

# 数据收集条款、个人信息保护 与消费者福利\*

刘征驰 叶宇阳 聂辉华

**摘要:**如何协调数据收集与个人信息保护之间的关系已成为数字经济时代的重大议题。本文从消费者福利的角度,讨论是否应该限制企业强制收集消费者数据。通过建模刻画企业与消费者在不同类型数据收集条款下的收益与成本,本文发现当不受个人信息保护法律规制时,企业将根据数据边际私人收益和边际公共收益来选择数据收集条款。而福利分析表明,相比于自愿收集条款,虽然强制收集条款意味着消费者个人权利受限,但可能通过提高数据公共收益来增加消费者福利。随后本文讨论个人信息保护法律对强制收集条款的规制如何影响消费者福利,并发现实施力度较弱的规制可以在某些数据收集场景中实现消费者福利帕累托改进,而过强的规制力度无助于实现消费者福利的帕累托改进。本文的政策启示是,应当继续推进数据要素的市场化改革,并适当放松对数据收集条款的规制力度。

**关键词:**数据收集条款 自愿收集 强制收集 个人信息保护 消费者福利

## 一、引言

数据已成为数字经济时代的关键资源,被视作与劳动、资本、土地等相提并论的生产要素。数据收集涉及对个人信息的利用<sup>①</sup>,因此规范对消费者数据的收集行为以保障消费者福利已成为理论与政策的关注重点(崔等,2019;多西斯、桑德-赞特曼,2023;阿西莫格鲁等,2022;阿西莫格鲁,2024)。在常见的个人数据收集场景中,数据收集者(企业)可能采用的数据收集条款一般分为两类。第一类是自愿收集条款:如果消费者拒绝数据收集,企业不会限制消费者使用服务,企业甚至可能对同意数据收集的消费者给予补贴(邵小快、郑捷,2024)。例如,在外卖应用程序(App)中消费者可以自愿选择是否提交实时位置数据。第二类是强制收集条款:如果消费者拒绝数据收集,企业会限制乃至完全禁止消费者使用服务;只有消费者同意数据收集,才被准许享受完整的服务(何佳等,2022)。例如,导航App禁止拒绝隐私政策的消费者使用查询出行路线等功能。强制收集条款往往被视为数据收集企业对消费者的权利侵犯与福利攫取,许多法律和政策都对其进行规制。例如,欧盟的《通用数据保护条例》(简称GDPR)强调数据处理者征得对数据收集的同意时,需要保证该同意是自由做出的;《中华人民共和国个人信息保护法》(简称《个保法》)第十六条规定“禁止强制收集非必要个人信息,个人信息处理者不得以个人不同意授权使用其个人信息或者撤回同意为由,拒绝提供产品或者服务;处理个人信息属于提供产品或者服务所必需的除外”;《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”)也提出不得采用强制同意等方式过度收集个人信息<sup>②</sup>。因此一个值得讨论的问题是:如果法律限制企业在数据收集时使用强制收集条款,能否达到个人信息保护法律提升消费者福利的初衷?

从已有的司法实践来看,对《个保法》第十六条中“非必要”的解释为:如果企业在不收集某消费者数据的情况下,仍可以为该消费者提供服务,则数据是非必要的<sup>③</sup>。这种解释意味着当消费者提供的数据无法直

收稿时间:2025-2-17;反馈外审意见时间:2025-5-7、2025-6-19、2025-7-28;拟录用时间:2025-11-25。

\*本研究得到国家自然科学基金面上项目《三权分置视阈下数据流通链交易治理机制研究》(72471086)、《数字经济对企业本质、边界和内部组织的重塑》(72273144)、国家社会科学基金万人计划领军人才项目(22VRC157)以及北京市高校卓越青年科学家计划项目(JWZQ20240202001)的资助。聂辉华为本文通讯作者。

接给本人带来的足够的收益时,企业不能采用强制收集条款。但是消费者数据积累可以改善服务质量(徐宗本等,2014),提升匹配效率(罗曼纽克、斯莫林,2019),促进创新决策(林奇,2008),从而给消费者带来非竞争性的公共收益。如果限制使用强制收集条款,虽然可以保障消费者拒绝数据收集时正常使用企业服务的权利,但也可能导致消费者获得“搭便车”的机会,降低数据收集的整体效率,进而损害消费者的整体福利。

为了更深入地分析这一问题,本文构建一个垄断企业和若干异质性消费者参与的二阶段博弈模型。企业收集消费者数据,并利用数据为消费者提供服务。当消费者使用服务时,企业可以从中获利。数据收集给消费者带来收益和成本。数据收益体现在服务质量的提升,这包括两个方面。其一是数据直接给数据提供者本人带来的收益,称之为数据私人收益。例如视频 App 根据某个消费者的浏览记录推荐他(她)可能感兴趣的视频。其二是数据积累带来的非竞争性收益,称之为数据公共收益,可以由所有消费者共享。例如视频 App 综合所有消费者的浏览记录,筛选出高质量的视频并推送。在大数据与人工智能时代,数据积累带来的公共收益愈发凸显。此外,数据收集也会给消费者带来成本,这不仅包括消费者的隐私损失,也包括消费者因打开权限、填写信息、上传照片等活动耗费的成本,而且以上成本依赖于消费者的主观感知,因此不同的消费者在数据收集成本方面存在差异<sup>④</sup>。

作为分析的基准,本文首先讨论不存在个人信息保护的情形。企业可以在两种数据收集条款中进行选择,并面临不同的成本:若选择自愿收集条款,企业必须支出补贴才能提高数据收集总量;若选择强制收集条款,企业会限制拒绝数据收集的消费者使用完整的服务,因此不用支出补贴也能迫使更多消费者同意数据收集,但依然存在部分高成本消费者为了避免数据收集而放弃使用完整的服务,从而导致企业丧失一部分收益。本文发现数据边际私人收益与边际公共收益会影响企业对数据收集条款的选择。具体而言,给定数据边际公共收益,当数据边际私人收益高于某一阈值时,企业会选择强制收集条款;当数据边际私人收益低于某一阈值时,企业会选择自愿收集条款。此外,当企业选择自愿收集条款时,若数据边际公共收益较高,企业会提供补贴。

接下来通过分析消费者福利,本文发现在某些情形下,强制收集条款下的消费者整体福利高于自愿收集条款的情形。由于数据积累可以增加数据公共收益,而强制收集条款作为减少消费者“搭便车”的机制,其效果在某些情形下优于自愿收集条款,因此可能反而有利于消费者整体福利的增进。此外,消费者福利异质性分析表明,当数据边际私人收益足够高,或数据边际公共收益较低时,相较于自愿收集条款,强制收集条款可以增加低成本消费者的福利。据此可以推断,如果禁止企业采用强制收集条款,可能会损害低成本消费者的福利,这与其他要素市场上规制政策的作用类似,例如最低工资标准可能损害低成本劳动者(即愿意接受低工资的低技能劳动者)的福利<sup>⑤</sup>。

最后,本文讨论个人信息保护法律的一般情形,即规制力度介于完全禁止强制收集条款和完全放任强制收集条款之间时,企业的数据收集条款选择和消费者福利如何变化。本文用“拒绝数据收集时消费者可以使用企业服务的最低比例”,刻画个人信息保护法律对强制收集条款的规制力度,这一最低比例由立法及司法活动外生决定。企业若采用强制数据收集条款,必须为拒绝数据收集的消费者提供不低于该比例的服务。因此最低比例越高,即对强制收集条款的规制力度越大。本文发现,相对于实施个人信息保护前的情况,无论如何调整规制力度,都存在某些数据收集场景,消费者整体福利会因个人信息保护而减少。此外,为了更贴合个人信息保护法律的立法初衷,本文将消费者福利帕累托改进作为效率目标进行分析,并发现:实施规制力度较低的个人信息保护,即准许企业采用限制使用的强制收集条款(消费者拒绝数据收集也能使用部分服务),在某些数据服务场景中有助于消费者福利帕累托改进,而继续提高规制力度无法实现进一步的消费者福利帕累托改进。

本文剩余部分的安排如下:第二部分对相关文献进行综述,阐述本文的边际贡献;第三部分介绍模型的基本设定;第四部分研究不存在个人信息保护时企业对数据收集条款的选择;第五部分比较不同数据收集条款类型下,消费者福利如何变化;第六部分讨论个人信息保护如何改变企业数据收集条款选择并影响消费者福

利;第七部分对全文进行总结并提出相应政策建议。

## 二、文献综述

数字经济最突出的特征,表现为规模庞大的数据被收集、存储与分析,因而个人数据收集与保护之间的权衡已经成为最重要的经济议题之一(阿奎斯蒂等,2016)。从宏观视角看,数据收集权力作为数据产权的一部分(刘涛雄等,2023),其归属会影响经济增长(琼斯、托内蒂,2020;丛等,2021)。而从微观视角看,数据收集的影响体现在多个方面,例如企业可能将消费者数据用于精准识别消费者类别从而进行价格歧视(科尼策等,2012;蒙特斯等,2019;市桥,2020;李三希等,2021;宾皮基斯等,2024),帮助广告商向消费者精准推送广告(戈德法布、塔克,2011;沈、米格尔,2018;汪敏达等,2022),提高买卖双方的匹配效率(罗曼纽克、斯莫林,2019)等等。

关于企业对数据收集条款的选择问题,现有文献往往将数据收集条款类型视为外生,只考察企业在单一数据收集条款类型下的选择。部分文献主要考察自愿收集条款的情形,即企业向同意数据收集的消费者支付补贴,例如阿西莫格鲁等(2022)发现随着企业对消费者类型的先验信念与消费者真实类型之间的偏离程度增加,补贴额度上升;而随着数据对消费者的负外部性增加,补贴额度下降。市桥(2021)发现在数据对消费者有害的前提下,若不同消费者之间数据完全互补,则补贴额度为正;若消费者数据完全替代,则补贴额度为0。邵小快和郑捷(2024)考察双寡头竞争场景下的数据收集条款选择,发现当让所有客户都不披露隐私的协议和区别对待式隐私协议均可选时,后者严格优于前者。还有一些文献只考虑强制收集条款的情形,例如布洛赫和德芒热(2018)发现在强制收集条款下,垄断企业的数据收集量通常会超过消费者最优的水平。崔等(2019)、范梅塞尔等(2023)和李三希等(2023b)用数据收集量或数据收集强度的方式刻画企业数据收集条款,虽然具体设定存在差异,但都可视作强制收集条款的情形,即消费者只能接受企业规定的的数据收集量或数据收集强度,除非消费者放弃使用企业提供的服务。上述研究发现垄断企业的数据收集量通常会超过消费者最优的水平(崔等,2019;范梅塞尔等,2023),而市场竞争会降低数据收集强度(李三希等,2023b)。

不同于已有文献,本文同时考虑两类数据收集条款,并将企业的数据收集条款类型纳入企业决策。不同类型的数据收集条款将导致不同的成本:若选择自愿收集条款,企业必须支出补贴才能提高数据收集量,但不会“挤出”消费者;若选择强制收集条款,企业不用支出补贴也能迫使更多消费者同意数据收集,但会“挤出”部分高成本消费者,从而企业会丧失一部分收益。如果忽视这种成本差异,不仅无法解释现实中企业对数据收集条款类型的选择,还可能导致福利分析的结果出现偏差。

另一类与本文密切相关的文献是关于个人信息保护法律规制影响消费者福利的研究。从影响机制看,此类文献可以细分为以下3个子类。第一,法律规制影响企业的数据收集策略进而影响消费者福利。例如,布洛赫和德芒热(2018)发现,当数据收益较高时,若赋予消费者选择退出(opt-out)的权利,反而可能促使企业提高数据收集水平以补偿某些用户拒绝提交数据而造成的损失,导致低隐私敏感用户因数据利用水平提高而受损,而高隐私敏感用户受益。第二,法律规制通过改变企业的产品定价策略来影响消费者福利。例如,李三希等(2021)发现在垂直差异化的双寡头竞争市场上,无个人信息保护下消费者总福利大于自愿性个人信息保护下的消费者总福利。由于厂商预期支付意愿较高的消费者将隐藏个人信息进入匿名市场,因此将相应提高在匿名市场的定价,进而导致所有新消费者福利较无个人信息保护时要差,且新消费者的福利损失大于老消费者的福利增加部分。蔡勇等(2025)将消费者隐私关切和厂商数据利用能力差异纳入分析,发现在水平差异化的双寡头竞争市场上,当新消费者规模较低且厂商数据利用能力较低时,与无数据保护相比,自愿性数据保护下部分老消费者选择隐藏数据,降低隐私关切,导致自愿性数据保护下的消费者总体剩余更高。张剑虎和林平(2023)证明“单独同意”法律条款虽然会增加保护消费者数据的企业数量及其市场份额,但由于企业能够将“单独同意”的合规成本转嫁给消费者,从而导致产品价格上升、消费者福利降低。第三,法律规制通过改变企业的产品质量及定价策略来影响消费者福利。例如,孔蒂和雷韦尔贝里(2021)发现在消费者数据与产品质量之间存在较强的互补性的前提下,隐私监管(opt-in)会促使企业通过提高产品质量激励消费者分享数据,于是



分享数据的消费者因高质量产品受益,但需支付更高价格;而匿名消费者因统一价格上涨可能受损。总体来看,仅当产品质量显著提升时,消费者剩余可能增加。

与上述研究不同,本文在消费者层面刻画了数据这一新型生产要素的非竞争性。基于这一假设,本文中法律规制通过两条路径影响消费者福利。首先,法律规制通过改变企业对数据收集条款的选择,影响消费者的数据收集权益(例如消费者拒绝数据收集时仍可以使用企业服务),进而直接影响消费者福利;其次,企业数据收集条款会影响数据收集总量,由于数据具有非竞争性,数据收集总量会影响每个消费者可获取的服务质量,最终影响消费者福利。

此外,现有研究只对比个人信息保护的两种特殊情形:不存在个人信息保护的情形与个人信息保护法律完全禁止企业采用强制收集条款的情形。基于这种对比,现有研究发现实施个人信息保护会损害部分消费者的福利。而本文用连续变量刻画个人信息保护对强制收集条款的规制力度,发现在某些数据收集场景中,对强制收集条款实施力度较弱的规制有助于实现消费者福利帕累托改进,而过强的规制力度则无法进一步实现消费者福利帕累托改进。这一结论表明,通过控制规制力度并推进数据分类治理,个人信息保护可以同时提升不同类型消费者的福利。

在经验研究方面,目前的研究主要围绕欧盟的隐私保护政策(如欧盟的《通用数据保护条例》)展开,且大多考察其对企业层面的影响。首先,隐私保护政策减少了企业的数据收集与数据处理量,弱化其“数据密集型”的程度(阿里多尔等,2023;德米雷尔等,2024)。其次,隐私保护政策降低了算法服务精度。以广告投放算法为例,戈德法布和塔克(2011)发现欧盟的隐私保护政策削弱了企业识别消费者类型的能力。最后,隐私保护政策削弱了中小数字企业的竞争力。波伊克特等(2022)针对网络技术服务市场的研究发现,尽管欧盟的《通用数据保护条例》实施后所有企业都遭受了损失,但巨头企业损失较小,并且其市场份额显著增加。以上研究的结论表明,个人信息保护法律可以改变企业数据收集行为并影响企业绩效。企业行为的变化可能对消费者福利产生影响,但尚未得到具体的经验证据。

本文的边际创新体现在以下3个方面。

第一,比较不同规制力度下,个人信息保护法律对消费者福利的影响。本文发现在某些数据收集场景中,较弱的规制力度可以实现消费者福利帕累托改进。这一结论为放松个人信息保护规制以及数据分类治理提供了理论指引,也为法学领域适当放宽最小必要原则的观点提供了经济学基础(丁晓东,2019;武腾,2021)。

第二,在消费者层面刻画数据的非竞争性,揭示企业在数据收集中的作用,进而讨论数据收集权利问题。现有文献通常从企业层面讨论数据的非竞争性及其与数据权力配置的关系(琼斯、托内蒂,2020;李三希等,2023a),但较少从消费者层面讨论这一问题。消费者数据的积累可以为消费者带来非竞争性的公共收益,因此厌恶数据收集的理性消费者具有“搭便车”的倾向,可能导致数据收集缺乏效率。企业可以利用补贴或强制的手段提高数据收集量,从而可能提升消费者福利。这有助于进一步理解数据要素背景下企业组织的作用,也为研究数据确权问题提供了新思路。

第三,将企业的数据收集条款类型选择纳入企业决策,为现实中企业差异化的数据收集条款选择提供了一个解释,拓展了关于企业数据收集的经济学研究。现有文献通常将数据收集条款类型视为外生,并未解释企业如何在自愿收集条款和强制收集条款之间做出选择。本文则发现企业会根据数据的边际私人收益和边际公共收益的相对高低来选择数据收集条款类型。

### 三、模型基本设定

本部分首先介绍消费者效用函数与企业收益函数的基准情形,然后分别介绍上述两种函数在自愿收集条款和强制收集条款下的变化,并引入模型的其他假设。

#### (一)消费者效用

本文构建模型刻画一个由垄断企业<sup>®</sup>和若干消费者参与的数据收集过程。消费者具有异质性,其类型由

参数  $c \in [0, 1]$  刻画,  $c$  表示数据收集给消费者带来的边际成本, 其分布的密度函数为  $f(c)$ 。不失一般性, 消费者人数被标准化为 1 单位, 并假设  $c$  在  $[0, 1]$  内均匀分布, 即:

$$f(c) = 1, c \in [0, 1] \quad (1)$$

若某个  $c$  类型的消费者使用企业所提供的数据服务, 则其效用函数为<sup>⑦</sup>:

$$u_c(d_c) = \alpha d_c + \beta \bar{d} - c d_c \quad (2)$$

消费者效用函数的第一部分  $\alpha d_c$  刻画了某个消费者的数据给其本人带来的收益, 即数据私人收益。例如某消费者若同意视频 App 收集浏览记录, 那么推荐算法会为其推荐感兴趣的视频; 又例如只有消费者向导航程序提交个人位置信息与目的地, 导航程序才能为消费者规划路径。 $\alpha$  为外生的常数, 刻画了边际私人收益的高低, 令  $\alpha > 0$ <sup>⑧</sup>。 $d_c$  是企业对某个  $c$  类型的消费者的数据收集量。由于在现实的数据收集情形中, 消费者面临的往往是“同意或拒绝”的二元选择, 因此与某些文献类似(市桥, 2021; 孔蒂、雷韦尔贝里, 2021; 阿西莫格鲁等, 2022), 本文将数据收集量  $d_c$  定义为 0-1 变量。若消费者同意企业收集个人数据, 则  $d_c = 1$ ; 若消费者拒绝企业收集个人数据, 则  $d_c = 0$ 。从技术角度看, 若某个消费者拒绝数据收集, 则无法获取该部分收益, 即  $d_c = 0$  时该部分收益为 0。

消费者效用函数的第二部分  $\beta \bar{d}$  刻画了数据公共收益(专指数据给消费者带来的非竞争性公共收益, 可以由全体消费者共享), 其中  $\beta$  为外生的常数且  $\beta > 0$ , 刻画了边际公共收益的高低。 $\bar{d} = \int_0^1 d_c f(c) dc$  表示该企业对所有消费者收集数据的总量。数据公共收益与数据总量  $\bar{d}$  正相关, 这是因为数据规模的扩大有助于改善服务质量(徐宗本等, 2014), 提升匹配效率(罗曼纽克、斯莫林, 2019), 促进创新决策(林奇, 2008)。例如, 某用户使用导航程序寻找最优驾驶线路时, 其他用户的出行数据有助于为其规避拥堵线路; 在网购时, 众多消费者提交的商品评价信息有利于减少信息不对称; 新能源汽车企业收集的用户驾驶信息有助于提升其智能驾驶性能, 从而提高每个驾驶者的体验。数据公共收益是非竞争性的, 可以由所有消费者共享<sup>⑨</sup>。从技术角度看, 即使某个消费者拒绝数据收集, 也不影响其获取该部分收益。

消费者效用函数的第三部分  $-c d_c$  刻画了数据收集给消费者带来的成本, 这不仅包括消费者的隐私损失, 也包括消费者因填写信息、提交评论、上传照片等活动耗费的成本。 $c$  越大则该消费者越厌恶数据收集。当然, 若某个消费者拒绝数据收集( $d_c = 0$ ), 则该消费者付出的成本为 0。由于消费者成本  $c d_c \in [0, 1]$ , 若  $\alpha + \beta \geq 1$ , 则意味着数据边际收益非常大, 企业采用强制收集条款不会“挤出”任何消费者, 最终的均衡结果是企业必定会采用强制收集条款。因此假设  $\alpha + \beta < 1$  以排除这种极端情形。

## (二) 企业收益

当消费者使用数据服务时, 某个  $c$  类型消费者给企业带来的收益为:

$$r_c = \lambda (\alpha d_c + \beta \bar{d}) \quad (3)$$

该函数旨在刻画企业作为基于数据的服务提供者, 其收益与消费者效用存在如下关系: 对于某个消费者而言, 基于数据的服务越能满足其需求, 则该消费者越有可能为企业带来收益。例如, 对于以广告收入为主的视频网站, 若能基于数据识别消费者偏好, 就能向消费者更精准地推送优质视频, 使消费者获取更好的观看体验; 广告也可以被更精准地投放, 于是广告品牌商会向视频网站支付更多广告费(按点击次数或转化率收费的广告计费方式就体现了这一特点)<sup>⑩</sup>。

为了简化, 假设  $\lambda$  外生。当  $\lambda \geq 1$ , 意味着企业从每个消费者获取的收益非常大, 因此企业不愿意用强制收集条款挤出消费者, 最终的均衡结果是企业必定会采用自愿收集条款。因此, 令  $0 < \lambda < 1$  以排除这种极端情形。

## (三) 数据收集条款

企业有两类数据收集条款可以选择, 在每种数据收集条款下, 消费者效用和企业收益会与式(2)及式(3)的基准情形存在差异。

第一类是自愿收集条款, 用符号  $v$  标记: 如果消费者同意收集个人数据, 企业会给予其额度为  $p$  的补贴( $p \geq$

0)<sup>⑩</sup>;如果消费者拒绝收集个人数据,企业不会限制消费者使用企业服务,但不会给消费者提供补贴。借鉴阿西莫格鲁等(2022)的思路,本文假定自愿收集条款下的消费者效用等于消费者的数据收益加上其获得的补贴。因此在这一类条款下,若某个 $c$ 类型消费者同意数据收集( $d_c=1$ ),则其效用为 $u_c^v=\alpha+\beta\bar{d}+p-c$ ;若某个 $c$ 类型消费者拒绝数据收集( $d_c=0$ ),则其效用为 $u_c^v=\beta\bar{d}$ 。

因此在自愿收集条款下,某个 $c$ 类型消费者的效用函数可以表示为:

$$u_c^v(d_c)=\alpha d_c+\beta\bar{d}+pd_c-cd_c \quad (4)$$

结合式(3)的企业收益函数,企业的净收益为<sup>⑪</sup>:

$$\begin{aligned} \Pi^v = \int_0^{\hat{c}} r_c f(c) dc - \int_0^{\hat{c}} pf(c) dc = & \underbrace{\int_0^{\hat{c}} \lambda(\alpha + \beta\bar{d}) f(c) dc}_{\text{同意数据收集的消费者给企业带来的收益}} \\ & + \underbrace{\int_{\hat{c}}^1 \lambda\beta\bar{d} f(c) dc}_{\text{拒绝数据收集的消费者给企业带来的收益}} - \underbrace{\int_0^{\hat{c}} pf(c) dc}_{\text{补贴给企业带来的成本}} \end{aligned} \quad (5)$$

其中, $\hat{c}$ 代表自愿收集条款下对数据收集与否无差异的消费者。

第二类是强制收集条款,用符号 $m$ 标记:如果消费者同意收集个人数据,消费者可以正常使用企业服务;如果消费者拒绝收集个人数据,企业将禁止消费者使用部分甚至全部企业服务。用参数 $\theta$ ( $0 \leq \theta \leq 1$ )表示消费者在拒绝数据收集的情形下可以使用服务的比例。在这一类条款下,若某个 $c$ 类型消费者同意数据收集( $d_c=1$ ),则其效用为 $u_c^m=\alpha+\beta\bar{d}-c$ ;若某个 $c$ 类型消费者拒绝数据收集( $d_c=0$ ),则其效用为 $u_c^m=\theta\beta\bar{d}$ 。根据 $\theta$ 的取值,可以进一步划分强制收集条款的类型:当 $\theta=0$ 时,条款类型为完全禁止的强制收集条款; $0 < \theta < 1$ 时,条款类型为限制使用的强制收集条款。特别地,当 $\theta=1$ 时,拒绝数据收集的消费者不会被限制使用企业服务,因此相当于补贴为0的自愿收集条款。

因此在强制收集条款下,某个 $c$ 类型消费者的效用函数可以表示为:

$$u_c^m(d_c)=\alpha d_c+\beta\bar{d}-(1-\theta)(1-d_c)\beta\bar{d}-cd_c \quad (6)$$

当消费者拒绝数据收集时,消费者只能使用比例为 $\theta$ 的服务,企业从这部分消费者获得的收益也相应减少。因此企业的净收益为:

$$\begin{aligned} \Pi^m = \int_0^{\hat{c}^m} r_c f(c) dc = & \underbrace{\int_0^{\hat{c}^m} \lambda(\alpha + \beta\bar{d}) f(c) dc}_{\text{同意数据收集的消费者给企业带来的收益}} + \underbrace{\int_{\hat{c}^m}^1 \lambda\beta\bar{d} f(c) dc - \int_{\hat{c}^m}^1 \lambda(1-\theta)\beta\bar{d} f(c) dc}_{\text{拒绝数据收集的消费者给企业带来的收益}} \end{aligned} \quad (7)$$

其中, $\hat{c}^m$ 代表强制收集条款下对数据收集与否无差异的消费者。

根据等式(7),在强制收集条款下,企业不用支出补贴,但会因限制拒绝数据收集的消费者使用服务而损失部分收益( $-\int_{\hat{c}^m}^1 \lambda(1-\theta)\beta\bar{d} f(c) dc$ )。

假设数据边际私人收益( $\alpha$ )、边际公共收益( $\beta$ )、消费者成本分布密度函数( $f(c)$ )是企业 and 所有消费者的共同知识,但具体某个消费者的成本只为其本人所知,而不被企业和其他消费者所知。此外,每个消费者都能够完全观测并理解企业的数据收集条款。模型时序如下:第一阶段,企业选择数据收集条款:若选择自愿收集条款,企业会设置 $p$ 的值;若选择强制收集条款,企业会设置 $\theta$ 的值;第二阶段,各个消费者独立做出决策,选择是否同意企业收集个人数据,即选择 $d_c$ 的取值。企业和消费者进而分别获取净收益和效用。

#### 四、不存在个人信息保护时企业对数据收集条款的选择

当不存在个人信息保护时,企业可以任意选择数据收集条款类型,并根据净收益最大化来决定数据收集条款。本文首先分别研究企业在自愿收集条款和强制收集条款下的最优设定:在自愿收集条款下,企业设定补贴额度 $p$ ;在强制收集条款下,企业设定消费者在拒绝数据收集时所能享受的服务比例 $\theta$ 。在此基础上,本文比较企业在两种条款下的净收益,可以得到企业净收益最大化的数据收集条款。

##### (一)自愿收集

自愿收集条款下,当且仅当 $u_c^v(d_c=1) > u_c^v(d_c=0)$ 时, $c$ 类型消费者会同意企业收集个人数据。于是无差异

消费者的成本为:

$$\hat{c}^* = \alpha + p \quad (8)$$

由于  $c \in [0, 1]$ ,  $p = 1 - \alpha$  即可保证所有消费者同意数据收集, 因此对于企业而言,  $p$  的有效取值区间是  $[0, 1 - \alpha]$ 。于是企业收集的数据总量为:

$$\bar{d}^* = \int_0^{\hat{c}^*} f(c) dc = \alpha + p \quad (9)$$

将式(8)和式(9)代入式(5)可得企业的净收益为:

$$\Pi^* = (p + \alpha)(\alpha + \beta)\lambda - (p + \alpha)p \quad (10)$$

根据式(10)可知企业增加补贴  $p$  的边际净收益为:

$$\frac{\partial \Pi^*}{\partial p} = (\alpha + \beta)\lambda - (2p + \alpha) \quad (11)$$

观察式(11)可知, 企业增加补贴  $p$  的边际净收益随  $p$  增加而递减。满足净收益最大化的补贴额度存在两种不同的情形: 当  $0 < \beta \leq (1 - \lambda)\alpha/\lambda$  时,  $\frac{\partial \Pi^*}{\partial p}(p = 0) = (\alpha + \beta)\lambda - \alpha \leq 0$ , 因此对于任意  $p > 0$  都有  $\frac{\partial \Pi^*}{\partial p} < 0$ , 于是满足净收益最大化的补贴额度为 0; 当  $(1 - \lambda)\alpha/\lambda < \beta < 1$  时,  $\frac{\partial \Pi^*}{\partial p}(p = 0) > 0$ , 且由于  $\alpha + \beta < 1$  以及  $\lambda < 1$ , 可知  $\frac{\partial \Pi^*}{\partial p}(p = 1 - \alpha) = (\alpha + \beta)\lambda + \alpha - 2 < 0$ , 因此必存在唯一的  $p^* \in [0, 1 - \alpha]$  满足  $\frac{\partial \Pi^*}{\partial p}(p = p^*) = 0$  且  $\frac{\partial^2 \Pi^*}{\partial p^2}(p = p^*) < 0$ 。因此满足净收益最大化的补贴额度  $p^*$  为:

$$p^* = \begin{cases} 0, & 0 < \beta \leq \frac{(1 - \lambda)\alpha}{\lambda} \\ \frac{1}{2}(\alpha(\lambda - 1) + \beta\lambda), & \frac{(1 - \lambda)\alpha}{\lambda} < \beta < 1 \end{cases} \quad (12)$$

将式(12)代入式(9), 可知当企业采用自愿收集条款时, 数据收集总量为:

$$\bar{d}^* = \begin{cases} \alpha, & 0 < \beta \leq \frac{(1 - \lambda)\alpha}{\lambda} \\ \frac{1}{2}(\alpha(\lambda + 1) + \beta\lambda), & \frac{(1 - \lambda)\alpha}{\lambda} < \beta < 1 \end{cases} \quad (13)$$

将式(12)代入式(10), 可知当企业采用自愿收集条款时, 企业净收益为:

$$\Pi^* = \begin{cases} \lambda\alpha(\alpha + \beta), & 0 < \beta < \frac{(1 - \lambda)\alpha}{\lambda} \\ \frac{1}{4}(\alpha + (\alpha + \beta)\lambda)^2, & \frac{(1 - \lambda)\alpha}{\lambda} < \beta < 1 \end{cases} \quad (14)$$

将消费者剩余定义为在区间  $[0, 1]$  内消费者效用函数的定积分, 其经济意义为消费者从数据服务中获取的收益加上获取的补贴, 再减去消费者在数据收集时付出的成本。于是消费者剩余为:

$$CS^* = \int_0^1 u_c^* f(c) dc = \begin{cases} \frac{1}{2}\alpha(\alpha + 2\beta), & 0 < \beta < \frac{(1 - \lambda)\alpha}{\lambda} \\ \frac{1}{8}(\alpha + (\alpha + \beta)\lambda)(\alpha + 4\beta + (\alpha + \beta)\lambda), & \frac{(1 - \lambda)\alpha}{\lambda} < \beta < 1 \end{cases} \quad (15)$$

根据以上分析, 当企业采用自愿收集条款时, 可以得到以下引理<sup>⑧</sup>。

引理 1: 若企业采用自愿收集条款, 企业会根据数据边际公共收益的相对大小决定是否给予消费者补贴: 当数据边际公共收益相对较小时 ( $0 < \beta \leq (1 - \lambda)\alpha/\lambda$ ), 企业不会给予补贴, 即补贴额度  $p = 0$ ; 当数据边际公共收益相对较大时 ( $(1 - \lambda)\alpha/\lambda < \beta < 1$ ), 补贴额度  $p > 0$  且  $p = (\alpha(\lambda - 1) + \beta\lambda)/2$ 。

引理 1 的经济学直觉很显然: 当数据公共收益较小时, 通过补贴获取更多数据的收益较小, 因此企业倾向于不设置补贴; 只有当数据公共收益增加到一定程度时, 通过补贴获取额外的数据才是合算的。

引理 2: 若企业采用自愿收集条款且  $p^* > 0$ , 则净收益最大化的补贴额度  $p^*$  与企业盈利能力  $\lambda$  及数据边际公



共收益 $\beta$ 正相关: $\partial p^*/\partial \lambda > 0$ ,  $\partial p^*/\partial \beta > 0$ ;  $p^*$ 与数据边际私人收益 $\alpha$ 负相关: $\partial p^*/\partial \alpha < 0$ 。

引理2的经济学直觉是: $\lambda$ 提高意味着企业从数据中获取的收益增大,因此企业愿意提高补贴收集更多数据; $\beta$ 提高意味着数据公共收益随数据量上升而增加越快,因此企业愿意提高补贴收集更多数据,从而提高企业收益;数据边际私人收益 $\alpha$ 越大,同意数据收集的消费者越多,意味着发放补贴时需要向更多的人发放,导致通过补贴增加数据收集总量的边际成本增加,因此企业会降低补贴额度以减少成本。

引理3:在自愿收集条款下,数据边际私人收益每提高1单位,企业数据收集总量的增加量大于0但小于或等于1单位: $0 < \partial \bar{d}/\partial \alpha \leq 1$ 。

引理3的经济学直觉是:数据来源于两部分消费者,一部分消费者即使没有补贴也愿意同意数据收集,当边际私人收益提高时,该部分消费者同步增加;另一部分消费者因补贴而同意数据收集,根据引理2,当边际私人收益提高时,企业会降低补贴额度,导致该部分消费者减少,因此企业数据收集总量的增加速度低于边际私人收益提高的速度。

## (二)强制收集

强制收集条款下,当且仅当 $u_c^*(d_c=1) > u_c^*(d_c=0)$ 时, $c$ 类型消费者会同意企业收集个人数据。于是无差异消费者的成本为:

$$\hat{c}^m = \alpha + (1-\theta)\beta\bar{d}^m \quad (16)$$

由于企业收集的数据总量为:

$$\bar{d}^m = \int_0^1 d_c f(c) dc = \int_0^{\hat{c}^m} f(c) dc = \hat{c}^m \quad (17)$$

将式(17)代入式(16),求解可知在强制收集条款下,无差异消费者的成本以及企业收集的数据总量为:

$$\hat{c}^m = \bar{d}^m = \frac{\alpha}{1-(1-\theta)\beta} \quad (18)$$

将式(18)代入式(7),可知强制收集条款下,企业的净收益为:

$$\Pi^m = \frac{\lambda \alpha (\alpha + \beta \theta (1 - (1-\theta)\beta))}{(1 - (1-\theta)\beta)^2} \quad (19)$$

引理4:当企业选择强制收集条款时,会选择 $\theta$ 最低的强制收集条款;特别地,若不存在个人信息保护法律规制,当企业选择强制收集条款时,会选择完全禁止的强制收集条款( $\theta=0$ )。

引理4在现实中可以得到验证。现实中若企业采用强制收集条款,往往是消费者不同意服务条款,企业就不向消费者提供服务。具体表现为消费者无法正常打开App,或只被允许进入“访客模式”,查看初始界面的少量信息,但几乎无法使用任何重要功能。

将 $\theta=0$ 代入式(17),可知当企业采用完全禁止的强制收集条款时,无差异消费者和企业收集数据总量为:

$$\bar{d}^m = \hat{c}^m = \frac{\alpha}{1-\beta} \quad (20)$$

将 $\theta=0$ 代入式(19),可知当企业采用完全禁止的强制收集条款时,企业的净收益为:

$$\Pi^m = \frac{\alpha^2 \lambda}{(1-\beta)^2} \quad (21)$$

在强制收集条款下,消费者不会获得补贴,因此消费者剩余的经济意义为消费者从数据服务中获取的收益减去消费者在数据收集时付出的成本。于是消费者剩余为:

$$CS^m = \int_0^1 u_c^m f(c) dc = \frac{\alpha^2}{2(1-\beta)^2} \quad (22)$$

引理5:在完全禁止的强制收集条款下,数据边际私人收益每提高1单位,企业数据收集总量的增加量大于1单位: $\partial \bar{d}^m/\partial \alpha > 1$ 。

引理5表明在完全禁止的强制收集条款下,边际私人收益增加对于数据收集总量具有“乘数效应”,因此企业的数据收集总量增加速度高于边际私人收益提高的速度。这一机制与哈久和赖特(2023)提出的“数据赋



能的学习”(Data-enabled learning)相似,即企业利用数据提高服务质量,从而使得企业能收集到更多数据。

### (三)企业净收益最大化的数据收集条款选择

比较企业在两种条款下的净收益,可知企业净收益最大化的数据收集条款选择如图1所示。

图1中纵轴表示数据边际私人收益( $\alpha$ )的取值,横轴表示数据边际公共收益( $\beta$ )的取值,过点(0,1)和点(1,0)的实线方程为 $\alpha=1-\beta$ ,该线和横轴纵轴所包围的三角区域(不包括边界)即为本文讨论的 $\alpha$ 与 $\beta$ 的取值范围( $\alpha>0, \beta>0$ 且 $\alpha+\beta<1$ )。图1中过点(0,0.5)和点(1,0)的曲线表示 $\alpha=\hat{\alpha}_1$ ,过原点的虚线表示 $\alpha=\beta\lambda/(1-\lambda)$ 。于是有命题1。

命题1:当不存在个人信息保护时,对于任意 $\lambda \in (0,1)$ ,企业根据数据边际私人收益和边际公共收益选择数据收集条款:

①对于边际私人收益相对较高( $\alpha>\hat{\alpha}_1$ )的数据收集场景,企业会选择完全禁止的强制收集条款。

②对于边际私人收益相对较低( $0<\alpha<\hat{\alpha}_1$ )的数据收集场景,企业会选择自愿收集条款。且当边际公共收益相对较低时( $0<\beta\leq(1-\lambda)\alpha/\lambda$ ),补贴额度 $p=0$ ;当边际公共收益相对较高时( $(1-\lambda)\alpha/\lambda<\beta<1$ ),补贴额度 $p>0$ 。其中,

$$\hat{\alpha}_1 = \begin{cases} \frac{(1-\beta)^2}{2-\beta}, & 0 < \beta \leq 1 - \sqrt{\lambda} \\ \frac{\lambda\beta(1-\beta)(2\sqrt{\lambda} + (1-\beta)(1+\lambda))}{4\lambda - (1-\beta)^2(1+\lambda)^2}, & 1 - \sqrt{\lambda} < \beta < 1 \end{cases} \quad (23)$$

命题1的经济学直觉如下:根据引理3,随着数据边际私人收益提高,企业通过补贴提高数据收集总量的效果较差;而根据引理5,企业使用完全禁止的强制收集条款效果较好。所以当边际私人收益较高时,与自愿收集条款相比,企业会更倾向于选择完全禁止的强制收集条款。反之,当数据边际私人收益较低时,企业会更倾向于选择自愿收集条款。

根据命题1,企业会依据数据收集场景选择数据收集条款类型。首先,关于企业在强制收集条款和自愿收集条款之间的选择,可以考虑以下两个现实例子。

第一个例子是实时位置数据。在导航App中,用户实时位置直接关系到路线引导,因此边际私人收益相对较高,企业会采用完全禁止的强制收集条款;而在外卖App中,用户实时位置数据虽然有助于用户找到附近商家,但并不影响用户订购外卖,因此边际私人收益相对较低,企业会采用自愿收集条款。例如在高德地图App,若消费者使用驾车、骑行、步行、公交地铁、停车记录等导航或路线规划服务,会将位置信息作为必要信息予以收集;而在美团App,消费者可以选择是否开启定位服务,若消费者拒绝授权,仍可以通过手动输入收货地址使用外卖等基本服务,只在网约车、单车、地图导航等场景下,位置信息是必要信息。

第二个例子是手机号码信息。在导航App中,用户手机号码与导航主要功能无关,因此边际私人收益相对较低,企业会采用自愿收集条款;而在外卖App中,用户手机号码关系到用户是否能收到外卖,因此边际私人收益相对较高,企业会采用完全禁止的强制收集条款。例如在高德地图App,消费者一般无需提交手机号码信息创建账号,也能正常使用导航和路线规划等核心功能;而在美团App,若消费者拒绝提交手机号码,则无法创建美团账号,不能使用外卖、团购及酒店住宿等核心功能。

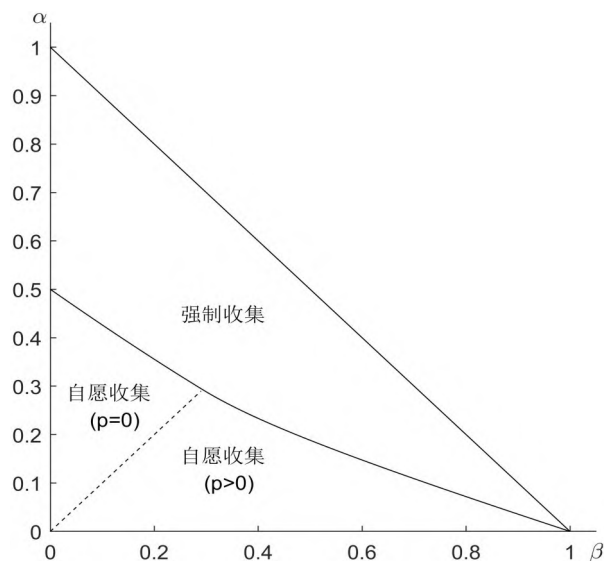


图1 不存在个人信息保护时企业的数据收集条款选择

其次,关于企业对自愿收集条款补贴额度的设置,可以考虑评论信息的例子。无论是在视频 App 还是购物 App,评论信息给提交评论的用户本人带来的收益较低,因此企业会采用自愿收集条款。在视频 App 中,评论对于评判视频质量的效果相对较弱,因此边际公共收益较低,企业一般不会给予补贴;而在购物 App,评论信息对于其它消费者而言具有很强的信号传递作用,因此边际公共收益较高,企业会给予补贴。例如,在哔哩哔哩 App,发布评论既不能获取优惠券或现金红包等奖励,也无法获得经验值进而解锁更多权益;而在京东,消费者提交评论可获得名为“京豆”的代金券,可以在购物时进行抵扣;在淘宝,消费者提交评论可获得“淘气值”,而随着淘气值增加,消费者可以获取更多的优惠券及其它会员服务<sup>⑩</sup>。

## 五、不同数据收集条款下的消费者福利比较

### (一)消费者整体福利

假定以下两种情形:企业只选择自愿收集条款,以及企业只选择完全禁止的强制收集条款,并对比两种情形下的消费者整体福利。可知两种条款下的消费者福利相对大小如图 2 所示,于是有命题 2。当企业盈利能力较弱时( $0 < \lambda < \sqrt{3} - 1$ ),福利对比如图 2a 所示;当企业盈利能力较强时( $\sqrt{3} - 1 \leq \lambda < 1$ ),福利对比如图 2b 所示。

命题 2:在某些情形下,完全禁止的强制收集条款下的消费者整体福利高于自愿收集条款下的消费者整体福利:

①若企业盈利能力较弱( $0 < \lambda < \sqrt{3} - 1$ ),当数据边际私人收益较高时( $\alpha > \hat{\alpha}_2$ ),完全禁止的强制收集条款下的消费者剩余高于自愿收集条款下的消费者剩余;

②若企业盈利能力较强( $\sqrt{3} - 1 \leq \lambda < 1$ ),当数据边际私人收益较高( $\alpha > \hat{\alpha}_2$ ),且数据边际公共收益较低时( $0 < \beta < \frac{1+\lambda-2\sqrt{(1+\lambda)^2-3}}{3}$ ),完全禁止的强制收集条款下的消费者剩余高于自愿收集条款下的消费者剩余。

其中:

$$\hat{\alpha}_2 = \begin{cases} \frac{2(1-\beta)^2}{2-\beta}, 0 < \beta \leq 1 - \frac{\sqrt{\lambda}}{2-\sqrt{\lambda}} \\ \frac{(-1+\beta)^2\beta(1+\lambda)(2+\lambda)+2\sqrt{(-1+\beta)^2\beta^2(1-2\beta(1+\lambda)^2+\beta^2(1+\lambda)^2+2\lambda(3+\lambda))}}{(-3+\beta+(-1+\beta)\lambda)(1+\beta+(-1+\beta)\lambda)}, 1 - \frac{\sqrt{\lambda}}{2-\sqrt{\lambda}} < \beta < 1 \end{cases} \quad (24)$$

命题 2 的经济学直觉如下:数据对于消费者具有非竞争性,自愿收集条款和强制收集条款是两种提高数据收集效率的方式,有助于克服部分消费者的“搭便车”问题,增加数据收集总量。在某些情形下,强制收集条款对数据收集效率的提升相对于自愿收集条款更为明显,超过了它对消费者福利的负面影响,反而成为更有利于整体消费者福利的数据收集条款。

结合命题 1 与命题 2 可得到以下推论。

推论 1:当  $\max\{\hat{\alpha}_1, \hat{\alpha}_2\} < \alpha < 1 - \beta$  时,相对于自愿收集条款,完全禁止的强制收集条款下企业净收益与消

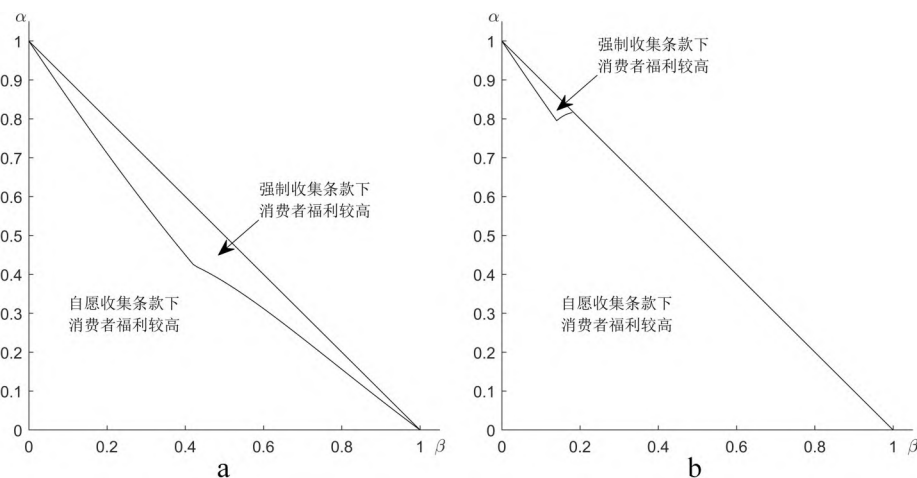


图 2 不同数据收集条款下的消费者整体福利比较(a:  $0 < \lambda < \sqrt{3} - 1$ ; b:  $\sqrt{3} - 1 \leq \lambda < 1$ )

费者剩余都相对较高。

推论1的经济学含义在于,不能将强制收集条款简单视为企业对消费者福利的攫取。在某些情形下,相对于自愿收集条款,强制收集条款是一种更有效率的制度安排,可以同时增加企业净收益和消费者福利。

## (二)异质性消费者福利

由于消费者具有异质性,即使在同一数据收集条款下,不同类型消费者的福利也不相同。下面分别考察数据收集条款如何影响不同消费者的福利。

当  $0 < \beta \leq (1-\lambda)\alpha/\lambda$  时,在自愿收集条款下,补贴额度等于0,某个  $c$  类型消费者效用为:

$$u_c^v = \begin{cases} \alpha + \alpha\beta - c, 0 < c \leq \alpha \\ \alpha\beta, \alpha < c \leq 1 \end{cases} \quad (25)$$

当  $(1-\lambda)\alpha/\lambda < \beta < 1$  时,在自愿收集条款下,补贴额度大于0,某个  $c$  类型消费者效用为:

$$u_c^v = \begin{cases} \alpha + \frac{1}{2}(\alpha(\lambda-1) + \beta\lambda) + \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda)\beta - c, 0 < c \leq \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda) \\ \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda)\beta, \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda) < c \leq 1 \end{cases} \quad (26)$$

在完全禁止的强制收集条款下,某个  $c$  类型消费者效用为:

$$u_c^m = \begin{cases} \alpha + \frac{\alpha\beta}{1-\beta} - c, 0 < c \leq \frac{\alpha}{1-\beta} \\ 0, \frac{\alpha}{1-\beta} < c \leq 1 \end{cases} \quad (27)$$

命题3:异质性消费者在不同数据收集条款下的福利对比结果如下。

①当  $0 < \beta \leq (1-\lambda)\alpha/\lambda$  时,对于成本较低的消费者 ( $0 \leq c < (1-\beta+\beta^2)\alpha/(1-\beta)$ ),其在自愿收集条款下的福利低于完全禁止的强制收集条款下的福利;对于成本较高的消费者 ( $(1-\beta+\beta^2)\alpha/(1-\beta) < c \leq 1$ ),其在自愿收集条款下的福利高于完全禁止的强制收集条款下的福利。

②当  $(1-\lambda)\alpha/\lambda < \beta < 1$  时,若数据边际私人收益较低 ( $0 < \alpha \leq \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ ),所有消费者 ( $0 \leq c \leq 1$ ) 在自愿

收集条款下的福利均不低于在完全禁止的强制收集条款下的福利;若数据边际私人收益较高 ( $\frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$

$< \alpha < \min\left\{\frac{\beta\lambda}{1-\lambda}, 1-\beta\right\}$ ),对于成本较低的消费者 ( $0 \leq c < \frac{2\alpha - (\alpha(\lambda+1) + \beta\lambda)(1-\beta)\beta}{2(1-\beta)}$ ),其在自愿收集条款下的福

利低于完全禁止的强制收集条款下的福利,对于成本较高的消费者 ( $\frac{2\alpha - (\alpha(\lambda+1) + \beta\lambda)(1-\beta)\beta}{2(1-\beta)} < c \leq 1$ ),其在自

愿收集条款下的福利高于完全禁止的强制收集条款下的福利。

命题3的经济学含义是:当数据边际私人收益足够高,或数据边际公共收益较低时,相比自愿收集条款,强制收集条款可以增加低成本消费者的福利,而减少高成本消费者的福利。原因在于强制收集条款会“挤出”高成本消费者,从而损害其福利;但强制收集条款可以提高企业的数据收集总量,从而增加数据公共收益,有利于未被“挤出”的低成本消费者。

从数据要素的视角进一步引申上述命题。强制收集条款是企业市场机制下收集数据要素的方式之一(此时的市场交易体现为“服务换数据”),企业通过强制收集条款将高成本消费者“挤出”,以较低成本获取数据要素,与企业雇佣职员获取劳动要素具有相似性。企业放弃雇佣高薪资要求的员工,无疑会损害此类员工的利益,但一般被视为合理的经济行为。如果认为市场机制应在要素配置中起决定性作用,那么应该承认强制收集条款作为企业数据收集方式具有一定的合理性。

命题3还表明,在数据要素市场上限制使用强制收集条款,其效果类似于其它要素市场上的规制政策,也可能伤害低成本的要素供给者。例如在劳动要素市场,最低工资标准可能损害低成本劳动者(愿意接受低工



资的低技能劳动者)的福利(克莱门斯、迈克尔,2019)。

## 六、个人信息保护法律对企业数据收集条款及消费者福利的影响

为了讨论个人信息保护法律规制的一般情形,本文假定立法与司法活动要求企业必须为拒绝数据收集的消费者提供部分服务。个人信息保护法律对强制收集条款的规制力度越大,则该部分服务占全部服务的比例越高。于是企业对数据收集条款的选择就受到个人信息保护法律的约束:企业若采用强制收集条款,必须为拒绝数据收集的消费者提供不低于该比例的服务。接下来的分析遵循以下思路:首先,个人信息保护会影响企业的数据收集条款选择;其次,分析个人信息保护在哪些数据收集场景中能够增加消费者整体福利;最后,个人信息保护法律旨在保护消费者的个人权益,因此消费者福利帕累托改进更贴合“正义的效率涵义”(波斯纳,2014)。于是本文讨论了个人信息保护是否能够实现消费者福利帕累托改进,并基于此为完善个人信息保护法律提供政策启示。

### (一)个人信息保护与企业数据收集条款选择

假定法律规定了拒绝数据收集的情形下,消费者可以使用企业服务的比例下限 $\hat{\theta}$ ,即规定 $\theta \geq \hat{\theta}$ ,其中 $\hat{\theta} \in [0, 1]$ 。 $\hat{\theta}$ 越高意味着个人信息保护法律对强制收集条款的规制力度越高。 $\hat{\theta}=1$ 时,意味着法律完全禁止企业采用强制收集条款。 $\hat{\theta} \in (0, 1)$ 时,企业可以选择限制使用的强制收集条款,即必须为拒绝收集数据的消费者提供一部分服务<sup>⑤</sup>。 $\hat{\theta}=0$ 时,意味着法律完全放任企业采用强制收集条款,相当于没有个人信息保护的状态,因此本文只讨论 $\hat{\theta} \in (0, 1]$ 的情况。

根据引理4,当个人信息保护法律要求 $\theta \geq \hat{\theta}$ 时,企业如果选择强制收集条款,会选择限制使用的强制收集条款( $\theta = \hat{\theta}$ )。

个人信息保护对企业数据收集条款选择的影响可以由图3表示,其中曲线BC代表实施个人信息保护前,企业选择完全禁止的强制收集条款( $\theta=0$ )与自愿收集条款的分界线;虚线CD表示实施个人信息保护( $\hat{\theta} \in (0, 1]$ )后,企业选择限制使用的强制收集条款( $\theta = \hat{\theta}$ )与自愿收集条款的分界线,于是共有3种情形。

命题4:对于任意 $\lambda \in (0, 1)$ ,规制力度 $\hat{\theta}$ ( $\hat{\theta} \in (0, 1]$ )的个人信息保护法律使企业数据收集条款产生如下改变。

①若 $0 < \alpha < \hat{\alpha}_1$ ,数据收集条款不变。该情形对应图3区域OBC。

②若 $\hat{\alpha}_1 < \alpha < \hat{\alpha}_3$ ,企业放弃完全禁止的强制收集条款( $\theta=0$ ),选择自愿收集条款。该情形对应图3区域BCD。

③若 $\hat{\alpha}_3 < \alpha < 1 - \beta$ ,企业选择限制使用的强制收集条款( $\theta = \hat{\theta}$ )。该情形对应图3区域ACD<sup>⑥</sup>。

其中, $\hat{\alpha}_1$ 表达式如式(23)所示, $\hat{\alpha}_3$ 对应实施个人信息保护后企业选择限制使用的强制收集条款( $\theta = \hat{\theta}$ )区域与选择自愿收集条款区域的分界线。限于篇幅, $\hat{\alpha}_3$ 的表达式以及关于命题4证明的详细内容参见《管理世界》网络发行版附录2。

### (二)个人信息保护与消费者整体福利

根据实施个人信息保护前后企业数据收集条款选择的变化,分3种情形讨论个人信息保护对消费者福利的影响。

①若 $0 < \alpha < \hat{\alpha}_1$ ,法律规制后数据收集条款不变,于是消费者福利不变。该情形对应图4a的加粗线内区域OBC。

②若 $\hat{\alpha}_1 < \alpha < \hat{\alpha}_3$ ,法律规制使企业放弃完全禁

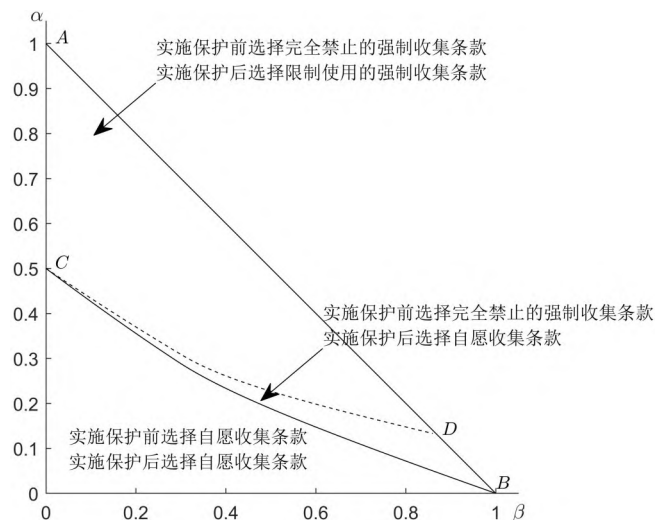


图3 个人信息保护对企业数据收集条款选择的影响

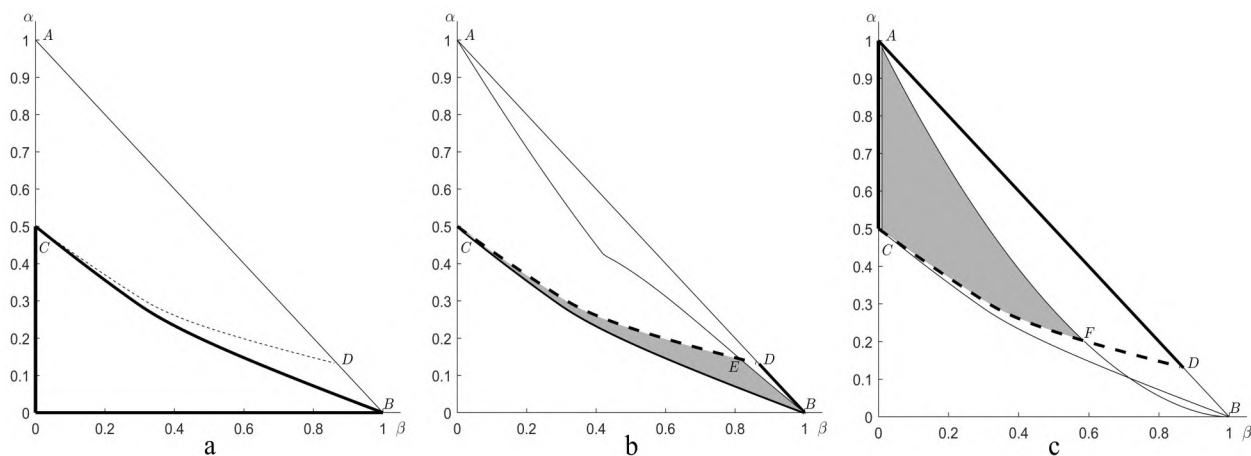


图4 个人信息保护影响消费者整体福利的3种情形

止的强制收集条款,选择自愿收集条款。根据命题2,消费者福利可能增加(图4b的阴影区域BCE),也可能减少(图4b的区域BDE)<sup>⑩</sup>。

③若 $\hat{\alpha}_3 < \alpha < 1 - \beta$ ,法律规制使企业选择限制使用的强制收集条款( $\theta = \hat{\theta}$ )。消费者福利可能增加(图4c的阴影区域ACF),也可能减少(图4c的区域ADF)。

最终,个人信息保护对消费者福利的影响如图5所示。于是可以得到命题5:

命题5:对于任意 $\hat{\theta} \in (0, 1]$ 以及任意 $\lambda \in (0, 1)$ ,当 $\alpha \in (\min\{\hat{\alpha}_2, \hat{\alpha}_3, 1 - \beta\}, \min\{\hat{\alpha}_3, 1 - \beta\})$ 或 $\alpha \in \left( \frac{2(1 - \beta(1 - \hat{\theta}))(1 - \beta)^2}{2 - \beta(1 - \hat{\theta}) - \beta}, 1 - \beta \right)$ 时,实施个人信息保护后消费者整体福利减少。

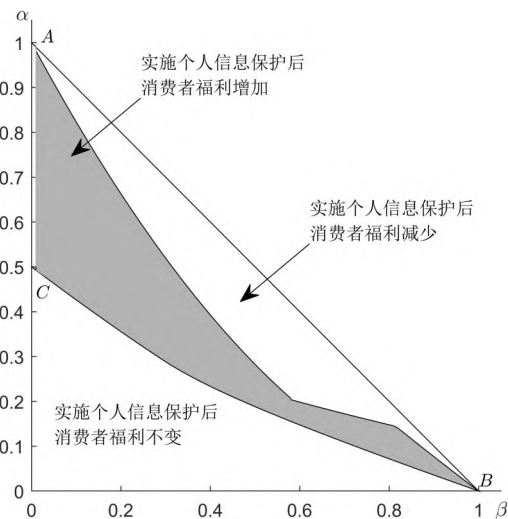


图5 个人信息保护对消费者整体福利的影响

命题5的经济学含义在于,个人信息保护对消费者群体具有“双刃剑”效果:它保障了消费者搭便车的权利,使消费者能够共享数据公共收益;但同时削弱了企业利用强制收集条款提升数据收集效率的能力,可能导致服务质量下降,进而导致消费者福利减少。因此,不存在一种任何数据收集场景中都能保障消费者福利的规制力度。

根据命题5,无论如何调整规制力度,个人信息保护都可能在某些数据收集场景中损害消费者整体福利。例如求职、婚恋交友等类型App采用强制收集条款收集消费者身份、学历等个人信息,此类信息有助于帮助消费者匹配合适的求职单位或交友对象,因此边际私人收益和边际公共收益都较高。如果限制企业采用强制收集条款收集此类信息,少部分消费者可能会在隐藏真实身份的情形下进入企业,导致匹配效率下降,甚至可能诱发“柠檬市场”,损害消费者整体福利。

命题5也表明,如果只根据数据边际私人收益来确定个人信息保护法律的规制范围,可能在某些数据收集场景中损害消费者福利。例如,若法律规制使得 $\alpha$ 低于某一阈值时,企业不能采用强制收集条款。那么根据图5所示,即使该阈值很低,当数据边际公共收益足够高时,规制依然会减少消费者福利。

### (三)个人信息保护与消费者福利帕累托改进

根据命题3,个人信息保护对不同消费者的福利影响可能不一致。因此,如果要在没有任何消费者福利减少的前提下,增加部分消费者的福利,即实现消费者福利的帕累托改进,必须分析实施个人信息保护对每一类消费者的影响。经计算可知,个人信息保护实现消费者福利帕累托改进的区域如图6阴影区域所示,于是可

以得到命题6。

命题6:实施个人信息保护可以实现消费者福利的帕累托改进的充要条件是:

- ①企业盈利能力较强( $1/4 < \lambda < 1$ );
- ②数据边际私人收益适中( $\hat{\alpha}_1 < \alpha < \hat{\alpha}_4$ )且边际公共收益较高( $\beta > \hat{\beta}_1$ );
- ③实施个人信息保护使企业放弃完全禁止的强制收集条款,转而采用自愿收集条款( $\alpha < \hat{\alpha}_3$ )。

其中, $\hat{\beta}_1$ 表示图6中点G的横坐标,当 $1/4 < \lambda < 1$ 时,点G为曲线BC和曲线OGB在区间 $\beta \in (0, 1)$ 内的唯一

交点。 $\hat{\alpha}_1$ 的表达式如式(23)所示, $\hat{\alpha}_4 = \min \left\{ \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}, 1-\beta \right\}$ 。

进一步分析当个人信息保护法律对强制收集条款的规制力度改变时,消费者福利帕累托改进的区域如何随之变化。可知变化趋势如图7所示,其中加粗虚线表示实施个人信息保护时,企业采用限制使用的强制收集条款( $\theta = \hat{\theta}$ )与自愿收集条款的分界线,阴影区域表示消费者福利帕累托改进的区域。于是有推论2。

推论2:存在某个临界规制力度(即存在某个阈值 $\hat{\theta}^* \in (0, 1)$ ),当规制力度低于临界力度时( $\hat{\theta} < \hat{\theta}^*$ ),随着规制力度提高( $\hat{\theta}$ 提高),消费者福利帕累托改进的区域增大;当规制力度超过临界力度时( $\hat{\theta} \geq \hat{\theta}^*$ ),随着规制力度提高( $\hat{\theta}$ 提高),消费者福利帕累托改进的区域不变。

推论2表明,在某些数据收集场景中,实施力度较弱的规制,即准许企业采用限制使用的强制收集条款,有助于消费者福利帕累托改进,而继续增加规制力度无助于实现进一步的消费者福利帕累托改进。

命题6和推论2的经济学含义在于,如果以消费者福利的帕累托改进为目标,应当审慎地把握个人信息保护的范围与力度,对不同数据收集情形进行差异化保护。典型的情形可用图8表示。

第一,如图8浅色阴影区域所示,对于边际私人收益较高的数据,企业会采用强制收集条款,个人信息保护会损害低成本消费者的福利,因此对其采取保护需要谨慎。例如导航App采用强制收集条款收集消费者位置信息,如果消费者拒绝收集位置信息,则无法使用导航和规划最优路线等服务。位置信息对于导航是必要数据,边际私人收益较高。如果限制企业采用强制收集条款收集位置信息,少部分只需查看路线的消费者可能会拒绝提交位置信息,于是企业规划最优路线时无法将这部分消费者造成的拥堵压力计算在内,从

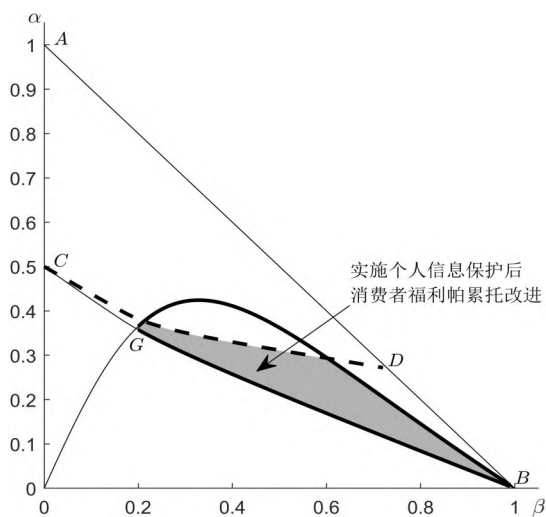


图6 实施个人信息保护后消费者福利帕累托改进区域

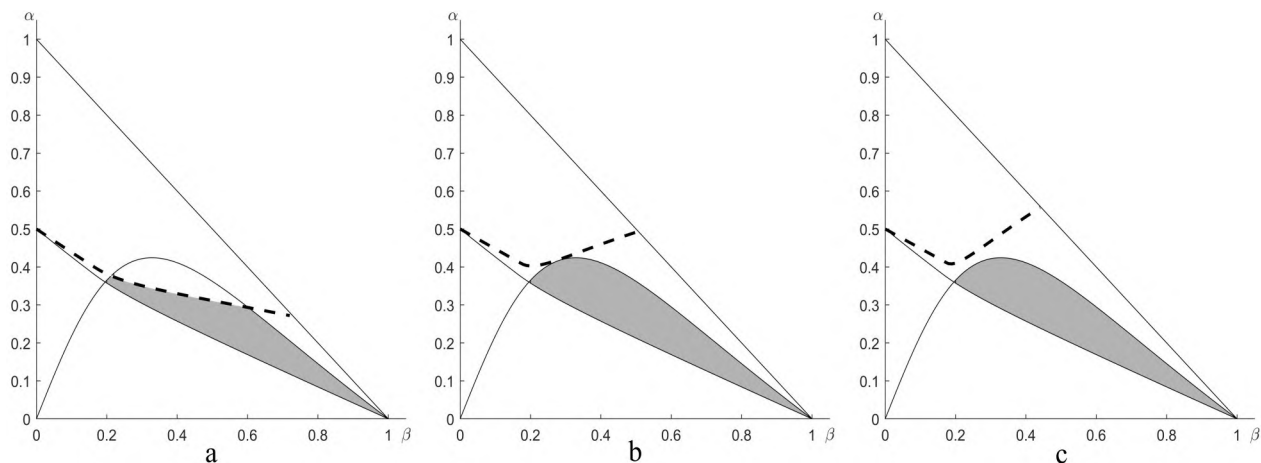


图7 规制力度影响消费者福利帕累托改进区域(a: $\hat{\theta} < \hat{\theta}^*$ ; b: $\hat{\theta} = \hat{\theta}^*$ ; c: $\hat{\theta} > \hat{\theta}^*$ )



而导致愿意提交位置信息的消费者无法获取准确的最优路线规划结果。这意味着个人信息保护的后果是损害部分消费者福利而保护另一部分消费者,因此需要谨慎考虑。

第二,如图8左下方白色区域所示,对于边际私人收益较低而边际公共收益较高的数据,企业会采用自愿收集条款,个人信息保护不会改变企业的数据收集条款选择。在此类数据收集场景中,应该充分尊重企业与消费者之间基于合意原则达成的协议。

第三,如图8深色阴影区域所示,对于边际私人收益适中而边际公共收益较高的数据,企业会采用强制收集条款,一定程度的个人信息保护会使企业转而采用自愿收集条款,从而实现消费者福利的帕累托改进。以商业保险企业收集医疗数据为例。由于医疗数据涉及重要隐私,对于消费者个人而言,收集医疗数据的私人收益不高。但收集医疗数据有助于减少信息不对称,规避恶意骗保等风险,推进保险市场健康发展,从而让大多数消费者获取更优质的服务。因此,在保险行业中,医疗数据对于消费者具有较高的公共收益。根据本文结论,应该对保险企业收集医疗数据的行为进行一定程度的规制,禁止企业采用完全强制性的数据收集条款(例如,若消费者拒绝企业收集医疗数据,企业则拒绝提供任何保险服务)。但是,可以允许企业对拒绝数据收集的消费者采取一定程度的差异化对待。例如,允许企业向同意数据收集的消费者设置更低门槛的赔付条件,既为企业提供获取数据的渠道,也为消费者保留一定的自主选择空间,从而在释放数据要素生产力与保护消费者个人隐私之间取得平衡。

此外,若企业盈利能力较弱,个人信息保护也无法实现消费者福利的帕累托改进。例如对于盈利能力较弱甚至亏损的企业,如果法律规制其强制收集行为,企业不会采用补贴额度足够高的自愿收集条款,可能损害消费者福利。因此,政府可以为尚未盈利的初创企业提供数据收集补贴,补偿因法律规制而增加的成本,促进数据要素利用,培育更多服务于消费者的新兴业态。

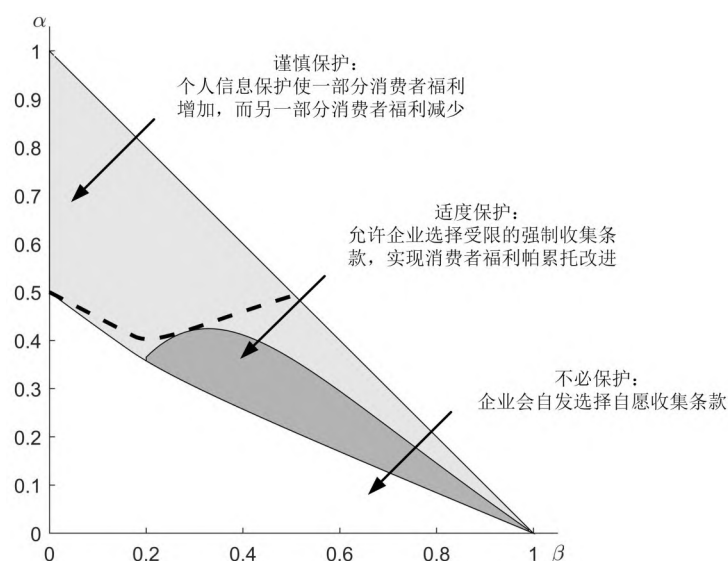


图8 差异化个人信息保护实现消费者福利帕累托改进的典型情形

## 七、结论与启示

数据收集是企业价值创造的重要途径,也引发对于消费者福利损失的担忧,因此已成为法律与政策关注的重点。本文比较企业的两种数据收集条款:自愿收集条款和强制收集条款,研究企业的数据收集条款的内生选择及其对消费者福利的影响,并分析个人信息保护法律的规制效果。本文发现,若数据的边际私人收益较高,则企业采用强制收集条款;若数据的边际私人收益较低,企业采用自愿收集条款,且补贴额度受数据边际公共收益影响。而福利分析表明,相比于自愿收集条款,虽然强制收集条款意味着消费者个人权利受限,但可能通过提高数据公共收益增加消费者福利;倘若禁止强制收集条款,可能减少消费者整体福利尤其是低成本消费者的福利。本文随后讨论个人信息保护法律对强制收集条款进行规制的一般情形并发现:实施力度较弱的规制有助于在某些数据收集场景中实现消费者福利帕累托改进,而过强的规制力度无助于进一步实现消费者福利帕累托改进。

本文的研究为推进数据要素市场化配置改革和完善个人信息保护法律提供了政策启迪。市场主体之间的契约关系是市场机制的基石,而自愿收集条款和强制收集条款都是企业收集数据的重要契约条款形式。在保障消费者知情权的基础上,应该允许企业灵活采用不同类型的数据收集条款。这不仅有利于促进数据要素

有效配置,而且并不违背个人信息保护的效率目标。具体而言,可以从以下3个方面着手。

第一,将数据对于消费者的公共收益纳入立法考量,限定个人信息保护法律的规制范围。当前《个保法》对强制收集条款的规制范围取决于个人信息的“必要性”。但在某些情形中,虽然从个人角度看数据并非服务所必要,但服务却依赖于若干消费者数据的积累。因此,在认定“非必要信息”时,不仅要考虑数据对于提供者本人的直接收益,还要考虑数据积累为其他消费者带来的公共收益。当数据公共收益较高时,应该承认个人信息对于企业服务的必要性,允许采用带有强制色彩的数据收集条款。

第二,充分考虑企业类型与数据应用场景的异质性,推进数据分类治理。在不同的数据收集场景中,数据收益和隐私成本的相对大小具有重大差异。应该依据实际情形,判断自发的市场秩序是否能保障消费者福利,从而厘清市场有效与市场失灵的边界,并对不同的数据收集场景采取差异化的监管措施。例如,应该准许导航App强制收集位置数据,而对保险企业收集医疗数据的行为应该给予一定程度的规制。此外,推进数据分类治理意义不仅局限于保障消费者福利,也有助于市场机制在数据要素配置过程中发挥决定性作用,并让法律规制和政府监管得以有的放矢。

第三,降低初创企业的“用数成本”,促进数据产业高质量发展。在互联网、大数据与人工智能等新兴产业,许多尚未实现盈利的初创企业可以孵育出市场潜力巨大的新产品和新业态。但是,初创企业往往难以负担高昂的数据收集成本。因此在实施个人信息保护法律的情况下,政府可以设立针对初创企业的数据收集补贴,促进数据要素高水平利用,从而充分解放新质生产力,推动更多服务于消费者的新业态涌现。

当然,本文模型还存在一些局限性。第一,本文没有考察市场竞争对企业数据收集条款的影响。现实中某些企业会以隐私保护作为竞争手段,可能策略性地调整数据收集条款。例如苹果公司宣称自己的产品注重保护用户隐私,以此获取竞争优势。第二,本文假定企业从消费行为中获取的收益比率是外生的。而在现实中,尽管企业的佣金比率和广告分成受到市场环境的制约,但考虑企业内生决定其收益比率将会更合理。第三,本文假定只存在两类数据收集条款,这是对现实情形的一种简化。在真实的数据场景中,企业提供的隐私条款可能存在多个维度的差异,例如是否允许撤销同意并清除数据,是否允许与第三方分享个人信息,是否允许消费者下载个人数据等,在未来的研究中可以进一步刻画这些数据收集条款并研究其对消费者福利的影响<sup>⑧</sup>。

(作者单位:刘征驰、叶宇阳,湖南大学经济与贸易学院;聂辉华,中国人民大学经济学院、中国人民大学企业与组织研究中心)

#### 注释

①数据和信息的定义存在差别,但在本文的情境中,消费者的个人信息会被企业以数据的形式收集利用,因此不对二者做出严格区分。

②《General Data Protection Regulation》,参见<https://gdpr-info.eu/>;《中华人民共和国个人信息保护法》,参见[http://www.npc.gov.cn/npc/c2/c30834/202108/t20210820\\_313088.html](http://www.npc.gov.cn/npc/c2/c30834/202108/t20210820_313088.html);《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》,参见[https://www.gov.cn/zhengce/2022-12/19/content\\_5732695.htm](https://www.gov.cn/zhengce/2022-12/19/content_5732695.htm)。

③由于篇幅所限,相关案例详细内容参见《管理世界》网络发行版附录1。

④不同消费者对于企业收集数据的反感程度不一致,例如中国消费者协会发布的2023年消费维权年主题调查结果显示,不支持当前APP产品强制要求授权和索取个人信息的消费者比例为46.7%,意味着还有约半数消费者对数据收集并无特别反感。因此可以合理推测,数据收集给不同消费者带来的成本是有差异的。参见《2023年消费维权年主题调查报告》, <https://www.cca.org.cn/Detail?catalogId=475800865030213&contentType=article&contentId=521578466259013>。

⑤经典价格理论认为,最低工资标准会导致原本愿意接受较低工资的低技能劳动者更难就业,从而损害其福利。某些经验研究支持该观点(克莱门斯、迈克尔,2019),但也有文献发现相反的证据(卡德、克鲁格,1994)。因此,不妨说至少在某些情形下,最低工资标准会损害低技能劳动者的福利。

⑥在大数据技术背景下,企业可以利用数据向用户提供独特的服务,形成所谓“数据型垄断”(施图克、格伦斯,2016),因此本文假设企业是垄断的。

⑦式(2)的消费者效用函数隐藏了以下假设:当没有任何消费者数据被收集时,消费者使用企业服务的效用为0,等于其不使用企业服务的保留效用。其经济含义为,如果没有消费者数据,企业将无法提供独特服务。

⑧本文排除了 $\alpha=0$ 的情形,以保证总能得到数据收集总量大于0的均衡结果。

⑨本文假设数据边际公共收益为一个常量,这是一种为了便于度量的简化假设。在很多机器学习场景中,数据积累带来的公共

收益体现为模型准确率随数据量提升而提高,但准确率提高速度缓慢下降(范里安,2019)。但是,数据量提升也可能带来如生成式人工智能这样的破坏性创新,此时边际公共收益可能呈现出递增的特点。为了便于从整体上比较数据边际公共收益与边际私人收益的大小,本文假设数据边际公共收益为一个常量。

⑩“基于数据的服务越能满足其需求,则该消费者越有可能为企业带来收益”,这一关系不仅仅存在于广告收入的商业模式中,也存在其他商业模式中。例如,在收取服务费的商业模式中,企业借助数据向消费者提供更优质的服务,也可以收取更高的服务费从而增加收益;在企业将聚合数据产品化后进行交易的商业模式中,企业利用数据提高服务质量,使得企业能够收集到更多数据,从而企业可以获取更多收益。目前数据要素的价值,集中体现为各种机器学习算法与人工智能服务快速发展;而在广告收入模式中,企业利用基于数据的算法识别消费者偏好并精准推送广告,正是数据要素产生价值的典型业态,因此本文基于广告收入的商业模式构造企业收益函数。此外,在本文《管理世界》网络发行版附录3,还基于其他两种商业模式分别构建了新的企业收益函数(收取服务费和聚合数据产品化),发现均能得到与正文模型类似的结论。

⑪实际上,也可能存在 $p < 0$ 的情形,即企业对同意数据收集的消费者收取额外的费用。例如某些企业存在会员服务,相比普通消费者,会员可能需要提交更多的信息,并支付一笔会员费,但可以享受更优质的服务。当考虑 $p < 0$ 的情形时,本文模型的基本结论不会改变。但 $p < 0$ 的情形中, $p$ 的经济意义不再是企业为数据收集而支出的补贴,而是企业为服务收取的额外费用。服务收费问题并非本文所要讨论的重点;况且从数据收集的角度而言, $p < 0$ 时的补贴额度依然可以视为0,于是本文限定 $p \geq 0$ 。

⑫企业还存在其他成本,例如存储数据等,但其主要表现为固定成本,边际成本接近于0(夏皮罗、范里安,1999),对企业选择数据收集条款影响较小,因此不在此讨论。

⑬本文引理、命题和推论的详细证明内容参见《管理世界》网络发行版附录2。

⑭关于高德地图对位置信息和手机号码的收集,可参见《高德地图基本功能隐私政策》, [https://terms.alicdn.com/legal-agreement/terms/privacy\\_policy\\_full/20250314145608599/20250314145608599.html](https://terms.alicdn.com/legal-agreement/terms/privacy_policy_full/20250314145608599/20250314145608599.html);关于美团对位置信息和手机号码的收集,可参见《美团隐私政策概要》, <https://fe-config.meituan.com/bm/config/3c8761gdjhcj2ced86.html>;关于哔哩哔哩的经验值获取方式,可登陆账号并进入个人中心查看;关于“京豆”的说明,可参见 <https://in.m.jd.com/help/app/score.html>;关于“淘气值”的说明,可在“淘宝”App中依次点击“我的淘宝”——“会员中心”后查看。

⑮例如当消费者拒绝同意某些App的隐私政策时,可以进入“基本模式”或“访客模式”,使用部分基础服务。

⑯当 $\theta=1$ 时,企业只能选择 $\theta=1$ 的强制收集条款,相当于 $p=0$ 的自愿收集条款,因此情形③即为情形②的一部分,但不影响本文其他结论。

⑰当 $\lambda \in (\sqrt{3}-1, 1)$ 且规制力度较低时,情形②中消费者福利不可能减少,进而导致图4b及图5的阴影区域形状略有变化,但不影响本文其他结论。

⑱中外文人名(机构名)对照:崔(Choi);多西斯(Dosis);桑德-赞特曼(Sand-Zantman);阿西莫格鲁(Acemoglu);罗曼纽克(Romanyuk);斯莫林(Smolín);林奇(Lynch);阿奎斯蒂(Acquisti);琼斯(Jones);托内蒂(Tonetti);丛(Cong);科尼策(Conitzer);蒙特斯(Montes);市桥(Ichihashi);宾皮基斯(Bimpikis);戈德法布(Goldfarb);塔克(Tucker);沈(Shen);米格尔(Miguel);布洛赫(Bloch);德芒热(Demange);范梅塞尔(Fainmesser);孔蒂(Conti);雷韦尔贝里(Reverberi);阿里多尔(Aridor);德米雷尔(Demirer);波伊克特(Peukert);哈久(Hagi);赖特(Wright);克莱门斯(Clemens);迈克尔(Michael);波斯纳(Posner);卡德(Card);克鲁格(Krueger);施图克(Stucke);格伦斯(Grunes);范里安(Varian);夏皮罗(Shapiro)。

#### 参考文献

- (1)丁晓东:《论个人信息法律保护的思想渊源与基本原理——基于“公平信息实践”的分析》,《现代法学》,2019年第3期。
- (2)何佳、高彧、孟涓涓、王曾:《个人信息披露决策:强制收集与挤出效应》,《经济研究》,2022年第5期。
- (3)李三希、王泰茗、刘小鲁:《数据投资、数据共享与数据产权分配》,《经济研究》,2023年a第7期。
- (4)李三希、武珂瑶、鲍仁杰:《大数据,个人信息保护和价格歧视——基于垂直差异化双寡头模型的分析》,《经济研究》,2021年第1期。
- (5)李三希、张仲元、焦倩:《竞争会降低企业收集消费者信息并增加隐私保护投资吗?》,《管理世界》,2023年b第7期。
- (6)刘涛雄、李若菲、戎珂:《基于生成场景的数据确权理论与分级授权》,《管理世界》,2023年第2期。
- (7)綦勇、侯泽敏、曹金霞:《差异化隐私数据保护政策下的厂商数据利用策略研究》,《中国管理科学》,2025年第3期。
- (8)邵小快、郑捷:《隐私偏好、差异化定价及消费者信息保护政策》,《经济研究》,2024年第2期。
- (9)汪敏达、李建标、陈志斌:《消费者个人信息保护与厂商广告策略的实验研究》,《中国工业经济》,2022年第4期。
- (10)武腾:《最小必要原则在企业处理个人信息实践中的适用》,《法学研究》,2021年第6期。
- (11)徐宗本、冯芷艳、郭迅华、曾大军、陈国青:《大数据驱动的管理与决策前沿课题》,《管理世界》,2014年第11期。
- (12)张剑虎、林平:《个人信息保护的市场机制与法律监管的福利效果》,《经济学(季刊)》,2023年第4期。
- (13)Acemoglu, D., Makhdomi, A., Malekian, A. and Ozdaglar, A., 2022, “Too Much Data: Prices and Inefficiencies in Data Markets”, *American Economic Journal: Microeconomics*, 14(4), pp.218~256.
- (14)Acemoglu, D., 2024, “Harms of AI”, in Bullock, J. B., eds: *The Oxford Handbook of AI Governance*, Oxford: Oxford University Press.
- (15)Acquisti, A., Taylor, C. and Wagman, L., 2016, “The Economics of Privacy”, *Journal of Economic Literature*, 54(2), pp.442~492.
- (16)Aridor, G., Che, Y. K. and Salz, T., 2023, “The Effect of Privacy Regulation on The Data Industry: Empirical Evidence from GDPR”, *RAND Journal of Economics*, 54(4), pp.695~730.
- (17)Bimpikis, K., Morgenstern, I. and Saban, D., 2024, “Data Tracking under Competition”, *Operations Research*, 72(2), pp.514~532.
- (18)Bloch, F. and Demange, G., 2018, “Taxation and Privacy Protection on Internet Platforms”, *Journal of Public Economic Theory*, 20(1), pp.52~66.



- (19) Card, D. and Krueger, A. B., 1994, "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania", *American Economic Review*, 84(4), pp.772~793.
- (20) Choi, J. P., Jeon, D. S. and Kim, B. C., 2019, "Privacy and Personal Data Collection with Information Externalities", *Journal of Public Economics*, 173, pp.113~124.
- (21) Clemens, J. and Michael, W., 2019, "The Minimum Wage and the Great Recession: Evidence of Effects on the Employment and Income Trajectories of Low-skilled Workers", *Journal of Public Economics*, 170, pp.53~67.
- (22) Cong, L. W., Xie, D. and Zhang, L., 2021, "Knowledge Accumulation, Privacy, and Growth in a Data Economy", *Management Science*, 67(10), pp.6480~6492.
- (23) Conitzer, V., Taylor, C. R. and Wagman, L., 2012, "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases", *Marketing Science*, 31(2), pp.277~292.
- (24) Conti, C. and Reverberi, P., 2021, "Price Discrimination and Product Quality under Opt-in Privacy Regulation", *Information Economics and Policy*, 55, No.100912.
- (25) Demirer, M., Hernández, D. J. J., Li, D. and Peng, S., 2024, "Data, Privacy Laws and Firm Production: Evidence from the GDPR", NBER Working Paper, No.32146.
- (26) Dosis, A. and Sand-Zantman, W., 2023, "The Ownership of Data", *Journal of Law, Economics, and Organization*, 39(3), pp.615~641.
- (27) Fainmesser, I. P. Galeotti, A., and Momot, R., 2023, "Digital Privacy", *Management Science*, 69(6), pp.3157~3173.
- (28) Goldfarb, A. and Tucker, C. E., 2011, "Privacy Regulation and Online Advertising", *Management Science*, 57(1), pp.57~71.
- (29) Hagiu, A. and Wright, J., 2023, "Data-enabled Learning, Network Effects, and Competitive Advantage", *RAND Journal of Economics*, 54(4), pp.638~667.
- (30) Ichihashi, S., 2020, "Online Privacy and Information Disclosure by Consumers", *American Economic Review*, 110(2), pp.569~595.
- (31) Ichihashi, S., 2021, "The Economics of Data Externalities", *Journal of Economic Theory*, 196, No.105316.
- (32) Jones, C. I. and Tonetti, C., 2020, "Nonrivalry and the Economics of Data", *American Economic Review*, 110(9), pp.2819~2858.
- (33) Lynch, C., 2008, "How Do Your Data Grow?", *Nature*, 455(7209), pp.28~29.
- (34) Montes, R., Sand-Zantman, W. and Valletti, T., 2019, "The Value of Personal Information in Online Markets with Endogenous Privacy", *Management Science*, 65(3), pp.1342~1362.
- (35) Peukert, C., Bechtold, S., Batikas, M. and Kretschmer, T., 2022, "Regulatory Spillovers and Data Governance: Evidence from the GDPR", *Marketing Science*, 41(4), pp.746~768.
- (36) Posner, R. A., 2014, *Economic Analysis of Law*, Valencia: Aspen Publishing.
- (37) Romanyuk, G. and Smolin, A., 2019, "Cream Skimming and Information Design in Matching Markets", *American Economic Journal: Microeconomics*, 11(2), pp.250~276.
- (38) Shapiro, C. and Varian, H. R., 1999, *Information rules: A Strategic Guide to the Network Economy*, Harvard Business Press.
- (39) Shen, Q. and Miguel Villas-Boas, J., 2018, "Behavior-based Advertising", *Management Science*, 64(5), pp.2047~2064.
- (40) Stucke, M. E. and Grunes, A. P., 2016, *Big Data and Competition Policy*, Oxford: Oxford University Press.
- (41) Varian, H. R., 2019, "Artificial Intelligence, Economics, and Industrial Organization", in Agrawal, A., Gans, J., and Goldfarb, A., eds: *The Economics of Artificial Intelligence: An Agenda*, Chicago: University of Chicago Press.

## Data Collection Terms, Personal Information Protection and Consumer Welfare

Liu Zhengchi<sup>a</sup>, Ye Yuyang<sup>a</sup> and Nie Huihua<sup>b,c</sup>

(a. School of Economics and Trade, Hunan University; b. School of Economics, Renmin University of China;

c. Center for Firm and Organization Studies, Renmin University of China)

**Abstract:** Balancing data collection and personal information protection has become a critical issue in the digital economy era. From the perspective of consumer welfare, this paper examines whether enterprises should be restricted in mandatorily collecting consumer data. By modeling the gains and costs for enterprises and consumers under different data collection terms, the study finds that enterprises choose data collection terms based on the marginal private and public benefits of data when unregulated by personal information protection laws. Welfare analysis reveals that although imposing restrictions on consumer rights, the mandatory collection terms may increase consumer welfare by increasing the public benefits of data compared to the voluntary collection terms. This paper then explores how personal information protection regulations on the mandatory collection terms affect consumer welfare, concluding that low-intensity regulations can achieve Pareto improvements in consumer welfare in certain data collection scenarios, whereas further increase of regulatory intensity fails to yield additional Pareto improvements. The policy implication is that regulators should promote reform for market-based allocation of data factors and appropriately relax restrictions on the mandatory collection terms.

**Keywords:** data collection terms; voluntary collection; mandatory collection; personal information protection; consumer welfare

# Data Collection Terms, Personal Information Protection and Consumer Welfare

*Liu Zhengchi<sup>a</sup>, Ye Yuyang<sup>a</sup> and Nie Huihua<sup>b,c</sup>*

(a. School of Economics and Trade, Hunan University; b. School of Economics, Renmin University of China;

c. Center for Firm and Organization Studies, Renmin University of China)

**Summary:** In common consumer data collection scenarios, the data collection terms that data collectors (enterprises) may adopt are generally divided into two categories. One is the voluntary collection terms: if consumers refuse data collection, the enterprise will not restrict their use of the service, and may even provide subsidies to consumers who agree to data collection. The other is the mandatory collection terms: only consumers who agree to data collection can fully use the services, otherwise they will be restricted or prohibited from using the services. The mandatory collection terms are regulated by personal information protection laws and policies. The question we are interested in is whether certain enterprises should be allowed to implement mandatory data collection terms? If the law prohibits enterprises from using mandatory collection terms, will it truly increase consumer welfare?

To address the above question, this paper constructs a two-stage game model involving a monopolistic enterprise and several heterogeneous consumers to explore the impact of personal information protection laws on consumer welfare. The enterprise collects consumers' data and uses the data to provide services to them. When consumers use the service, the enterprise can get profit. Data collection brings benefits and costs to consumers. The benefits of data bringing to consumers include two parts. One is the benefit that data directly brings to the data providers themselves, which is called data private benefit. The other is the non-competitive benefits brought by data accumulation, known as data public benefits, which can be shared by all consumers. In addition, data collection also brings different costs to each consumer.

The main finding of this paper is that the marginal private benefits and marginal public benefits of data will affect the choice of data collection terms by enterprises. Given the marginal public benefit of data, when the marginal private benefit of data exceeds a threshold, the enterprise will choose the mandatory collection terms. otherwise, the enterprise will choose the voluntary collection terms and provide subsidies when the marginal public benefit of data is high. The analysis of consumer welfare shows that in some cases, the overall welfare of consumers under mandatory collection terms is higher than that under voluntary collection terms. The reason is that data accumulation can increase the public benefits of data, while mandatory collection terms as a mechanism to reduce consumer free riding are beneficial for improving overall consumer welfare. In addition, heterogeneity analysis indicates that mandatory collection terms are more likely to increase the welfare of low-cost consumers. Therefore, prohibiting companies from adopting mandatory collection terms may harm the welfare of low-cost consumers.

The contribution of this paper lies in two aspects. Firstly, this paper incorporates the selection of data collection terms for enterprises into their decision-making process, and provides an explanation for the differentiated data collection terms selection of enterprises in reality, thus expanding the economic research on enterprise data collection. Secondly, this paper finds that in certain data collection scenarios, weaker regulatory efforts can achieve Pareto improvement in consumer welfare. This conclusion provides theoretical guidance for relaxing personal information protection regulations and data classification governance.

**Keywords:** data collection terms; voluntary collection; mandatory collection; personal information protection; consumer welfare

**JEL Classification:** L51, L86

## 附录1 《个保法》第十六条对强制收集条款的规制

《中华人民共和国个人信息保护法》第十六条规定:“个人信息处理者不得以个人不同意处理其个人信息或者撤回同意为由,拒绝提供产品或者服务;处理个人信息属于提供产品或者服务所必要的除外。”按照该条规定,如果个人信息是必要的,那么企业可以以个人不同意处理其个人信息或者撤回同意为由,拒绝提供产品或者服务,即可以采用强制收集条款;如果个人信息是非必要的,那么企业不得采用强制收集条款,而只能采用自愿收集条款。尽管为强制收集条款留下了空间,但许多司法实践中对数据“必要性”的认定标准比较严格,倾向于禁止强制收集条款。

在司法实践中,对个人信息的“必要性”的解释可以归纳为:如果某个消费者不提交个人信息,则企业无法实现该消费者所需求的服务,那么此时收集个人信息是必要的。例如,在上诉人胡红芳诉上海携程商务有限公司(以下简称携程公司)侵权责任纠纷一案中,上诉人要求携程公司停止收集上诉人非必要个人信息,增加用户不同意其现有《服务协议》《隐私政策》仍可继续使用携程APP的选项。对于此项主张,《胡红芳、上海携程商务有限公司侵权责任纠纷二审民事判决书》中有如下判决:

……就本案而言,携程公司如果不收集胡红芳的姓名、手机、身份信息等就无法实现酒店预订的服务事项。如允许消费者不提供任何个人信息,仍可使用携程APP要求携程公司提供服务,系对携程公司的过分苛责。企业经营者会因此丧失支撑起经营模式的基础资源,也无法通过合理的数据处理更好地保障消费者利益,如此既不利于数字经济的长远发展,最终也会损害消费者权益。

从该原则出发,可以得到如下推论:如果消费者未提交个人信息,或消费者可以随意更改个人信息的情况下,企业仍可以向消费者提供服务,那么此时收集个人数据是非必要的。例如在原告罗懿与被告北京大生知行科技有限公司(以下简称大生知行公司)隐私权、个人信息保护纠纷一案中,原告主张被告运营的51Talk网站和“多说英语”软件未经告知个人信息收集政策,强制收集用户年龄段、英语水平、学习需求等画像标签信息,对于该主张,《罗某与北京某公司一审民事判决书》中给出了以下判决:

……从涉案软件或网站功能设置本身上看,履行合同所必须的范围,应限定在软件或网络运营者提供的基本服务功能,或用户在有选择的基础上自主选择增加的附加功能。若收集的个人信息与该项基础服务和附加功能有直接关联,缺乏上述个人信息将导致相关功能无法实现,“履行服务所必须”之必要性才得以成立。本案中,被告抗辩其收集目的为针对不同用户需求、精准推送适合用户的个性化课程,属于为增进用户体验优化设置的一种信息推送模式,然而,被告软件或网站提供服务的基本功能为提供在线课程视频流和相关图文、视频等信息,对用户画像信息进行收集的目的并非用于支撑其基础服务功能,亦无证据表明本案原告曾自主选择使用该项优化设置功能,故被告以其软件或网站功能实现为由实施收集行为的依据不足。

而在上诉人王某与被上诉人深圳市腾讯计算机系统有限公司(以下简称腾讯公司)个人信息保护纠纷一案中,王某主张微视将对用户昵称、头像、性别、地区信息的收集使用设置为“强制选项”违法且侵害上诉人合法权益。对于此项主张,《王某、深圳市腾讯计算机系统有限公司个人信息保护纠纷民事二审民事判决书》中有如下判决:

……从举证责任分配上,被上诉人腾讯公司作为微视APP的运营者,应当证明若缺少用户“地区、性别”信息,微视APP将无法正常使用向用户提供服务,没有上述个人信息的参与,产品或服务的现有功能无法实现。仅有改善服务质量或提升用户体验及提供更好的服务不足以构成强制要求用户同意收集个人信息的理由……

从本案查明的事实看,微视APP获取用户“地区、性别”个人信息,但同时又允许用户在微信和微视APP随意更改和填写,说明“地区、性别”信息的准确性、一致性并非使用微视APP服务所必要。

据此,本院认为被上诉人腾讯公司其运营的微视APP强制获取用户地区、性别信息未满足互联网企业收集处理用户个人信息的必要性原则……

不难看出,以上案例倾向于根据数据对于数据提供者本人的边际私人收益来判断数据是否“非必要”,从而确定对强制收集条款的规制范围。

## 附录2 部分命题、引理、推论的证明

## 2.1 引理1的证明:

由式(11)易知,满足净收益最大化的补贴额度 $p^* = \frac{1}{2}(\alpha(\lambda-1) + \beta\lambda)$ 。可知当 $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$ 时, $p^* \leq 0$ ,此时企业应该令 $p=0$ 。  
 $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$ 时,此时企业应该令 $p > 0$ 。证毕。

## 2.2 引理2的证明:

若企业采用自愿收集条款且 $p^* > 0$ ,则由式(12)可知, $p^* = \frac{1}{2}(\alpha(\lambda-1) + \beta\lambda)$ 。于是,

$$\frac{\partial p^*}{\partial \lambda} = \frac{\partial \left( \frac{1}{2}(\alpha(\lambda-1) + \beta\lambda) \right)}{\partial \lambda} = \frac{1}{2}(\alpha + \beta) > 0 \quad (\text{A.1})$$

$$\frac{\partial p^*}{\partial \beta} = \frac{\partial \left( \frac{1}{2}(\alpha(\lambda-1) + \beta\lambda) \right)}{\partial \beta} = \frac{1}{2}\lambda > 0 \quad (\text{A.2})$$

$$\frac{\partial p^*}{\partial \alpha} = \frac{\partial \left( \frac{1}{2}(\alpha(\lambda-1) + \beta\lambda) \right)}{\partial \alpha} = \frac{1}{2}(\lambda-1) < 0 \quad (\text{A.3})$$

证毕。

## 2.3 引理3的证明:

由式(13)可知,当 $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$ 时



$$\frac{\partial \bar{d}^r}{\partial \alpha} = 1 \quad (\text{A.4})$$

$$\text{当 } \frac{(1-\lambda)\alpha}{\lambda} < \beta < 1 \text{ 时}$$

$$\frac{\partial \bar{d}^r}{\partial \alpha} = \frac{\partial \left( \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda) \right)}{\partial \alpha} = \frac{1}{2}(\lambda+1) \quad (\text{A.5})$$

由于  $0 < \lambda < 1$ , 可知  $0 < \frac{1}{2}(\lambda+1) < 1$ 。证毕。

#### 2.4 引理4的证明:

由等式(19)求企业净收益关于  $\theta$  的一阶偏导并化简可得:

$$\frac{\partial \Pi^m}{\partial \theta} = -\frac{\lambda\alpha\beta(2\alpha+\beta-1-\beta(1-\theta)(1-\theta))}{(1-\beta(1-\theta))^3} \quad (\text{A.6})$$

因为  $-\frac{\lambda\alpha\beta}{(1-\beta(1-\theta))^3} < 0$ , 可知当  $2\alpha+\beta-1 \leq 0$  时,  $\frac{\partial \Pi^m}{\partial \theta} \leq 0$  在  $\theta \in [0, 1]$  时恒成立, 则使企业净收益最大化的  $\theta^* = 0$ ; 当  $2\alpha+\beta > 1$  时, 若

$0 \leq \theta < \frac{1-2\alpha-\beta^2}{(1-\beta)\beta}$ , 则  $\frac{\partial \Pi^m}{\partial \theta} < 0$ , 若  $\frac{1-2\alpha-\beta^2}{(1-\beta)\beta} < \theta \leq 1$ , 则  $\frac{\partial \Pi^m}{\partial \theta} > 0$ , 企业净收益随  $\theta$  先递减后递增, 因此满足净收益最大化的  $\theta^*$  必为 0 或 1。考虑到  $\theta=1$  时, 企业的数据收集条款等同于  $p=0$  的自愿收集条款, 因此当企业选择强制收集条款时, 会且仅会选择  $\theta$  最低的强制收集条款。证毕。

#### 2.5 引理5的证明:

由等式(20)可知,

$$\frac{\partial \bar{d}^m}{\partial \alpha} = \frac{\partial \left( \frac{\alpha}{1-\beta} \right)}{\partial \alpha} = \frac{1}{1-\beta} \quad (\text{A.7})$$

由于  $0 < \beta < 1$ , 则  $\frac{1}{1-\beta} > 1$ 。于是引理4得证。证毕。

#### 2.6 命题1的证明:

为了证明命题1, 需要分别讨论企业在  $p=0$  的自愿收集条款、和  $p=\frac{1}{2}(\alpha(\lambda-1)+\beta\lambda)$  的自愿收集条款以及完全禁止的强制收集条款三者之间的净收益大小, 并得到满足企业净收益最大化的数据收集条款选择

(1) 首先分析企业在  $p=0$  的自愿收集条款和  $p=\frac{1}{2}(\alpha(\lambda-1)+\beta\lambda)$  的自愿收集条款中的选择。由引理1可知, 若  $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$ , 即  $\frac{\beta\lambda}{1-\lambda} < \alpha < 1$ , 企业会在二者之中选择  $p=\frac{1}{2}(\alpha(\lambda-1)+\beta\lambda)$  的自愿收集条款;  $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$ , 企业会在二者之中选择  $p=0$  的自愿收集条款。

(2) 其次分析当  $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$  时, 企业在  $p=0$  的自愿收集条款和完全禁止的强制收集条款中的选择。根据式(14)和式(21), 如果要满足  $\Pi^r - \Pi^m > 0$ , 可得:

$$\alpha < \frac{(1-\beta)^2}{2-\beta} \quad (\text{A.8})$$

可知满足  $0 < \alpha < \frac{(1-\beta)^2}{2-\beta}$  时,  $p=0$  的自愿收集条款优于完全禁止的强制收集条款;  $\frac{(1-\beta)^2}{2-\beta} < \alpha < 1$  时,  $p=0$  的自愿收集条款劣于完全禁止的强制收集条款。

(3) 最后分析当  $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$  时, 企业在  $p>0$  的自愿收集条款和完全禁止的强制收集条款中的选择。如果要满足  $\Pi_1 - \Pi_2 > 0$ , 即  $\frac{1}{4}(\alpha + (\alpha+\beta)\lambda)^2 - \frac{\alpha^2\lambda}{(1-\beta)^2} > 0$ 。求解可知, 当  $0 < \beta \leq (1-\sqrt{\lambda})^2$  时,  $\frac{1}{4}(\alpha + (\alpha+\beta)\lambda)^2 - \frac{\alpha^2\lambda}{(1-\beta)^2} > 0$  在定义域  $\alpha \in (0, 1-\beta)$  内恒成立; 当  $(1-\sqrt{\lambda})^2 < \beta < 1$  时, 方程  $\frac{1}{4}(\alpha + (\alpha+\beta)\lambda)^2 - \frac{\alpha^2\lambda}{(1-\beta)^2} = 0$  在定义域  $\alpha \in (0, 1-\beta)$  内有唯一解:

$$\alpha = \frac{\lambda\beta(1-\beta)(2\sqrt{\lambda} + (1-\beta)(1+\lambda))}{4\lambda - (1-\beta)^2(1+\lambda)^2} \quad (\text{A.9})$$

于是  $0 < \beta \leq (1-\sqrt{\lambda})^2$  时,  $p>0$  的自愿收集条款优于完全禁止的强制收集条款; 当  $(1-\sqrt{\lambda})^2 < \beta < 1$  时, 若  $0 < \alpha < \frac{\lambda\beta(1-\beta)(2\sqrt{\lambda} + (1-\beta)(1+\lambda))}{4\lambda - (1-\beta)^2(1+\lambda)^2}$ ,  $p>0$  的自愿收集条款优于完全禁止的强制收集条款, 若  $\frac{\lambda\beta(1-\beta)(2\sqrt{\lambda} + (1-\beta)(1+\lambda))}{4\lambda - (1-\beta)^2(1+\lambda)^2} < \alpha < 1$ ,  $p>0$  的自愿收集条款劣于完全禁止的强制收集条款。

(4) 综上所述, 可知企业选择完全禁止的强制收集条款的条件是  $\hat{\alpha}_1 < \alpha < 1-\beta$ , 其中,

$$\hat{\alpha}_1 = \begin{cases} \frac{(1-\beta)^2}{2-\beta}, & 0 < \beta \leq 1-\sqrt{\lambda} \\ \frac{\lambda\beta(1-\beta)(2\sqrt{\lambda} + (1-\beta)(1+\lambda))}{4\lambda - (1-\beta)^2(1+\lambda)^2}, & 1-\sqrt{\lambda} < \beta < 1 \end{cases} \quad (\text{A.10})$$

于是命题1中的情形①得证。若  $0 < \alpha < \hat{\alpha}_1$ , 由引理1可知, 若  $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$ , 即  $\frac{\beta\lambda}{1-\lambda} \leq \alpha < 1$ , 企业会在二者之中选择  $p > 0$  的自愿收集条款;  $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$ , 企业会在二者之中选择  $p = 0$  的自愿收集条款。于是命题1中的情形②得证。证毕。

## 2.7 命题2的证明:

根据命题1, 为了证明命题2, 需要做出以下两组比较:

①当  $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$  时, 企业选择  $p = 0$  的自愿收集条款时的消费者剩余, 与企业选择完全禁止的强制收集条款的消费者剩余;

②当  $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$  时, 企业选择  $p > 0$  的自愿收集条款时的消费者剩余, 与企业选择完全禁止的强制收集条款的消费者剩余。

(1) 首先比较当  $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$  时, 企业选择  $p = 0$  的自愿收集条款时的消费者剩余, 与企业选择完全禁止的强制收集条款的消费者剩余。

当  $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$  时, 企业若选择自愿收集条款, 则补贴额度  $p = 0$ 。比较  $p = 0$  的自愿收集条款和完全禁止的强制收集条款下的消费者剩余, 如果要满足  $CS^v - CS^m > 0$ , 可得:

$$\alpha < \frac{2(1-\beta)^2}{2-\beta} \quad (\text{A.11})$$

(2) 然后比较当  $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$  时, 企业选择  $p > 0$  的自愿收集条款时的消费者剩余, 与企业选择完全禁止的强制收集条款的消费者剩余。

当  $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$  时, 企业若选择自愿收集条款, 则补贴额度  $p > 0$ 。比较  $p > 0$  的自愿收集条款和完全禁止的强制收集条款下的消费者剩余, 可知如果要满足  $CS^v - CS^m > 0$ , 则有:

$$\frac{1}{8}(\alpha + (\alpha + \beta)\lambda)(\alpha + 4\beta + (\alpha + \beta)\lambda) - \frac{\alpha^2}{2(1-\beta)^2} > 0 \quad (\text{A.12})$$

整理可得:

$$[(1+\lambda)^2(1-\beta)^2 - 4]\alpha^2 + (1+\lambda)(4+2\lambda)(1-\beta)^2\beta\alpha + \lambda(4+\lambda)(1-\beta)^2\beta^2 > 0 \quad (\text{A.13})$$

由于  $\lambda \in (0, 1)$  且  $\beta \in (0, 1)$ , 因此  $(1+\lambda)^2(1-\beta)^2 - 4 < 0$ 。易知  $\alpha = 0$  时, 不等式(A.13)成立。则若  $\alpha = 1 - \beta$  时不等式(A.13)成立, 则不等式(A.13)在  $\alpha \in (0, 1 - \beta)$  时恒成立。

将  $\alpha = 1 - \beta$  代入不等式(A.13)左边, 可得:

$$-3\beta^2 + 2(1+\lambda)\beta + (1+\lambda)^2 - 4 > 0 \quad (\text{A.14})$$

于是可以分为  $0 < \lambda < \sqrt{3} - 1$  和  $\sqrt{3} - 1 \leq \lambda < 1$  两种情形讨论:

a. 当  $0 < \lambda < \sqrt{3} - 1$  时, 不等式(A.14)必不成立。则  $\frac{1}{8}(\alpha + (\alpha + \beta)\lambda)(\alpha + 4\beta + (\alpha + \beta)\lambda) - \frac{\alpha^2}{2(1-\beta)^2}$  在  $\alpha \in (0, 1 - \beta)$  时存在且仅存在一个零点(另一个零点在区间  $(-\infty, 0)$  内)。求解可知, 该零点为:

$$h(\beta, \lambda) = \frac{(-1+\beta)^2\beta(1+\lambda)(2+\lambda) + 2\sqrt{(-1+\beta)^2\beta^2(1-2\beta(1+\lambda)^2 + \beta^2(1+\lambda)^2 + 2\lambda(3+\lambda))}}{(-3+\beta+(-1+\beta)\lambda)(1+\beta+(-1+\beta)\lambda)} \quad (\text{A.15})$$

可知当  $0 < \alpha < h(\beta, \lambda)$  时,  $CS_2 - CS_3 > 0$ ; 当  $h(\beta, \lambda) < \alpha < 1 - \beta$  时,  $CS_1 - CS_2 < 0$ 。

b. 当  $\sqrt{3} - 1 \leq \lambda < 1$  时, 方程  $-3\beta^2 + 2(1+\lambda)\beta + (1+\lambda)^2 - 4 = 0$  存在两个不等的实根:

$$\frac{1+\lambda \pm 2\sqrt{(1+\lambda)^2 - 3}}{3} \quad (\text{A.16})$$

其中, 可知  $\frac{1+\lambda - 2\sqrt{(1+\lambda)^2 - 3}}{3} \in (0, 1)$ ; 当  $\lambda \in (0, \sqrt{5} - 2)$  时,  $\frac{1+\lambda + 2\sqrt{(1+\lambda)^2 - 3}}{3} \in (0, 1)$ , 由于  $\sqrt{5} - 2 < \sqrt{3} - 1$ , 因此  $\sqrt{3} - 1 \leq \lambda < 1$  时, 必有  $\frac{1+\lambda + 2\sqrt{(1+\lambda)^2 - 3}}{3} > 1$ 。因此在  $\sqrt{3} - 1 \leq \lambda < 1$  的前提下, 方程  $-3\beta^2 + 2(1+\lambda)\beta + (1+\lambda)^2 - 4 = 0$  在  $\beta \in (0, 1)$  的定义域内仅有一个实根  $\frac{1+\lambda - 2\sqrt{(1+\lambda)^2 - 3}}{3}$ 。

于是当  $\sqrt{3} - 1 \leq \lambda < 1$  时:

若  $0 < \beta < \frac{1+\lambda - 2\sqrt{(1+\lambda)^2 - 3}}{3}$ , 当  $0 < \alpha < h(\beta, \lambda)$  时满足  $CS_2 - CS_3 > 0$ ;  $h(\beta, \lambda) < \alpha < 1 - \beta$  时满足  $CS^v - CS^m < 0$ ; 若  $\frac{1+\lambda - 2\sqrt{(1+\lambda)^2 - 3}}{3} < \beta < 1$ ,  $0 < \alpha < 1 - \beta$  时总满足  $CS^v - CS^m > 0$ 。

(3) 综上所述可知:

①在  $0 < \lambda < \sqrt{3} - 1$  的前提下, 当  $0 < \alpha < h(\beta, \lambda)$  时,  $CS^v - CS^m > 0$ ; 当  $h(\beta, \lambda) < \alpha < 1 - \beta$  时,  $CS^v - CS^m < 0$ 。

②在  $\sqrt{3} - 1 \leq \lambda < 1$  的前提下, 若  $0 < \beta < \frac{1+\lambda - 2\sqrt{(1+\lambda)^2 - 3}}{3}$ , 当  $0 < \alpha < h(\beta, \lambda)$  时满足  $CS^v - CS^m > 0$ ;  $h(\beta, \lambda) < \alpha < 1 - \beta$  时满足  $CS^v - CS^m < 0$ ;

③在 $\sqrt{3}-1 \leq \lambda < 1$ 的前提下,若 $\frac{1+\lambda-2\sqrt{(1+\lambda)^2-3}}{3} < \beta < 1, 0 < \alpha < 1-\beta$ 时总满足 $CS^e - CS^m > 0$ 。

证毕。

## 2.8 命题3的证明:

(1)首先证明命题3情形①。

根据式(25)和式(27),可知当 $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$ 时,

a. 若 $0 < c \leq \alpha$ ,必有 $\alpha + \alpha\beta - c < \alpha + \frac{\alpha\beta}{1-\beta} - c$ ,即 $u_c^e < u_c^m$ 。

b. 若 $\frac{\alpha}{1-\beta} < c \leq 1$ ,必有 $\alpha\beta > 0$ ,即 $u_c^e > u_c^m$ 。

c. 若 $\alpha < c \leq \frac{\alpha}{1-\beta}$ ,当 $u_c^e > u_c^m$ ,则须满足 $\alpha\beta > \alpha + \frac{\alpha\beta}{1-\beta} - c$ ,即 $c > \frac{(1-\beta+\beta^2)\alpha}{1-\beta}$ 。由于 $\alpha \in (0, 1-\beta)$ ,且 $1-\beta < 1-\beta+\beta^2 < 1$ ,所以 $\frac{(1-\beta+\beta^2)\alpha}{1-\beta} \in \left(\alpha, \frac{\alpha}{1-\beta}\right)$ 。于是 $\alpha < c < \frac{(1-\beta+\beta^2)\alpha}{1-\beta}$ 时, $u_c^e < u_c^m$ ;  $\frac{(1-\beta+\beta^2)\alpha}{1-\beta} < c \leq \frac{\alpha}{1-\beta}$ 时, $u_c^e > u_c^m$ 。

综上所述, $0 \leq c < \frac{(1-\beta+\beta^2)\alpha}{1-\beta}$ 时, $u_c^e < u_c^m$ ;  $\frac{(1-\beta+\beta^2)\alpha}{1-\beta} < c \leq 1$ 时, $u_c^e > u_c^m$ 。于是命题3情形①得证。

(2)然后证明命题3情形②。

考虑自愿收集条款( $p > 0$ )下对是否同意采集数据无差异的消费者 $\tilde{c}$ ,并对比其在自愿收集条款( $p > 0$ )和完全禁止的强制收集条款下的效用。若 $u_c^e(c = \tilde{c}) > u_c^m(c = \tilde{c})$ ,则当 $0 \leq c \leq 1$ 时有 $u_c^e > u_c^m$ 。若 $u_c^e(c = \tilde{c}) = u_c^m(c = \tilde{c})$ ,由于 $\tilde{c} = \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda)$ ,则当 $0 \leq c \leq \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda)$ 时有 $u_c^e = u_c^m$ ,当 $\frac{1}{2}(\alpha(\lambda+1) + \beta\lambda) < c \leq 1$ 时有 $u_c^e > u_c^m$ 。若 $u_c^e(c = \tilde{c}) < u_c^m(c = \tilde{c})$ ,则存在一个 $\tilde{c}$ ,使 $u_c^e < u_c^m$ 。当 $0 \leq c < \tilde{c}$ 时有 $u_c^e < u_c^m$ ,当 $\tilde{c} < c \leq 1$ 时有 $u_c^e > u_c^m$ 。

由于 $u_c^e(c = \tilde{c}) > 0$ ,因此只需观察 $u_c^m(c = \tilde{c}) = \alpha + \frac{\alpha\beta}{1-\beta} - \tilde{c}$ 的情况

$$u_c^e(c = \tilde{c}) - u_c^m(c = \tilde{c}) = \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda)\beta - \left[\alpha + \frac{\alpha\beta}{1-\beta} - \frac{1}{2}(\alpha(\lambda+1) + \beta\lambda)\right]$$

可知,当 $u_c^e(c = \tilde{c}) - u_c^m(c = \tilde{c}) > 0$ 时,有 $\alpha < \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ 。

易证 $0 < \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)} < \frac{\beta\lambda}{1-\lambda}$ 。因此 $0 < \alpha < \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ 时, $u_c^e(c = \tilde{c}) > u_c^m(c = \tilde{c})$ ;  $\frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)} < \alpha < \frac{\beta\lambda}{1-\lambda}$ 时, $u_c^e(c = \tilde{c}) < u_c^m(c = \tilde{c})$ 。

求解 $\tilde{c}$ 可知:

$$\tilde{c} = \frac{2\alpha - (\alpha(\lambda+1) + \beta\lambda)(1-\beta)\beta}{2(1-\beta)} \quad (\text{A.17})$$

易证 $0 < \tilde{c} \leq \tilde{c}^m = \frac{\alpha}{1-\beta}$ 。由此命题3情形②得证。

综上,命题3证毕。

## 2.9 命题4的证明:

### 2.9.1 证明命题4情形①

显然的是,若实施个人信息保护前企业选择自愿收集条款,则实施个人信息保护后企业的条款选择不会改变。由于企业是理性的,在实施个人信息保护前,如果企业放弃了 $\theta \leq \hat{\theta}$ 的强制收集条款,而选择了自愿收集条款,那么当实施个人信息保护后,企业依然保持原先的条款选择。于是命题4情形①得证。

### 2.9.2 证明命题4情形②和情形③(其一):当 $0 < \beta \leq \frac{(1-\lambda)\alpha}{\lambda}$ 时

显然的是,实施个人信息保护会让企业在某些类型数据收集情形下放弃完全禁止的强制收集条款,转而采用自愿收集条款。在图1中企业选择强制收集条款的区域,完全禁止的强制收集条款必然是企业净收益最大化的选择,当法律规定 $\theta \geq \hat{\theta}$ 后,企业采用强制收集条款的净收益必然减少。而在图1中选择强制收集条款区域与选择自愿收集条款区域的分界线上, $\theta = 0$ 的强制收集条款和自愿收集条款对于企业是无差异的。因此当个人信息保护法律规定 $\theta \geq \hat{\theta}$ 后,企业在图1的分界线上必然会放弃强制收集条款,转而选择自愿收集条款。

根据引理5,当实施个人信息保护时,若企业选择强制收集条款,必会选择限制使用的强制收集条款且 $\theta = \hat{\theta}$ 。因此,只需计算实施个人信息保护时企业选择自愿收集条款与选择限制使用的强制收集条款( $\theta = \hat{\theta}$ )两个区域的分界线。

企业选择限制使用的强制收集条款( $\theta = \hat{\theta}$ )的净收益为:

$$\Pi^m(\theta = \hat{\theta}) = \frac{\lambda\alpha(\alpha + \beta\hat{\theta}(1-\beta(1-\hat{\theta})))}{(1-\beta(1-\hat{\theta}))^2} \quad (\text{A.18})$$

根据式(14)与式(A.18),比较自愿收集条款( $p=0$ )下的企业净收益与限制使用的强制收集条款( $\theta = \hat{\theta}$ )下的企业净收益,可知当 $\Pi^e - \Pi^m(\theta = \hat{\theta}) > 0$ 时有:



$$\alpha < \frac{(1-\beta)(1-\beta(1-\hat{\theta}))}{2-\beta(1-\hat{\theta})} \quad (\text{A.19})$$

即  $\alpha \in \left[0, \frac{(1-\beta)(1-\beta(1-\hat{\theta}))}{2-\beta(1-\hat{\theta})}\right)$  时, 企业会选择自愿收集条款 ( $p=0$ );  $\alpha \in \left[\frac{(1-\beta)(1-\beta(1-\hat{\theta}))}{2-\beta(1-\hat{\theta})}, 1\right)$  时, 企业会选择限制使用的强制收集条款 ( $\theta=\hat{\theta}$ )。

2.9.3 证明命题 4 情形②和情形③(其二): 当  $\frac{(1-\lambda)\alpha}{\lambda} < \beta < 1$  时

此处证明较为繁琐, 需要首先比较企业在不同数据收集条款下的净收益大小, 得到不等式解的一般形式; 然后排除其中无意义的解, 从而证明命题 4。

比较自愿收集条款 ( $p>0$ ) 下的企业净收益与限制使用的强制收集条款 ( $\theta=\hat{\theta}$ ) 下的企业净收益, 当  $\Pi^* - \Pi^m(\theta=\hat{\theta}) \geq 0$  时有:

$$\left((1-\beta(1-\hat{\theta}))^2(1+\lambda)^2 - 4\lambda\right)\alpha^2 + 2\left((1-\beta(1-\hat{\theta}))(1+\lambda) - 2\hat{\theta}\right)(1-\beta(1-\hat{\theta}))\beta\lambda\alpha + (1-\beta(1-\hat{\theta}))^2\beta^2\lambda^2 \geq 0 \quad (\text{A.20})$$

做如下设定:

$$k_1 = (1-\beta(1-\hat{\theta}))^2(1+\lambda)^2 - 4\lambda \quad (\text{A.21})$$

$$k_2 = 2\left((1-\beta(1-\hat{\theta}))(1+\lambda) - 2\hat{\theta}\right)(1-\beta(1-\hat{\theta}))\beta\lambda \quad (\text{A.22})$$

$$k_3 = (1-\beta(1-\hat{\theta}))^2\beta^2\lambda^2 \quad (\text{A.23})$$

求不等式 (A.20) 的解, 可得引理 A1。

引理 A1 (不等式 (A.20) 的解):

① 在  $\beta > \frac{(1-\sqrt{\lambda})^2}{(1-\hat{\theta})(1+\lambda)}$  的前提下,  $\alpha \in \left(0, \frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1}\right)$  时, 有  $\Pi^* \geq \Pi^m(\theta=\hat{\theta})$ 。

② 在  $\beta < \frac{(1-\sqrt{\lambda})^2}{(1-\hat{\theta})(1+\lambda)}$  的前提下, 若  $\beta \leq \max\left\{\frac{1+\lambda-2\hat{\theta}}{(1-\hat{\theta})(1+\lambda)}, \frac{\lambda+\hat{\theta}}{\hat{\theta}(1+\lambda)}\right\}$ , 则  $\alpha \in (0, 1-\beta)$  时有  $\Pi^* \geq \Pi^m(\theta=\hat{\theta})$ ; 若  $\beta > \max\left\{\frac{1+\lambda-2\hat{\theta}}{(1-\hat{\theta})(1+\lambda)}, \frac{\lambda+\hat{\theta}}{\hat{\theta}(1+\lambda)}\right\}$ ,

则  $\alpha \leq \frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1}$  或  $\alpha \geq \frac{-k_2 + \sqrt{k_2^2 - 4k_1k_3}}{2k_1}$  时, 有  $\Pi^* \geq \Pi^m(\theta=\hat{\theta})$ 。

③ 在  $\beta = \frac{(1-\sqrt{\lambda})^2}{(1-\hat{\theta})(1+\lambda)}$  的前提下, 若  $\beta \leq \frac{1+\lambda-2\hat{\theta}}{(1-\hat{\theta})(1+\lambda)}$ , 则  $\alpha \in (0, 1-\beta)$  时有  $\Pi^* \geq \Pi^m(\theta=\hat{\theta})$ ; 若  $\beta > \frac{1+\lambda-2\hat{\theta}}{(1-\hat{\theta})(1+\lambda)}$ , 则  $\alpha \in \left(0, \frac{k_3}{-k_2}\right)$  时, 有

$\Pi^* \geq \Pi^m(\theta=\hat{\theta})$ 。

接下来在引理 A1 的基础上, 排除其中无意义的解, 进而证明命题 4。

易证  $\beta \in (0, 1-\lambda)$  时, 企业选择自愿收集条款 ( $p=0$ ) 和限制使用的强制收集条款 ( $\theta=\hat{\theta}$ ) 分界线  $\alpha = \frac{(1-\beta)(1-\beta(1-\hat{\theta}))}{2-\beta(1-\hat{\theta})}$  与  $\alpha = \frac{\beta\lambda}{1-\lambda}$  在有且仅有一个交点。设该交点横坐标为  $\tilde{\beta}$ 。接下来分别讨论  $\tilde{\beta} < \beta < 1$  和  $0 < \beta \leq \tilde{\beta}$  的情形。

(1) 证明若  $\tilde{\beta} < \beta < 1$ , 实施个人信息保护后, 当  $\alpha \in (0, \hat{\alpha}_3)$  时, 企业选择自愿收集条款; 当  $\alpha \in (\hat{\alpha}_3, 1-\beta)$  时, 企业选择限制使用的强制收集条款 ( $\theta=\hat{\theta}$ )。其中  $\hat{\alpha}_3 = \min\left\{\frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1}, 1-\beta\right\}$ 。

(2) 证明若  $0 < \beta \leq \tilde{\beta}$ , 实施个人信息保护后, 当  $\alpha \in (0, \hat{\alpha}_3)$  时, 企业选择自愿收集条款; 当  $\alpha \in (\hat{\alpha}_3, 1-\beta)$  时, 企业选择限制使用的强制收集条款 ( $\theta=\hat{\theta}$ )。其中  $\hat{\alpha}_3 = \frac{(1-\beta)(1-\beta(1-\hat{\theta}))}{2-\beta(1-\hat{\theta})}$ 。

此处计算较为繁琐, 限于篇幅省略过程, 有需要的读者可联系作者。综上, 可得引理 A2。

引理 A2 (实施规制力度  $\hat{\theta}$  的个人信息保护后企业的数据收集条款选择):

当  $\alpha \in (0, \hat{\alpha}_3)$  时, 企业选择自愿收集条款; 当  $\alpha \in (\hat{\alpha}_3, 1-\beta)$  时, 企业选择限制使用的强制收集条款 ( $\theta=\hat{\theta}$ )。其中,

$$\hat{\alpha}_3 = \begin{cases} \frac{(1-\beta)(1-\beta(1-\hat{\theta}))}{2-\beta(1-\hat{\theta})}, & 0 < \beta \leq \tilde{\beta} \\ \min\left\{\frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1}, 1-\beta\right\}, & \tilde{\beta} < \beta < 1 \end{cases} \quad (\text{A.24})$$

根据引理 A2, 不难得到命题 4。证毕。

## 2.10 命题 5 的证明

限制使用的强制收集条款 ( $\theta=\hat{\theta}$ ) 下, 消费者剩余为:

$$CS^m(\theta=\hat{\theta}) = \frac{\alpha(\alpha+2\beta\hat{\theta}(1-\beta(1-\hat{\theta})))}{2(1-\beta(1-\hat{\theta}))^2} \quad (\text{A.25})$$

若限制使用的强制收集条款下消费者剩余 $CS^m(\theta=\hat{\theta})$ 高于完全禁止的强制收集条款下消费者剩余 $CS^m(\theta=0)$ ,则满足:

$$\alpha < \frac{2(1-\beta(1-\hat{\theta}))(1-\beta)^2}{2-\beta(1-\hat{\theta})-\beta} \quad (\text{A.26})$$

不难证明 $0 < \frac{2(1-\beta(1-\hat{\theta}))(1-\beta)}{2-\beta(1-\hat{\theta})-\beta} < 1$ ,因此可知 $0 < \frac{2(1-\beta(1-\hat{\theta}))(1-\beta)^2}{2-\beta(1-\hat{\theta})-\beta} < 1-\beta$ 。于是有引理A3。

引理A3:对于任意 $\hat{\theta} \in (0, 1]$ ,当数据边际私人收益较低时( $0 < \alpha < \frac{2(1-\beta(1-\hat{\theta}))(1-\beta)^2}{2-\beta(1-\hat{\theta})-\beta}$ ),限制使用的强制收集条款下消费者剩余高

于完全禁止的强制收集条款下的消费者剩余,即 $CS^m(\theta=\hat{\theta}) > CS^m(\theta=0)$ ;当数据边际私人收益较高时( $\frac{2(1-\beta(1-\hat{\theta}))(1-\beta)^2}{2-\beta(1-\hat{\theta})-\beta} < \alpha < 1-\beta$ ),限

制使用的强制收集条款下消费者剩余低于完全禁止的强制收集条款下消费者剩余,即 $CS^m(\theta=\hat{\theta}) < CS^m(\theta=0)$ 。

根据实施个人信息保护前后企业数据收集条款选择的变化,个人信息保护对消费者福利的影响分为3种情形:

①当 $\alpha \in (0, \hat{\alpha}_1)$ 时,实施个人信息保护前,企业选择自愿收集条款;实施个人信息保护后,企业依然选择自愿收集条款。这种情形下,消费者福利不变。该情形对应图4a的加粗线内区域OBC。

②当 $\alpha \in (\hat{\alpha}_1, \min\{\hat{\alpha}_3, 1-\beta\})$ 时,实施个人信息保护前,企业选择完全禁止的强制收集条款( $\theta=0$ );实施个人信息保护后,企业转而选择自愿收集条款。该情形对应图4b的加粗线内区域BCD。根据命题2,若 $\alpha \in (\hat{\alpha}_1, \min\{\hat{\alpha}_3, 1-\beta\})$ ,消费者整体福利增加(图4b的阴影区域BCE);否则,消费者福利减少(图4b的区域BDE)。

③当 $\alpha \in (\hat{\alpha}_3, 1-\beta)$ 时,实施个人信息保护前,企业选择完全禁止的强制收集条款( $\theta=0$ );实施个人信息保护后,企业选择限制使用的强制收集条款( $\theta=\hat{\theta}$ )。该情形对应图4c的加粗线内区域ACD。根据引理A3,若 $\alpha \in \left(\hat{\alpha}_3, \frac{2(1-\beta(1-\hat{\theta}))(1-\beta)^2}{2-\beta(1-\hat{\theta})-\beta}\right)$ ,消费者整体福利增加(图4c的阴影区域ACF);否则,消费者福利减少(图4c的区域ADF)。

于是命题5得证。

## 2.11 命题6的证明:

证明实施个人信息保护可以实现消费者福利的帕累托改进的必要条件为:

①企业盈利能力较强( $\frac{1}{4} < \lambda < 1$ );

②数据边际私人收益适中( $\hat{\alpha}_1 < \alpha < \hat{\alpha}_4$ )且边际公共收益较高( $\beta > \hat{\beta}_1$ );

③实施个人信息保护使企业放弃强制收集条款,转而采用自愿收集条款( $\alpha < \hat{\alpha}_3$ )。

(1)首先证明必要条件③。

考虑 $c < \alpha$ 消费者的福利。无论企业采用何种数据收集条款, $c < \alpha$ 的消费者都会同意数据收集,因此在 $p=0$ 的自愿收集条款和强制收集条款下,当 $\alpha$ 与 $\beta$ 不变时,此类消费者的福利与企业数据收集总量正相关。由式(18)可知,当企业采用强制收集条款时,若 $\theta$ 上升,企业数据收集总量减少,于是此类消费者福利减少。于是可知:

a.若在实施个人信息保护前后,企业数据收集条款不变,消费者福利无法实现帕累托改进。

b.若实施个人信息保护前,企业采用完全禁止的强制收集条款( $\theta=0$ );实施个人信息保护后,企业采用限制使用的强制收集条款( $\theta=\hat{\theta}$ ),则消费者福利无法实现帕累托改进。

因此消费者福利实现帕累托改进的必要条件是:实施个人信息保护后企业从采用完全禁止的强制收集条款( $\theta=0$ )转而采用自愿收集条款,即 $\alpha \in (\hat{\alpha}_1, \hat{\alpha}_3)$

必要条件③得证。

(2)然后证明必要条件①。

即证明当 $0 < \lambda \leq \frac{1}{4}$ 时,若企业放弃完全禁止的强制收集条款转而采取自愿收集条款,不存在消费者福利帕累托改进的可能。根据命题1和命题3,即证明 $0 < \lambda \leq \frac{1}{4}$ 时,对于 $\beta \in (1-\sqrt{\lambda}, 1)$ ,下列不等式恒成立:

$$\frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)} \leq \frac{\lambda\beta(1-\beta)(2\sqrt{\lambda}+(1-\beta)(1+\lambda))}{4\lambda-(1-\beta)^2(1+\lambda)^2}$$

化简可得

$$-\sqrt{\lambda}(\lambda+1)\beta^2 + (3\lambda+1)\beta + \lambda\sqrt{\lambda} - \sqrt{\lambda} - 1 \leq 0 \quad (\text{A.27})$$

计算可证必要条件①得证。

(3)接下来证明必要条件②。

根据命题3,若要满足所有消费者在 $p>0$ 的自愿收集条款下的福利均不低于在完全禁止的强制收集条款下的福利,则有 $\alpha \leq \hat{\alpha}_4$ 。结合对必要条件③的证明,可知 $\hat{\alpha}_1 < \alpha < \hat{\alpha}_4$ 。

证明图6中曲线BC和曲线OGB的在定义域  $\beta \in (0, 1)$  内唯一交点G。

由于  $\beta = 1 - \sqrt{\lambda}$  时, 不等式(A.27)不成立, 且  $-\sqrt{\lambda}(\lambda + 1) < 0$ , 可知当  $\lambda > \frac{1}{4}$  时,  $-\sqrt{\lambda}(\lambda + 1)\beta^2 + (3\lambda + 1)\beta + \lambda + \lambda\sqrt{\lambda} - \sqrt{\lambda} - 1$  在  $\beta \in (1 - \sqrt{\lambda}, 1)$  的定义域内有且仅有一个零点。于是可知  $\lambda > \frac{1}{4}$  时, 图6中曲线BC和曲线OGB的在定义域  $\beta \in (0, 1)$  内有且仅有唯一交点G。

令G点横坐标为  $\hat{\beta}$ , 结合命题3, 存在消费者福利帕累托改进时必有  $\beta > \hat{\beta}$ 。

必要条件②得证。

综上得证个人信息保护可以实现消费者福利的帕累托改进的必要条件①②③。

根据命题3, 不难证明条件①②③为个人信息保护可以实现消费者福利的帕累托改进的充分条件。于是命题6得证。

## 2.12 推论2的证明:

(1) 首先证明若企业放弃完全禁止的强制收集条款转而采取自愿收集条款, 消费者福利帕累托改进区域的上界低于  $\frac{\beta\lambda}{1-\lambda}$ 。即证明:

$$\frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)} < \frac{\beta\lambda}{1-\lambda}$$

化简为:

$$2(1-\beta^2) < 2$$

由于  $0 < \beta < 1$ , 于是得证。

(2) 接下来证明必然存在  $\hat{\theta} \in (0, 1)$ , 当  $\theta \geq \hat{\theta}$  时有  $\hat{\alpha}_3 > \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ 。由于  $0 < \beta \leq \hat{\beta}$  时  $\hat{\alpha}_3 = \frac{(1-\beta)(1-\beta(1-\theta))}{2-\beta(1-\theta)} > \frac{\beta\lambda}{1-\lambda} > \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$  恒成立, 因此只需证明  $\tilde{\beta} < \beta < 1$  时, 存在  $\hat{\theta} \in (0, 1)$ , 当  $\theta \geq \hat{\theta}$  时有  $\frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1} > \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ 。

易知,

$$\lim_{\theta \rightarrow 1} \frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1} = \frac{\beta\lambda}{1-\lambda}$$

按照左极限定义, 对于  $\frac{\beta\lambda}{1-\lambda} - \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)} > 0$ , 必然存在  $E > 0$ , 使得满足  $1-E < x < 1$  的一切  $x$ , 恒有  $\left| \frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1} - \frac{\beta\lambda}{1-\lambda} \right| < \frac{\beta\lambda}{1-\lambda} - \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ 。由于  $\tilde{\beta} < \beta < 1$  时  $\frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1} < \frac{\beta\lambda}{1-\lambda}$ ,  $\left| \frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1} - \frac{\beta\lambda}{1-\lambda} \right| < \frac{\beta\lambda}{1-\lambda} - \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$  即为  $\frac{-k_2 - \sqrt{k_2^2 - 4k_1k_3}}{2k_1} > \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ 。

于是必然存在  $\hat{\theta} \in (0, 1)$ , 当  $\theta \geq \hat{\theta}$  时有  $\hat{\alpha}_3 > \frac{(1-\beta^2)\beta\lambda}{2-(1-\beta^2)(\lambda+1)}$ 。推论2得证。

## 附录3 关于不同商业模式下企业收益函数的讨论

现实中, 企业通过数据获取收益的模式主要包括: (1) 向消费者直接收取服务费用或交易佣金; (2) 通过精准投放获取广告收入; (3) 将聚合数据产品化后进行交易。本文的企业收益函数比较贴合广告收入的商业模式。对于其它两种商业模式, 我们也构建相应的企业收益函数, 以检验本文结论的稳健性。

### 3.1 企业收取服务费

在这种商业模式中, 企业收益与数据公共品的价值并不直接挂钩, 而是来源于向消费者收取的服务费。在某些情形下, 消费者若拒绝数据收集, 则可能需要支付不同的价格, 但不会被限制使用服务, 即自愿收集条款。例如, 美国前进保险公司 (Progressive Insurance) 推出 Snapshot 计划, 使用安装在汽车中的设备或移动应用程序来追踪投保人的驾驶行为, 同意收集个人驾驶数据的投保人可以获得保费折扣。而在某些情形下, 消费者若拒绝数据收集, 并不会受到价格歧视, 但会被限制使用服务, 即强制收集条款。例如云服务商 Microsoft Azure 和亚马逊云服务 (AWS) 会在服务条款中要求收集用户的使用量、性能日志、错误报告等数据, 如果不同意则会被拒绝服务。

基于上述商业模式, 我们假定在自愿收集条款下, 企业可以根据消费者是否同意数据收集而设定不同额度的服务费; 而在强制收集条款下, 企业设定统一的服务费额度, 但可以限制拒绝数据收集的消费者使用服务。

接下来, 我们分别计算自愿收集条款与强制收集条款下的企业净收益和消费者福利。通过对比, 可以得知企业对数据收集条款类型的选择, 并验证强制收集条款是否可能提升消费者福利。

#### (1) 自愿收集条款。

模型时序如下:

第一阶段, 企业选择自愿收集条款, 并设定服务费额度。对同意数据收集的消费者, 设定服务费  $s_1^*$ ; 对拒绝数据收集的消费者, 设定服务费  $s_2^*$ 。

第二阶段, 各个消费者独立做出决策, 选择是否使用服务, 以及是否同意企业收集个人数据。企业和消费者进而分别获取净收益和效用。



若消费者使用服务,且同意数据收集,其效用为: $u_c^*(d_c=1)=\alpha+\beta\bar{d}-c-s_1^*$ ;若消费者使用服务,但拒绝数据收集,则 $u_c^*(d_c=0)=\beta\bar{d}-s_2^*$ ;若消费者不使用服务(当然也不会发生数据收集),则 $u_c=0$ 。

与原文模型一致, $\alpha$ 为外生常数,刻画了边际私人收益的高低; $\beta$ 为外生常数,刻画了边际公共收益的高低; $c$ 刻画了数据收集给消费者带来的成本; $\bar{d}=\int_0^1 d_c f(c)dc$ ,表示自愿收集条款下该企业对所有消费者收集数据的总量,其中 $d_c$ 为0-1变量。若消费者同意企业收集个人数据,则 $d_c=1$ ;若消费者拒绝企业收集个人数据,则 $d_c=0$ 。令 $\alpha\in(0,1),\beta\in(0,1),c\in[0,1]$ 且 $f(c)=1$ 。

当消费者使用服务且同意数据收集时,某个 $c$ 类型消费者给企业带来的收益为: $r_c=s_1^*$ ;当消费者使用服务但拒绝数据收集时, $r_c=s_2^*$ ;当消费者不使用服务时, $r_c=0$ 。

显然,企业不会对拒绝数据收集的消费者设定过高的服务费,以避免他们不使用服务从而导致企业损失部分收益。此外与原文类似,我们设定企业提供服务的边际成本为0。因此企业净收益函数为:

$$\Pi^*=\int_0^{\bar{c}} s_1^* f(c)dc+\int_0^{\bar{c}} s_2^* f(c)dc$$

其中, $\bar{c}$ 是满足 $u_c^*(d_c=1)=u_c^*(d_c=0)$ 的无差异消费者类型。于是最优服务方案是如下数学规划的解:

$$\text{Max}_{[s_1^*,s_2^*]} \int_0^{\bar{c}} s_1^* f(c)dc+\int_0^{\bar{c}} s_2^* f(c)dc$$

$$\text{s.t. } s_2^* \leq \beta\bar{d}$$

$$\text{其中, } \bar{d}=\int_0^{\bar{c}} f(c)dc。$$

计算可知企业的最优策略为设定 $s_1^*=\frac{\alpha-\beta+\beta(\alpha+\beta)}{2},s_2^*=\frac{\beta(\alpha+\beta)}{2}$ 。此时数据收集总量为 $\bar{d}=\frac{\alpha+\beta}{2}$ ,企业净收益为 $\Pi^*=\frac{(\alpha+\beta)^2}{4}$ ,消费者剩余为 $CS^*=\frac{(\alpha+\beta)^2}{8}$ 。

(2)强制收集条款。

博弈时序为:

第一阶段,企业选择强制收集条款,并设定服务费 $s^m$ 以及消费者拒绝数据收集的情形下可以使用服务的比例 $\theta$ 。

第二阶段,各个消费者独立做出决策,选择是否使用服务,以及是否同意企业收集个人数据。企业和消费者进而分别获取净收益和效用。

若消费者使用服务,且同意数据收集,则 $u_c^m(d_c=1)=\alpha+\beta\bar{d}^m-c-s^m$ ;若消费者使用服务,但拒绝数据收集,则 $u_c^m(d_c=0)=\theta\beta\bar{d}^m-s^m$ ;若消费者不使用服务,则 $u_c^m=0$ 。其中 $\bar{d}^m$ 表示强制收集条款下该企业对所有消费者收集数据的总量。

当消费者使用服务时,某个 $c$ 类型消费者给企业带来的收益为: $r_c=s^m$ 。可知企业净收益函数存在两种可能的情形:其一是拒绝数据收集的消费者仍愿意支付 $s^m$ 以使用部分服务,此时 $\Pi^m=s^m$ ;其二是拒绝数据收集的消费者不愿意支付 $s^m$ ,此时 $\Pi^m=\int_0^{\bar{c}} s^m f(c)dc$ ,其中, $\bar{c}$ 是满足 $u_c^m(d_c=1)=0$ 的无差异消费者类型。

存在两种可能的均衡,第一种是 $\theta\neq 0$ ,企业的服务费恰好使拒绝数据收集的消费者愿意使用服务,即 $s^m=\theta\beta\bar{d}^m$ 。计算可知企业的最优策略是设定 $\theta=1,s^m=\alpha\beta$ ,此时数据收集总量为 $\bar{d}^m=\alpha$ ,企业净收益为 $\Pi^m=\alpha\beta$ 。

第二种可能的均衡是 $\theta=0$ ,即企业禁止拒绝数据收集的消费者使用服务,计算可知企业的最优策略是设定 $\theta=0,s^m=\frac{\alpha}{2}$ ,此时数据收集总量为 $\bar{d}^m=\frac{\alpha}{2(1-\beta)}$ ,企业净收益为 $\Pi^m=\frac{\alpha^2}{4(1-\beta)}$ 。

比较可知,当 $\alpha\leq 4\beta(1-\beta)$ 时,企业的最优策略是设定 $\theta=1,s^m=\alpha\beta$ ,此时数据收集总量为 $\bar{d}^m=\alpha$ ,企业净收益为 $\Pi^m=\alpha\beta$ ,消费者剩余为 $CS^m=\frac{\alpha^2}{2}$ ;当 $\alpha>4\beta(1-\beta)$ 时,企业的最优策略是设定 $\theta=0,s^m=\frac{\alpha}{2}$ ,此时数据收集总量为 $\bar{d}^m=\frac{\alpha}{2(1-\beta)}$ ,企业净收益为 $\Pi^m=\frac{\alpha^2}{4(1-\beta)}$ ,消费者剩余为 $CS^m=\frac{\alpha^2}{8(1-\beta)^2}$ 。

由于 $\theta=1,s^m=\alpha\beta$ 即等同于一种 $s_1^*=s_2^*$ 的自愿收集条款,且对于企业而言,该策略劣于自愿收集条款下 $s_1^*=\frac{\alpha-\beta+\beta(\alpha+\beta)}{2},s_2^*=\frac{\beta(\alpha+\beta)}{2}$ 的策略,因此企业不会选择该策略。于是只需比较 $\theta=0,s^m=\frac{\alpha}{2}$ 的强制收集条款与自愿收集条款即可。

(3)结论。

首先对比企业净收益可知,对于边际私人收益相对较高( $\alpha>\max\{1-\beta+\sqrt{1-\beta},4\beta(1-\beta)\}$ )的数据收集场景,企业选择强制收集条款;对于边际私人收益相对较低( $\alpha<\max\{1-\beta+\sqrt{1-\beta},4\beta(1-\beta)\}$ )的数据收集场景,企业选择自愿收集条款,且当边际公共收益相对较高时( $\beta>\alpha$ ), $s_1^*<s_2^*$ ,即企业向同意数据收集的消费者提供更低的服务费。该结论与原文命题1类似。

然后对比消费者剩余,当 $\alpha>1-\beta$ 时,强制收集条款下的消费者福利高于自愿收集条款下的消费者福利。该结论与原文命题2类似,即在某些情形下,强制收集条款下的消费者整体福利高于自愿收集条款下的消费者整体福利。

于是可知,强制收集条款可能提升福利的结论,在企业收取服务费的商业模式中依然成立。因此,对强制收集条款进行规制可能损害消费者福利。

### 3.2 企业将聚合数据产品化后进行交易

在某些情形下,企业可以将聚合数据产品化,然后向市场上的其它数据需求者出售从而获取收入。为了刻画这种商业模式,参考李三希等(2021)的做法,不妨假设企业从数据获取的收益与数据收集总量成正比。那么在自愿收集条款下,企业的净收益函数为:

$$\Pi^*=\lambda\bar{d}-p\bar{d}$$

其中 $\bar{d}$ 为企业在自愿收集条款下的数据收集总量, $p$ 为补贴额度, $\lambda$ 则刻画了企业的盈利能力,假设 $0<\lambda<1$ 。

而在强制收集条款下,企业的净收益函数为:

$$\Pi^m = \lambda \bar{d}^m$$

其中 $\bar{d}^m$ 为企业在强制收集条款下的数据收集总量。模型其它设定如消费者效用函数以及博弈时序与原文一致,不再详述。

(1)自愿收集条款。

可知在自愿收集条款下,满足净收益最大化的补贴额度 $p^*$ 为:

$$p^* = \begin{cases} \frac{\lambda - \alpha}{2}, 0 < \alpha < \lambda \\ 0, \lambda \leq \alpha < 1 \end{cases}$$

企业的数据收集总量为:

$$\bar{d}^* = \begin{cases} \frac{\lambda + \alpha}{2}, 0 < \alpha < \lambda \\ \alpha, \lambda \leq \alpha < 1 \end{cases}$$

企业净收益为:

$$\Pi^* = \begin{cases} \frac{(\lambda + \alpha)^2}{4}, 0 < \alpha < \lambda \\ \lambda \alpha, \lambda \leq \alpha < 1 \end{cases}$$

消费者剩余为:

$$CS^* = \begin{cases} \frac{(\lambda + \alpha + 4\beta)(\lambda + \alpha)}{8}, 0 < \alpha < \lambda \\ \frac{\alpha(\alpha + 2\beta)}{2}, \lambda \leq \alpha < 1 \end{cases}$$

(2)强制收集条款。

在强制收集条款下,企业会设定 $\theta=0$ ,即企业禁止拒绝数据收集的消费者使用服务。此时企业的数据收集总量为:

$$\bar{d}^m = \frac{\alpha}{1-\beta}$$

企业的净收益为:

$$\Pi^m = \frac{\lambda \alpha}{1-\beta}$$

消费者剩余为:

$$CS^m = \frac{\alpha^2}{2(1-\beta)^2}$$

(3)结论。

首先对比企业净收益可知,对于边际私人收益相对较高( $\alpha > \hat{\alpha}_s$ )的数据收集场景,企业选择强制收集条款;对于边际私人收益相对较低( $\alpha < \hat{\alpha}_s$ )的数据收集场景,企业选择自愿收集条款。其中 $\alpha = \hat{\alpha}_s$ 是方程 $(\lambda - \alpha)^2 - \beta(\lambda + \alpha)^2 = 0$ 在区间 $(0, \lambda)$ 内的唯一解。该结论与原文命题1类似。但不同的是,在这种商业模式下,企业的补贴额度与数据的边际公共收益无关。

然后对比消费者剩余,当 $\alpha > \hat{\alpha}_s$ 时,强制收集条款下的消费者福利高于自愿收集条款下的消费者福利。其中 $\alpha = \hat{\alpha}_s$ 是方程 $[4 - (1-\beta)^2]\alpha^2 - (1-\beta)^2(2\lambda + 4\beta)\alpha - (1-\beta)^2(\lambda + 4\beta)\lambda = 0$ 在区间 $(0, \lambda)$ 内的唯一解。该结论与原文命题2类似,即在某些情形下,强制收集条款下的消费者整体福利高于自愿收集条款下的消费者整体福利。

于是可知,强制收集条款可能提升福利的结论,在企业将聚合数据产品化后进行交易的商业模式中依然成立。因此,对强制收集条款进行规制可能损害消费者福利。