以上的数据流通是以数据互换而非数据交易的形式进行的。因此,从已有相关研究与具体实践来看,数据交易作为构建数据基础制度的重要部分,主要目标仍在于通过交易流通提高数据价值、推进数据要素的充分利用,如果将数据交易作为一种数据垄断治理方式,其效果仍有待观察。

(二) 用户数据使用限制

限制企业使用用户数据,主要目的是通过直接限制企业对于用户数据的使用或滥用,减少企业利用数据进行价格歧视等的垄断行为,同时减少用户隐私问题造成的负效用。

1. 限制使用用户数据的有效性。Kirpalani & Philippon(2020)分析了限制使用用户数据在双边平台中的影响。用户共享的数据可以使平台上的买家和卖家更好地匹配,随着用户数据变得更加丰富和精确,平台不仅可以利用这些信息来生产或推广自己的产品,削弱卖家的竞争力,同时平台也可以“俘获”越来越多的买家,卖家更加难以在平台之外匹配到其他消费者,卖家的外部选择权减少,越来越依赖于平台,平台将由此收取更高的信息或服务价格,这不仅降低了卖方的收益,也降低了消费者福利。因此,Kirpalani & Philippon(2020)认为,适当限制平台对用户数据使用的监管政策可以降低卖家对平台的依赖性,削弱平台的垄断定价能力,从而可以增加整体社会福利。

2. 限制使用用户数据的损失。一方面,限制用户数据的使用也必须考虑失去数据的生产力价值带来的损失。Eisenmann et al.(2011)认为,直接限制平台使用用户数据可以为用户提供更强的隐私保护,但是这种策略的一个潜在缺点是,可能会阻止企业创建大型和丰富的数据库,从而降低数据的使用效率,降低用户福利。在 Jones & Tonetti(2020)纳入数据要素的内生增长模型中,模拟了限制数据使用、禁止企业将数据出售给任何第三方的政策效果。研究发现,虽然这类政策成功产生了隐私收益,但由于数据的非竞争性特征,数据没有被规模性使用而导致经济效率低下,反而带来了更大的经济增长效率损失。

另一方面,由于大企业和小企业承担监管和规制成本的能力不同,类似的监管或限制政策可能适得其反,加剧了企业间的不平衡。Campbell et al.(2015)在模型中模拟了强化用户数据监管或限制的政策如何影响数据密集型行业的市场结构,研究结果表明,尽管政策给所有企业都增加了成本,但小企业和新企业受到的不利影响最大,这种负面影响反而使大企业间接受益。Sun et al.(2023)进行了一项大规模的实地实验,与阿里巴巴合作,通过在其主页推荐算法中禁止使用用户数据来模拟政策监管。研究发现,没有使用用户数据的推荐会导致客户参与度和市场交易量大幅下降,客户对推荐产品的点击率立即下降了75%,更为重要的是,对于规模较小、评级较低的商家,客户参与度和市场交易的降幅更大。因此,限制使用用户数据的监管政策可能会对电子商务相关的创业和就业产生负面影响,并且政策对于不同参与个体的影响是不平等的,可能使得市场份额向规模较大的商家倾斜。Johnson et al.(2023)认为,大企业有更多的技术和资源来应对政策监管或限制,其实证研究了欧洲《通用数据保护条例》(General Data Protection Regulation,简称GDPR)施行以后网站的技术供应商的变化。研究发现,由于大型技术供应商能够提供更好的法规遵从性,能够更好地应对法律风险,因此网站更青睐于选择大型技术供应商作为合作伙伴,在GDPR实施之后,为网站提供技术支持的技术供应商市场的行业集中度显著提高。这意味着,限制使用用户数据的类似监管可能会加剧某些行业的市场垄断状况。

(三)数据可携带权与用户数据可迁移性

从前文对已有文献的总结可以看出,无论是数据共享还是直接限制用户数据的使用,都会打击企业对于数据的投资和开发的积极性。由于降低了数据的生成和使用,最终会导致社会福利下降。特别是,数据共享虽然可以缓解数据垄断,但是由于数据共享的对象是企业支付了相应数据投资成本而占有的数据,任何比例的数据共享都使得竞争企业可以“搭便车”,从而降低了企业进行数据投资的积极性。那么,如果企业间共享的是不包含企业数据投资成本的原始用户数据呢?

一些研究认为,如果通过企业间互相提供API接口等方式,令企业只共享原始的用户数据,或许可以避免对企业数据投资的打击。Graef & Prüfer(2021)认为只共享原始的用户数据是合适的。如果共享的是企业处理过的数据,可能会一开始就挤出企业的数据投资,但是共享关于用户选择或特征的原始数据不会导致这种威胁,因为在用户与产品或服务提供商(即企业)的交互过程中,这些数据可以自动记录下来,是企业与用户交互的免费副产品。如果只共享原始的用户数据,还会激励竞争对手开发自己的模型来分析用户信息,这可能会导致差异化产品,从而为消费者提供更多的选择。因此,这种数据共享对竞争和创新的好处可能会超过对主导企业的负面影响,Graef & Prüfer(2021)明确建议,当一家企业拥有超过30%市场份额时,就要求其共享原始的用户数据,并详细讨论了强制这种数

据共享的治理结构设计。Prüfer & Schottmüller(2022)在其模型中进一步证明了这种用户数据共享如何保障市场竞争从而增加社会福利。强制共享用户数据可以有效降低市场的进入壁垒,新进入者会持续进入并与在位者竞争,这使得在位的垄断企业保持创新动机,于是,在数字市场中,创新率永久为正的真正竞争是可能的。

可见,共享原始的用户数据似乎可以较好地解决数据垄断问题。而从实践层面上看,实现用户数据共享需要两个方面的具体制度安排:一是赋予用户数据可携带权,即用户享有对数据的控制权,用户可以决定是否允许企业收集存储其个人数据、对个人数据进行转移、交易的对象和价格;二是需要保障用户数据在不同平台或企业之间的可迁移性。

数据可携带权和用户数据可迁移性是当前世界各主要国家或组织的数据制度和法律规范中的重要内容。数据可携带权是指数据主体有权从数据控制者处获取个人信息副本,以及请求数据控制者直接将其个人信息传输给另一实体。我国2021年施行的《个人信息保护法》明确赋予个人“数据可携带权”。欧盟的《通用数据保护条例》(GDPR)以及美国的《加州消费者隐私法案》(California Consumer Privacy Act,简称CCPA)等诸多制度法案也都要求建立和保障用户的数据可携带权,并且提议企业,尤其是大型数字平台企业,应该提供相应的数据接口,方便用户将数据迁移到其他企业或平台,强化平台间的兼容性。

学界对个人数据可携带权和用户数据可迁移性展开了较为丰富的研究,但未达成一致结论。本节主要综述当强化用户数据权利时,即给予用户对数据的控制、转移、交易等权利时,对企业相关决策的作用和影响。从已有研究来看,一些文献强调了赋予用户数据权利的有效性,但由于隐私悖论、数据的外部性特征以及企业的策略性应对等原因,一些文献则认为类似的政策并不能导致数据的有效使用,对数据垄断治理的效果可能有限。

1. 强化用户数据权利的有效性。数据垄断问题的表现之一就是企业对大量用户数据的独占,因此,通过保障用户数据的可迁移性以及令用户自主决定和处置用户数据,以防止用户数据被锁定在单个企业之中,通过用户数据在企业间的互联互通达到破除数据垄断的效果;同时,由于是用户基于自身效用最大化的自主决策,也兼顾了隐私对用户造成的负效用问题。在Jones & Tonetti(2020)的模型中,如果将数据产权分配给消费者,可以使得数据的使用接近于社会计划者。尽管消费者会考虑到自己的隐私问题,但受到市场的激励,他们会将自己的数据广泛出售给诸多企业,从而充分挖掘数据的价值,实现近乎最佳的分配,而消费者福利也只是略低于最优水平。Jullien & Sand-Zantmann(2021)认为,平台之间的可互操作性(即原始用户数据在平台间的开发共享)可以阻止市场发生倾斜。Dosis & Sand-Zantman(2023)则认为,当产生数据的市场(即用户一侧)相对于使用数据的市场(即企业一侧)更重要时,消费者拥有数据权利可能会改善社会福利。

2. “数字隐私悖论”使得用户数据权利不能被有效使用。通过数据可携带权强化用户的数据权利并由用户对数据进行处置、交易的前提是,用户能够对个人数据的价值进行正确、合理的判断。一些研究已经表明,用户对于个人数据的处置和利用存在非理性。Athey et al.(2017)通过经济学实验发现,尽管用户声称重视个人数据及隐私,但是在现实中却很容易选择共享个人数据和披露个人信息,其将这种现象称之为“数字隐私悖论”(digital privacy paradox)。Chen et al.(2021)结合用户调查和支付宝使用的行为数据进行研究发现,被调查者自述的对个人数据和隐私的重视与数据共享授权数量之间没有关系,用户对数据及隐私的重视和担忧并没有改变其将个人数据共享给支付宝平台上第三方小程序的行为。可见,用户对于个人数据的保留价格实际上可能很低,企业通过很小的激励就能诱使用户共享其数据,或者放弃其数据权利。如果由于“隐私悖论”等原因,用户无法理性、有效地行使数据权利,那么也就无法与企业进行有效博弈,从而减少企业对用户数据的锁定、独占和滥用。

3. 数据的外部性特征可能使得用户数据权利没有意义。个人数据具有外部性,企业从个人用户获取的数据不仅提供了关于该用户的信息,还提供了关于具有类似特征或行为的其他用户的信息。这意味着,即使没有掌握某一部分用户的数据,企业仍然可以通过其他用户的数据推断并掌握该部

分用户的偏好等信息。从隐私的角度,Acemoglu et al.(2022)表明,当用户分享自己的数据时,其不仅损害了自己的隐私,也损害了与其信息相关的其他个人的隐私,此时,每个人都会忽视自己的隐私问题,并放弃自己的数据和信息,因为其他人授权与平台共享和使用数据的决定已经暴露了他的很多信息。Miklós-Thal et al.(2023)在数据市场模型中考虑了用户数据的多维性,将用户数据区分为敏感数据和非敏感数据,与数据的外部性特征类似,随着企业收集更多数据,了解个人数据不同维度之间的相关性,越来越善于从用户的非敏感数据中推断出用户的隐私敏感数据,用户对数据授权决策的两极分化便越严重。一方面,随着企业从非敏感数据中推断敏感数据的能力的提高,看重隐私的用户将决定避免授权企业共享和使用任何数据,即使是表面上不敏感的数据,从而导致“数字隐士”(digital hermits)的占比增加;另一方面,一些用户意识到只保留部分数据是没有意义的,从而将授权企业共享和使用所有数据,导致企业从一开始就收集敏感数据,并提供较低的数据价格。这些研究表明,由于数据的外部性特征,且用户无法阻止其他人授权共享数据,用户将会意识到,个人决定是否授权与平台共享和使用数据的数据权利是没有意义的。

数据的外部性特征也可能使得强化用户数据权利的做法只保护了一部分用户的利益,而对另一部分用户造成福利损失。Liu et al.(2020)在 Gul & Pesendorfer(2001)的基础上引入异质性消费者和“诱惑商品”(如赌博、电子烟、酒等)。在模型中,一部分消费者是“意志坚强的”,总是会抵制诱惑商品,一部分消费者是“意志薄弱的”,可能会沉迷于诱惑商品,诱惑商品将对消费者造成伤害。企业将利用数据进行广告、推销等,提高商品与消费者的匹配度。当用户行使数据权利并可以自主决定是否共享数据时,意志坚强的消费者将始终决定共享数据,因为这可以提高消费需求与企业提供的正常商品的匹配效率而获得正效用,消费者选择退出共享是不理性的。由于数据的外部性特征,意志坚强的消费者数据共享的决策也使得诱惑商品的卖家更容易瞄准意志薄弱的消费者,如果对意志薄弱的消费者的伤害大于意志坚强的消费者的收益,那么整体社会福利将降低。因此,Liu et al.(2020)认为,尽管 GDPR 和 CCPA 给予每个消费者选择加入或退出数据共享的选择权利,但可能无法为易受到诱惑的消费者提供足够的保护。如果考虑这部分人的伤害和损失,赋予用户数据权利并不一定会导致最理想的社会效率。类似地,Parlour et al.(2022)在贷款市场中考察了用户数据控制权的政策效果。在用户数据共享以后,来自金融科技企业的竞争增加了低信用质量用户的信贷供应,但是由于银行可以更准确地进行价格歧视,并且可能提高价格以应对来自金融科技企业的竞争,原本高信用质量的用户在贷款市场上的剩余可能会降低。因此,旨在让消费者更直接控制和共享数据的政策会产生相反效果。

4. 用户数据权利与企业的策略性应对。如前所述,由于隐私悖论、数据的外部性特征等原因,用户没有足够的动机来行使其数据权利。那么,如果用户对于数据权利的行使决策是完全理性的,用户数据权利能够避免数据垄断问题吗?从已有文献研究结论来看,如果进一步考虑企业的策略性行为,赋予或强化用户对数据的控制权,同样也不一定能提高社会整体福利。

第一,企业可能减少对数据收集的投资或补贴。与直接限制使用用户数据的政策一样,一些文献研究认为,令用户掌握数据的控制权可能扭曲企业的数据投资激励。Hagiu & Wright(2021)认为,赋予或强化用户数据权利这种让消费者更容易选择不分享数据的政策使得企业更难以收集数据,企业将减少对于消费者提供的隐私补贴,这将使得消费者福利减少。He et al.(2023)以开放银行为例,利用跨国数据集研究发现,在用户数据控制权下的数据共享增加了进入者,但减少了事前的数据生产。Ke & Sudhir(2023)将消费者拥有和控制个人数据的特征纳入前瞻性企业和消费者的动态两阶段模型发现,这种数据权利通过限制企业价格歧视和保障隐私提高了消费者福利,但也抑制了企业提供较低的初始价格以鼓励或补贴用户参与的动机。在激烈竞争的市场中,强化数据权利将有利于消费者,但是在市场垄断的情况下,这种数据权利会出人意料地伤害消费者。

第二,企业可以转移数据获取成本。一方面,企业可以将更高的数据获取成本通过产品或服务价格转嫁给用户。在 Cong & Mayer(2023)模型中,各个均衡值(特别是用户福利)均与数据价格或用户选择数据分享的比例无关。即使用户拥有数据且可以出售数据而获得收益,平台可以通过更高的服务

价格将收集数据的额外成本转嫁给用户,这正好抵消了用户出售数据而获得的额外收益。因此,当政策将数据权利赋予用户时,用户福利和大企业的市场份额保持不变。另一方面,在一部分用户行使数据权利提高企业对其数据的获取、使用的难度或成本以后,企业可以策略性地调整对其他用户的安排以应对这种变化。早期的一些文献主要以广告为例,分析了当用户试图回避广告营销时企业如何调整其策略并转移成本(Hann et al.,2008;Wilbur,2008;Anderson & Gans,2011;Johnson,2013)。Goh et al.(2015)实证研究了美国“不打电话”(Do Not Call,简称DNC)登记政策的影响,用户登记注册DNC后,销售人员就不被允许向登记用户打电话进行营销推广,这反而表明未登记的用户更容易接受电话营销,因而销售人员打给这部分用户的电话越来越多。可见,当用户行使数据权利进行相关决策时,可能对其他用户造成了负外部性,企业实际可以将成本转移给其他用户。Montes et al.(2019)研究了在不完全竞争条件下消费者拥有数据控制权对企业定价行为的影响,同样发现,在垄断市场情况下,让消费者更容易保护自己的数据对消费者产生了模棱两可的影响:一些人获益,另一些人受损。

第三,利用数据的外部性特征,企业仍然可以过度收集用户数据。一些文献讨论了企业如何利用数据的外部性特征与用户互动并过度获取和占有用户数据。Choi et al.(2019)在Bergemann & Bonatti(2015)的数据定价模型基础上考虑了用户数据的外部性,将用户数据分为有外部性的数据和没有外部性的数据,数据收集需要用户同意且用户充分意识到这种同意的后果。研究结果表明,如果企业向选择共享的用户提供一定的价格折扣,由于数据的社会性和用户协调的失败,过度收集用户数据的情况仍然可以在垄断市场均衡中产生,这仍然导致了福利损失。类似地,Ichihashi(2021)也证明了,在数据存在外部性的情况下,企业可以通过灵活地设计数据收集策略来减少用户拒绝共享数据的外部选择。Bergemann et al.(2022)分析了数据中介和具有相关偏好的多个用户之间的信息交易,数据的外部性放大了个人数据对于数据中介的价值,具有市场力量的数据中介机构可以通过对用户数据的整合来操纵用户对于数据决策的权衡,最终以接近于零的报酬水平诱使单个用户放弃有关其偏好的精确信息。从而,赋予用户对其数据的控制权不足以带来数据的有效使用。

如果进一步考虑企业的异质性,赋予用户数据权利的作用将更为复杂。He et al.(2023)以欧洲“开放银行”(open bank)倡议为例,分析了让用户(借款人)决定是否要共享数据对信贷市场竞争及社会福利的影响。开放银行允许借款人自己决定是否要与第三方分享其银行数据,这使得所有的贷款人(传统银行和金融科技企业)在关于获取借款人信息方面具有公平的竞争地位,这加剧了贷款人之间的竞争,从而有利于借款人。但是,其研究发现,由于金融科技企业拥有更先进的数据分析算法和金融科技,它将在筛选借款人方面超过银行,开放银行通过允许借款人共享其银行数据大大提高了金融科技企业作为“挑战者”的竞争力,这反而阻碍了竞争并使得所有的借款人的福利降低。这说明,赋予用户数据控制权的政策本意是促进市场竞争,由于企业的异质性,结果却可能适得其反。其背后的更深层含义是,企业对于数据的需求和使用具有不同的模式,对于数据的分析使用能力也不同,当数据政策变化时,不同企业将以不同的数据投资决策进行反应(Fainmesser et al.,2023),并受到不同程度的影响。因此,赋予或强化用户数据权利的相关政策手段将影响和重新塑造企业间的竞争互动格局(Casadesus-Masanell & Hervas-Drane,2015),相关政策应该进一步考虑对企业的异质性影响。

5. 集体行动。造成用户数据权利难以发挥作用的原因之一是,由于用户个人的非理性和数据的外部性特征,用户难以达成协调一致的行动与企业进行议价。Cong et al.(2022)在一个内生增长模型中证明,由于消费者在承担隐私成本的同时未能充分内化知识溢出,再加上企业的市场力量,导致数据定价过低和数据供应效率低下,导致创新部门就业不足和长期增长无效率。

一些文献进一步提出,在个人数据可携带权和用户数据可迁移的前提下,用户应该采取集体行动,构建类似于工会组织的“用户联盟”。通过用户联盟和企业之间的“竞争”或谈判,用集体行动来抵消个人理性的缺失并内化个人数据的外部性,在避免企业数据垄断的同时实现数据的充分使用。MacCarthy(2011)认为,仅仅依靠个人同意来确定信息收集和使用行为的合法性是不够的。Fairfield & Engel(2015)建议将隐私标记为公共产品,呼吁采取集体选择的方法来解决这个问题。Economides &

Lianos(2021)提出,由于享有主导地位的 digital 平台与个人用户之间的议价能力不对称,赋予用户个人数据控制权是不够的,数据的定价不应该完全由占主导地位的、拥有优越议价能力的 digital 平台来决定,建议由国家主管部门促成用户的集体谈判,以“收集协会”(collecting societies)的形式与平台就用户的个人数据补偿进行谈判。Cong & Mayer(2023)介绍并评估了用户联盟作为一种解决方案的可行性。用户联盟类似于一个提供补贴或者向用户征税以激励有效的数据共享的计划者,用户联盟考虑的是所有用户的共同收益,它将内化个人数据贡献的外部性。当企业之间是对称的,市场垄断不成为主要问题时,用户联盟可以设置一个补贴水平,通过补贴用户来激励用户的数据共享;当企业之间差距较大,垄断问题成为主要矛盾时,用户联盟将设置一个征税水平,以抑制用户数据共享和企业的数据收集,从而削弱主导企业的数据优势来纠正不平等的数据格局。

四、总结与研究展望

当企业依靠对数据资源的占有实施市场垄断行为,并利用各种手段排斥市场竞争而导致经济效率下降,以及利用各种手段侵占消费者剩余和侵犯用户隐私导致用户福利水平下降,数据垄断问题就出现了。在数据垄断治理方面,从学界已有相关研究来看,数据共享和限制使用用户数据这两种手段总是面临着两方面的权衡:缓解数据垄断而提高社会福利,以及扭曲企业的数据投资激励而导致社会福利下降。个人数据可携带权和用户数据可迁移性是当前世界各主要国家实施的数据制度的重点之一,从已有相关研究的结论上看,这似乎是缓解数据垄断问题的最有希望的方向,尽管由于隐私悖论、数据外部性等问题,在某些方面仍然不尽如人意。

如何通过构建数据治理机制,在数据的共享、交易、访问和使用等方面建立制度性安排,在最大限度发挥数据价值的同时,治理和避免数据垄断问题等负面后果,是学界关注的重点和热点之一。无论是从已有研究的进展来看,还是从当前社会的现实需求来看,对数据垄断问题的相关研究方兴未艾。本文认为,未来研究仍至少有以下三个方面可以进一步拓展:

第一,加强异质性分析,扩展和丰富数据垄断治理政策的影响及后果的相关结论,以便于精准施策。从当前已有文献来看,数据垄断治理的研究和分析仍然缺乏异质性维度。其一,国家间的异质性。已有研究主要聚焦于欧美发达经济体,特别是以欧盟的 GDPR 等政策为对象,展开了较为丰富的研究。但是发展中经济体与发达经济体在数字经济发展程度、用户隐私偏好等方面存在显著差异,相关数据治理政策可能具有不同的传导机制和后果。其二,行业间的异质性。尽管数字技术已经向各行业的各方面进行渗透,但不同行业的数字化程度不同,不同行业对于数据要素的依赖程度也不同。在不同行业之间,数据垄断问题的表现形式、严重程度、政策干预的必要性都需要进一步探讨。后续研究可以进一步考虑在模型中引入用户隐私偏好的差异、引入企业对数据依赖性的差异等,讨论相应的数据垄断治理政策对不同用户群体的福利的影响,以及对异质性企业的数据投资、价格和产量等决策的影响,提高数据垄断治理的机制设计与政策手段的准确性及有效性。

第二,加强数据垄断治理相关政策效果的评估,为政策的影响及作用提供实证证据的支撑。不论是数据要素在经济中的作用,还是数据垄断治理政策的影响,已有文献绝大多数都是基于理论模型进行考察和分析。对于数据政策的作用及其后果的研究,相应的经验证据仍较少,尤其是缺乏微观实证证据的支撑。一些实证研究文献大多使用调研数据、实验数据进行研究,这在一定程度上限制了研究结论的代表性和可推广性。后续的研究可以先以指标测度为基础,准确测度数据资产在企业资产中的占比情况、数字技术在企业中的运用情况等,体现和反映企业与数据相关的各类决策;然后运用相应的政策评估计量方法,分析数据垄断治理政策对企业的数据投资等相关决策的影响。

第三,加强对数据垄断治理政策的跨国溢出效应的研究,探索数据垄断治理的跨国合作方式。一方面,当前的诸多数字平台和数据垄断企业都是全球性公司,其运营的相关投入、知识产权、服务的用户对象等等往往都遍布全球不同国家,相应地,这些跨国企业也可以利用不同国家间的制度差异来规避垄断治理政策。因此,对于这些企业的监管和规制,需要超国家监管机构或跨国性的制度

安排才能有效实现数据垄断治理的目标。从当前来看,关于数据治理跨国合作的相关研究仍较少。后续研究可以全球性的平台企业为对象,研究相关数据政策如何影响其商业战略和全球业务分布,探索跨国合作数据治理的机构设置、制度安排、方式手段。另一方面,已有文献对于数据垄断治理政策的讨论主要集中于对本国企业和用户的影响,缺乏对政策跨国溢出效应的研究。后续研究可以进一步讨论一国的数据治理政策如何影响国际数据流动,在微观层面,研究相关政策引起的国际数据流动的变化如何影响跨国企业的相应决策;在宏观层面,可以研究国际数据流动的变化是否会像国际贸易的变化一样影响一国的经济表现。

参考文献:

- Acemoglu, D. et al.(2022), “Too much data: Prices and inefficiencies in data markets”, *American Economic Journal: Microeconomics*, 14(4):218–256.
- Acquisti, A. et al.(2016), “The economics of privacy”, *Journal of Economic Literature*, 54(2):442–492.
- Anderson, S.P. & J.S.Gans(2011), “Platform siphoning: Ad-avoidance and media content”, *American Economic Journal: Microeconomics*, 3(4):1–34.
- Athey, S. et al.(2017), “The digital privacy paradox: Small money, small costs, small talk”, NBER Working Paper, No.23488.
- Autor, D. et al.(2020), “The fall of the labor share and the rise of superstar firms”, *Quarterly Journal of Economics*, 135(2):645–709.
- Bai, J. et al.(2021), “Digital resilience: How work-from-home feasibility affects firm performance”, NBER Working Paper, No.28588.
- Bamberger, K.A. & O.Lobel(2017), “Platform market power”, *Berkeley Technology Law Journal*, 32(3): 1051–1092.
- Begenau, J. et al.(2018), “Big data in finance and the growth of large firms”, *Journal of Monetary Economics*, 97(C): 71–87.
- Benzell, S.G. et al.(2023), “How APIs create growth by inverting the firm”, *Management Science*, <https://doi.org/10.1287/mnsc.2023.4968>.
- Bergemann, D. & A.Bonatti(2015), “Selling cookies”, *American Economic Journal: Microeconomics*, 7(3):259–294.
- Bergemann, D. et al.(2015), “The limits of price discrimination”, *American Economic Review*, 105(3):921–957.
- Bergemann, D. et al.(2022), “The economics of social data”, *RAND Journal of Economics*, 53(2):263–296.
- Bourreau, M. et al.(2018), “Taxation of a digital monopoly platform”, *Journal of Public Economic Theory*, 20(1):40–51.
- Brynjolfsson, E. & K.McElheran(2016), “The rapid adoption of data-driven decision-making”, *American Economic Review*, 106(5):133–139.
- Calvano, E. & M.Polo(2021), “Market power, competition and innovation in digital markets: A survey”, *Information Economics and Policy*, 54, No.100853.
- Campbell, J. et al.(2015), “Privacy regulation and market structure”, *Journal of Economics & Management Strategy*, 24(1):47–73.
- Casadesus-Masanell, R. & A.Hervas-Drane(2015), “Competing with privacy”, *Management Science*, 61(1):229–246.
- Chen, L. et al.(2021), “The data privacy paradox and digital demand”, NBER Working Paper, No.28854.
- Chen, Z. et al.(2022), “Data-driven mergers and personalization”, *RAND Journal of Economics*, 53(1):3–31.
- Choe, C. et al.(2024), “Softening competition through unilateral sharing of customer data”, *Management Science*, 70(1):526–543.
- Choi, J.P. et al.(2019), “Privacy and personal data collection with information externalities”, *Journal of Public Economics*, 173(C):113–124.
- Cong, L.W. & S.Mayer(2023), “Antitrust, regulation, and user union in the era of digital platforms and big data”, NBER Working Paper, No.30881.
- Cong, L.W. et al.(2021), “Knowledge accumulation, privacy, and growth in a data economy”, *Management Science*, 67(10): 6480–6492.
- Cong, L.W. et al.(2022), “Endogenous growth under multiple uses of data”, *Journal of Economic Dynamics and Control*, 141, No.104395.

- Crouzet, N. & J.Eberly(2019), "Understanding weak capital investment: The role of market concentration and intangibles", NBER Working Paper, No.25869.
- Crouzet, N. & J.Eberly(2021), "Intangibles, markups, and the measurement of productivity growth", *Journal of Monetary Economics*, 124(S):92–109.
- Crouzet, N. et al.(2022), "A model of intangible capital", NBER Working Paper, No.30376.
- De Ridder, M.(2024), "Market power and innovation in the intangible economy", *American Economic Review*, 114(1):199–251.
- Dosis, A. & W.Sand-Zantman(2023), "The ownership of data", *Journal of Law, Economics, and Organization*, 39(3):615–641.
- Duch-Brown, N. et al.(2017), "The economics of ownership, access and trade in digital data", JRC Digital Economy Working Paper, No.2017–01.
- Economides, N. & I.Lianos(2021), "Restrictions on privacy and exploitation in the digital economy: A market failure perspective", *Journal of Competition Law & Economics*, 17(4):765–847.
- Eeckhout, J. & L.Veldkamp(2022), "Data and market power", NBER Working Paper, No.30022.
- Eisenmann, T. et al.(2011), "Platform envelopment", *Strategic Management Journal*, 32(12): 1270–1285.
- Elsaify, M. & S.Hasan(2020), "Some data on the market for data", SSRN Working Paper, No.3568817.
- Fainmesser, I.P. & A.Galeotti(2016), "Pricing network effects", *Review of Economic Studies*, 83(1):165–198.
- Fainmesser, I. P. & A. Galeotti(2020), "Pricing network effects: Competition", *American Economic Journal: Microeconomics*, 12(3):1–32.
- Fainmesser, I. P. et al.(2023), "Digital privacy", *Management Science*, 69(6):3157–3173.
- Fairfield, J.A.T. & C.Engel(2015), "Privacy as a public good", *Duke Law Journal*, 65(3):385–457.
- Farboodi, M. & L.Veldkamp(2022), "A model of the data economy", NBER Working Paper, No.28427.
- Farronato, C. et al.(2024), "Dog eat dog: Balancing network effects and differentiation in a digital platform merger", *Management Science*, 70(1):464–483.
- Gao, G. & L.M.Hitt(2012), "Information technology and trademarks: Implications for product variety", *Management Science*, 58(6):1211–1226.
- Goh, K.Y. et al.(2015), "Privacy and marketing externalities: Evidence from do not call", *Management Science*, 61(12):2982–3000.
- Goldfarb, A. & V.F.Que(2023), "The economics of digital privacy", *Annual Review of Economics*, 15(1):267–286.
- Graef, I. & J.Prüfer(2021), "Governance of data sharing: A law & economics proposal", *Research Policy*, 50(9), No.104330.
- Gul, F. & W.Pesendorfer(2001), "Temptation and self-control", *Econometrica*, 69(6):1403–1435.
- Hagiu, A. & J.Wright(2021), "Data-enabled learning, network effects and competitive advantage", *RAND Journal of Economics*, 54(4):638–667.
- Hann, I.H. et al.(2008), "Consumer privacy and marketing avoidance: A static model", *Management Science*, 54(6): 1094–1103.
- He, Z. et al.(2023), "Open banking: Credit market competition when borrowers own the data", *Journal of Financial Economics*, 147(2):449–474.
- Ichihashi, S.(2020), "Online privacy and information disclosure by consumers", *American Economic Review*, 110(2): 569–595.
- Ichihashi, S.(2021), "The economics of data externalities", *Journal of Economic Theory*, 196, No.105316.
- Jin, Y. & S.Vasserman(2021), "Buying data from consumers: The impact of monitoring programs in us auto insurance", NBER Working Paper, No.29096.
- Johnson, G.A. et al.(2023), "Privacy and market concentration: Intended and unintended consequences of the GDPR", *Management Science*, 69(10):5695–5721.
- Johnson, J.P.(2013), "Targeted advertising and advertising avoidance", *RAND Journal of Economics* 44(1):128–144.
- Jones, C.I. & C.Tonetti(2020), "Nonrivalry and the economics of data", *American Economic Review*, 110(9):2819–2858.

- Jullien, B. & W.Sand-Zantman(2021), “The economics of platforms: A theory guide for competition policy”, *Information Economics and Policy*, 54, No.100880.
- Ke, T.T. & K.Sudhir(2023), “Privacy rights and data security: GDPR and personal data markets”, *Management Science*, 69(8):4389—4412.
- Khan, L.M.(2017), “Amazon’s antitrust paradox”, *Yale Law Journal*, 126(3):710—805.
- Khan, L.M.(2018), “Sources of tech platform power”, *Georgetown Law Technology Review*, 2(2):325—334.
- Kind, H. J. et al.(2008), “Efficiency enhancing taxation in two-sided markets”, *Journal of Public Economics*, 92(5—6):1531—1539.
- Kind, H.J. et al.(2010), “Tax responses in platform industries”, *Oxford Economic Papers* 62(4):764—783.
- Kirpalani, R. & T.Philippon(2020), “Data sharing and market power with two-sided platforms”, NBER Working Paper, No.28023.
- Liu, Z. et al.(2020), “Data privacy and temptation”, NBER Working Paper, No.27653.
- MacCarthy, M.(2011), “New directions in privacy: Disclosure, unfairness and externalities”, *I/S: A Journal of Law and Policy for the Information Society*, 6(3):425—512
- McIntosh, D.(2019), “We need to talk about data: How digital monopolies arise and why they have power and influence”, *Journal of Technology Law & Policy*, 23(2): 185—213.
- Miklós-Thal, J. et al.(2023), “Digital hermits”, NBER Working Paper, No.30920.
- Montes, R. et al.(2019), “The value of personal information in online markets with endogenous privacy”, *Management Science*, 65(3):1342—1362.
- Müller, O. et al.(2018), “The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics”, *Journal of Management Information Systems*, 35(2):488—509.
- Parlour, C.A. et al.(2022), “When fintech competes for payment flows”, *Review of Financial Studies*, 35(11):4985—5024.
- Prüfer, J. & C. Schottmüller(2022), “Competing with big data”, *Journal of Industrial Economics* 69(4):967—1008.
- Rammer, C. et al.(2022), “Artificial intelligence and industrial innovation: Evidence from German firm-level data”, *Research Policy*, 51(7),No.104555.
- Reimers, I. & J.Waldfoegel(2023), “A framework for detection, measurement, and welfare analysis of platform bias”, NBER Working Paper, No.31766.
- Rochet, J.C. & J. Tirole(2003), “Platform competition in two-sided markets”, *Journal of the European Economic Association*, 1(4):990—1029.
- Sun, T. et al.(2023), “The value of personal data in internet commerce: A high-stakes field experiment on data regulation policy”, *Management Science*, <https://doi.org/10.1287/mnsc.2023.4828>.
- Tambe, P. et al.(2020), “Digital capital and superstar firms”, NBER Working Paper, No.28285.
- Tirole, J.(2023), “Competition and the industrial challenge for the digital age”, *Annual Review of Economics*, 15(1): 573—605.
- Van Dijck, J. et al.(2019), “Reframing platform power”, *Internet Policy Review*, 8(2):1—18.
- Van Loo, R. & N.Aggarwal(2023), “Amazon’s pricing paradox”, *Harvard Journal of Law & Technology*, https://scholarship.law.bu.edu/faculty_scholarship/3645.
- Van Loo, R.(2017), “Rise of the digital regulator”, *Duke Law Journal*, 66(6):1267—1330.
- Veldkamp, L. & C.Chung(2019), “Data and the aggregate economy”, *Journal of Economic Literature*, 57(1):3—43.
- Veldkamp, L.(2023), “Valuing data as an asset”, *Review of Finance*, 27(5):1545—1562.
- Weyl, E.(2010), “A price theory of multi-sided platforms”, *American Economic Review*, 100(4):1642—1672.
- Wilbur, K.C.(2008), “A two-sided, empirical model of television advertising and viewing markets”, *Marketing Science* 27(3):356—378.
- Wu, L. et al.(2020), “Data analytics, innovation, and firm productivity”, *Management Science*, 66(5):2017—2039.
- Zhu, F. & Q.Liu(2018), “Competing with complementors: An empirical look at Amazon.com”, *Strategic Management Journal*, 39(10):2618—2642.

Research Progress on Data Monopoly

SHEN Kunrong and LIN Jianwei
(Nanjing University, Nanjing, China)

Summary: Strengthening platform antitrust and data governance has become a global consensus. In particular, how to maximize the value of data and manage the problem of data monopoly through the establishment and improvement of data systems is one of the hot spots in the academic circle. This paper systematically reviews and summarizes the research progress on data monopoly in existing foreign literature. Among them, the paper focuses on the governance methods of data monopoly and looks forward to the research directions in the future.

In the first part, this paper discusses the relationship between data and market monopoly and discusses the existing literature's standards for enterprises to use data to implement monopoly behavior. Data monopoly includes two aspects: one is the monopoly of data resources by enterprises; the other is that enterprises take advantage of data resources and use data as a tool to maintain and strengthen their monopoly position in the market and obtain monopoly profits. When data monopoly has a negative impact on social welfare, data monopoly becomes a problem that needs to be addressed.

In the second part, this paper introduces the causes and the negative effects of data monopoly. Data can cause or exacerbate market monopoly problems for three reasons: data can become a barrier to market entry, strengthen the advantages of monopolies, and become a tool for monopolies to carry out monopolistic behavior. The negative effects of data monopoly can be summarized from two aspects: Firstly, data monopoly will inhibit market competition and lead to the decline of economic efficiency. Secondly, enterprises rely on data monopoly to gain excess profits, which will occupy consumer surplus and ultimately damage user welfare, which also involves consumer privacy issues.

In the third part, this paper introduces the role of data systems in anti-monopoly. Specifically, this paper summarizes the existing research from three aspects: data sharing, limitation of user data use, and data portability. (1) Data sharing can reduce the inequality in the possession of data resources and alleviate the problem of data monopoly. However, data sharing has the potential to reduce the incentive for enterprises to make data investments, which will trigger new efficiency losses. (2) Restricting the use of user data by enterprises can reduce monopoly behaviors such as price discrimination by enterprises, thus improving user welfare. At the same time, it can reduce the negative utility caused by user privacy issues. On the other hand, restricting the use of user data must also consider the loss of social welfare caused by reducing the use of data. Moreover, because large and small firms differ in their ability to bear regulatory costs, policies can backfire, exacerbating imbalances among firms. (3) The portability of personal data and the transferability of user data between platforms are currently one of the key focuses of data systems implemented in major countries around the world. From the conclusions of existing research, it seems that this is the most promising direction to alleviate the problem of data monopoly. However, it is necessary to further consider the effects of privacy paradox, data externalities, and strategic behaviors of monopolies on policy effects.

Finally, this paper introduces the future research direction of data monopoly governance. Whether from the academic research progress or the current social reality, the research on data monopoly is in the rising period. There are at least three directions for future research. First, consider the heterogeneous effects of data monopoly governance policies between countries or industries; second, strengthen the evaluation of the effects of data monopoly governance policies to provide more empirical evidence for the consequences of policies; third, consider the transnational spillover effects of data governance policies and explore transnational cooperation ways of data governance. It is hoped that this paper can provide useful insights to accelerate the construction of the data system and improve the related research and policy practice of data governance.

Keywords: Data Monopoly; Data Sharing; Data Rights; Data Governance

JEL Classification: D12, D21, O33

(责任编辑:木子)

(校对:何伟)