

电商平台个性化定价的算法公平性风险与协同治理机制研究

倪佳

扬州大学商学院, 江苏 扬州

收稿日期: 2026年3月2日; 录用日期: 2026年3月16日; 发布日期: 2026年4月16日

摘要

数字经济背景下, 算法技术的迭代推动电商平台个性化定价模式广泛应用, 成为平台实现利润最大化的核心手段, 但也因定价过程的算法化、隐蔽化引发一系列公平性问题, 破坏电商生态系统的稳定运行。本文从价格歧视到动态博弈的视角界定个性化定价内涵, 厘清个性化定价、价格歧视与“大数据杀熟”的概念边界, 同时从程序、分配、信息三维度解析算法公平性的核心要义。研究系统识别出电商个性化定价中存在的分配不公、程序不公与信息不公三类核心风险, 并基于技术-组织-制度三维框架剖析风险的生成机理, 进一步探讨算法公平性缺失对消费者、平台自身及市场整体造成的多层次影响效应。最终, 本文构建起政府、平台、消费者与第三方社会力量协同参与的多维治理机制, 以期规范电商平台算法定价行为、维护数字市场公平竞争秩序、促进电商行业可持续发展提供理论参考与实践路径。

关键词

电商平台, 个性化定价, 算法公平性, 协同治理

Research on the Fairness Risk and Collaborative Governance Mechanism of Personalized Pricing Algorithm in E-Commerce Platform

Jia Ni

Business School of Yangzhou University, Yangzhou Jiangsu

Received: March 2, 2026; accepted: March 16, 2026; published: April 16, 2026

文章引用: 倪佳. 电商平台个性化定价的算法公平性风险与协同治理机制研究[J]. 电子商务评论, 2026, 15(4): 623-632.
DOI: 10.12677/ecl.2026.154438

Abstract

In the context of the digital economy, the iterative advancement of algorithmic technologies has driven the widespread adoption of personalized pricing models on e-commerce platforms, becoming a core strategy for maximizing profits. However, the algorithmic and opaque nature of pricing processes has triggered a series of fairness issues, disrupting the stable operation of e-commerce ecosystems. This paper defines the essence of personalized pricing from the perspective of price discrimination and dynamic game theory, clarifying the conceptual boundaries between personalized pricing, price discrimination, and “big data price discrimination”. It also analyzes the core principles of algorithmic fairness from three dimensions: procedural, distributive, and informational. The study systematically identifies three core risks in e-commerce personalized pricing: distributive injustice, procedural injustice, and informational injustice. Based on a three-dimensional framework encompassing technology, organization, and institutions, it examines the mechanisms underlying these risks and further explores the multi-level impacts of algorithmic fairness deficiencies on consumers, platforms, and the broader market. Ultimately, the paper proposes a multidimensional governance mechanism involving collaborative participation from governments, platforms, consumers, and third-party social forces. This framework aims to provide theoretical references and practical pathways for regulating algorithmic pricing behaviors on e-commerce platforms, maintaining fair competition in digital markets, and promoting sustainable development in the e-commerce industry.

Keywords

E-Commerce Platforms, Personalized Pricing, Algorithmic Fairness, Collaborative Governance

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在数字经济浪潮中，数据已成为关键生产要素。电商平台依托大数据与人工智能算法，精准刻画用户的消费偏好、支付能力及价格敏感度，并据此实施个性化定价。这一模式在提升资源配置效率、增强用户体验的同时，也因算法的“黑箱效应”与平台的逐利导向，引发了日益突出的算法公平性风险。2018年，携程被曝光“同房不同价”引起用户的不满。2019年，其机票预订系统中价格在短时间内剧烈波动的现象又登上热搜[1]。这些事件不仅暴露了算法定价透明度的缺失，更凸显了消费者在算法决策中知情权与公平交易权的保障困境。

从实践层面看，当前电商平台个性化定价的算法应用缺乏统一规范约束，定价行为边界模糊，现有监管体系难以适配算法技术的迭代速度。从理论层面看，现有研究多聚焦于个性化定价的经济效应或价格歧视的法律规制，针对算法技术嵌入定价过程引发的公平性风险分类、生成机理及协同治理机制的系统性研究仍有待深化。算法公平性不仅是数字经济时代市场公平原则的核心体现，更是电商生态系统可持续发展的重要基础。基于此，本文以电商平台个性化定价的算法公平性为研究核心，通过界定核心概念、识别风险类型、剖析生成机理、分析影响效应，最终构建多维协同治理机制，旨在回应数字市场发展中的现实问题，弥补现有研究的理论空白，为政府监管部门、电商平台及相关主体提供决策参考。

2. 核心概念与辨析

2.1. 个性化定价的内涵与本质

2.1.1. 个性化定价的内涵

尽管个性化定价已成为数字平台的普遍实践，但其概念内涵在学界仍存在一定分歧。2013年英国的非营利组织(OFT)将其定义为企业利用观察到的、自愿提供的、推断的或收集到的个人行为或特征信息，为不同的消费者制定企业认为他们愿意支付的不同价格[2]。此定义强调基于用户分组的差别定价，本质上对应经济学中的第三类价格歧视。然而，经合组织(OECD)在2018年发布的《数字时代的个性化定价》中将个性化定价重新界定为：“企业利用用户行为和特征信息，最终实现定价与用户个人支付意愿正相关的价格歧视行为”[3]。这一定义更具包容性，能够涵盖不同实现形式的个性化定价。

本文沿用 OECD 的这一定义框架。它不仅涵盖了显性的“直接标价差异”，更洞察到了通过红包、优惠券、积分等形成的隐性“价格组合歧视”。基于此，电商平台的个性化定价可被视为传统价格歧视在数字经济背景下的技术升维形态：它依托大数据与机器学习算法，实现对用户特征的精准捕捉与价格的动态调整，是算法、数据与市场环境共同驱动下的动态博弈过程。

2.1.2. 个性化定价的本质

(1) 价格歧视是个性化定价的经济学内核，通常分为三种：一级价格歧视，即因人定价，要求企业精准掌握每个消费者的保留价格，对每人收取其愿意支付的最高价，以攫取全部消费者剩余；二级价格歧视，即因量定价，通过设计不同的“数量-价格”组合(如会员套餐、满减梯度)，引导消费者自我选择，间接实现差异化收费；三级价格歧视，即因类定价，依据可观测的群体特征(如地域、年龄、学生身份)对不同类别用户设定不同价格。

传统经济技术条件下，企业受信息收集和分析能力限制，难以精准捕捉消费者个体支付意愿，价格歧视实践多在二级、三级层面，一级因实施条件严苛只存在于理论中。而在数字经济背景下，企业能广泛获取并深度分析用户多维信息，完成精准画像，大幅提升识别个体支付意愿的能力。这让电商平台个性化定价突破传统群体划分局限，在实践中趋近一级价格歧视“千人千价”的理想目标[4]。

(2) 个性化定价是一种更加精细的价格歧视形式，它基于不同层次细分，最终目的是实现更加精准的价格歧视以获得企业的最大利润。一方面，平台利用差异化定价策略，将商品与服务精准匹配给不同支付意愿和需求特征的消费者，提升交易匹配效率，减少资源闲置与浪费。另一方面，平台靠算法捕捉消费者支付意愿，通过差异化定价挖掘不同消费群体潜力，实现收益最大化，是电商平台在数字市场提升盈利、巩固地位的核心策略。

2.2. 算法公平性的多维维度

结合电商定价场景，算法公平性包含三维核心内涵：程序公平，即定价算法设计、执行全流程透明、可解释、可追溯；分配公平，即定价结果符合市场公平，避免歧视性定价导致消费者剩余过度被剥夺；信息公平，即平台与消费者信息对称，杜绝信息操纵诱导非理性消费。三者互为支撑，程序公平为基础，信息公平为前提，分配公平为最终目标。

2.3. 核心概念辨析：个性化定价、价格歧视、“大数据杀熟”

在平台定价讨论中，“个性化定价”“价格歧视”与“大数据杀熟”常被混为一谈，实则三者概念层级不同需加以厘清。

价格歧视是三者中范畴最广的经济学概念，指当某种商品任一单位的需求价格独立于其他单位的销

售价格时，垄断厂商利用这种差异向不同接受者实行不同收费标准或销售价格的行为[5]。传统上可分为三级，其中一级价格歧视最接近“千人千价”的理想状态。

个性化定价是电子商务兴起后形成的概念，本质上属于价格歧视的技术化、精细化形态，词性偏向中性。合理运用可提升交易匹配效率，但过度使用可能导致经营者过度攫取消费者剩余、挤压竞争空间。

“大数据杀熟”则是个性化定价偏离公平原则的极端异化表现，特指平台利用信息不对称优势，对老用户、高粘性用户实施隐蔽性高价的失范行为，具有明显的剥削性与道德贬义。

三者关系为层层嵌套：价格歧视范畴最广，个性化定价是其在数字环境的具体应用，“大数据杀熟”是个性化定价中违反公平原则的极端子集。换言之，并非所有价格歧视都体现为个性化定价，也并非所有个性化定价都构成“杀熟”，只有系统性侵害消费者公平交易权的行为才属于“大数据杀熟”。

3. 电商个性化定价中的算法公平性风险分类与生成机理

个性化定价是基于数据与算法的动态定价机制，在提升资源配置效率、匹配用户需求、促进市场细分等方面具有显著的经济价值。然而，当这种机制在运行中缺乏透明度、可问责性以及基本的伦理约束，且被单一地导向利润最大化时，就可能突破合理边界，引发算法公平性风险。本章将系统梳理个性化定价在电商场景中可能引发的公平性风险，并深入分析这些风险是如何产生的。

3.1. 电商个性化定价的算法公平性核心风险分类

3.1.1. 分配不公风险：消费者剩余的非合理剥夺

在电商生态中，分配公平要求价格差异具有正当依据，而非单纯基于平台对用户支付意愿的精准预测。个性化定价算法通过挖掘海量数据，实现了从“三级价格歧视”向“一级价格歧视”的转变，由此引发了分配不公的风险。

首先，算法定价的本质是对消费者剩余的深度侵蚀。传统市场中，商家因信息搜集成本限制，难以精准掌握个体消费者的支付意愿。而电商平台借助用户画像技术，整合购买历史、搜索路径、支付频率、设备终端等多维度信息，可精准测算消费者的“保留价格”。算法依据这一最高支付意愿实时调价，将原本属于消费者的福利，即保留价格与市场价格的差额，最大限度转化为平台利润，打破了“同物同价”的市场公平基准，使价值分配向拥有算法优势的平台方极度倾斜[6]。

其次，这类分配不公在特定群体中表现尤为突出。算法往往会识别出价格敏感度低的用户，如急需特定商品的消费者、对平台形成路径依赖的高粘性老用户，并对其定向施加价格溢价。这不仅在经济层面造成了社会福利的净损失，更在伦理层面触及了社会公平的底线，导致数字时代“贫者越贫，富者越富”的效应在消费端变相体现。

3.1.2. 程序不公风险：算法黑箱对知情权与选择权的侵蚀

程序不公是个性化定价算法公平性风险的底层技术问题，核心表现为定价算法的设计、训练、执行与调整全流程具有高度隐蔽性与不可解释性，即算法“黑箱效应”，这直接剥夺了消费者的定价知情权与自主选择权[7]。

电商平台的定价算法依托大数据与机器学习技术构建，其模型训练的特征变量、价格调整的核心逻辑及算法迭代的决策依据均由平台单方面掌控，且未向消费者进行任何公开与解释。消费者既无法知晓平台为何对自己设定该价格，也无法追溯定价算法的决策过程，更无法参与定价规则的制定。在此情形下，消费者成为算法决策的被动接受者。面对差异化的定价结果，消费者缺乏有效的质疑与抗辩依据，其在交易中的主体地位被弱化，而平台则凭借算法技术形成了定价过程中的信息与权力垄断，违背了程序公平对透明性、可解释性与可追溯性的要求。

3.1.3. 信息不公风险：信息优势下的消费者非理性消费诱导

信息公平是实现市场有效竞争的重要前提。然而在个性化定价场景中，平台凭借其信息优势，通过算法对信息流进行选择呈现，从而实现深层的信息操纵。

一方面，平台借助算法开展精准的心理操纵。通过实时监测消费者行为数据，算法能够捕捉个体的心理防线与认知偏好。例如，当算法对用户实施个性化加价时，往往会同步推送“限时抢购”“仅剩几件”等制造紧迫感的诱导性信息。这种依托算法能力的“数字诱导”，利用消费者的损失厌恶与从众心理，迫使他们在信息不充分、情绪焦虑的状态下做出非理性购买决策。此时，价格不再是价值的信号，而沦为操纵消费行为的工具。

另一方面，算法通过构建“信息茧房”加剧了价格欺诈风险。在算法推荐的驱动下，消费者难以接触同一商品在不同时间、不同账户下的真实价格基准。这种高度的信息封闭性，使得平台可利用信息租金实施动态操纵。消费者在缺乏有效比价工具和全局市场信息的情况下，往往误以为眼前的“个性化价格”是基于某种优惠政策的最低价，实则可能支付了远高于市场平均水平的溢价。这种对信息流动权的掌控，本质上是对市场公平竞争环境的破坏。

3.2. 算法公平性风险的生成机理

电商个性化定价中的算法公平性风险并非单一因素所致，而是技术缺陷、组织目标与制度滞后共同作用的结果，三者相互关联、层层助推，形成风险生成的闭环。

3.2.1. 技术层面：算法技术的固有缺陷与应用偏差

定价算法的技术特性与应用偏差，是风险产生的底层客观诱因。首先，训练数据中蕴含的历史偏见经算法学习后，会内化为其决策逻辑，放大市场既有的歧视，最终体现为畸形的定价偏差。若历史交易数据本身反映了消费者支付意愿的差异，算法便可能将这种差异固化为定价规则，使弱势群体承受更高溢价。其次，特征选择环节可能引入歧视性代理变量。平台常将设备型号、消费粘性、地理位置等与商品实际价值无关的指标纳入定价模型，以此推断用户价格敏感度，进一步加剧对用户的分层[8]。此外，算法模型本身缺乏可解释性。深度学习等复杂模型常被视为“黑箱”，其决策过程难以被开发者清晰解读，更不用说接受外部审计与消费者质询了。这种技术层面的不透明性，为程序不公埋下了结构性隐患。

3.2.2. 组织层面：平台逐利性与内部治理缺失

技术缺陷能否转化为现实风险，关键在于平台组织的动机与约束。当前电商平台的经营目标偏差与内部治理缺位，构成了风险生成的核心内在动因。

在激烈的市场竞争与资本回报压力下，多数平台将利润最大化作为算法设计的首要乃至唯一目标，个性化定价由此沦为攫取消费者剩余的工具。在此逻辑下，算法公平性往往被视为次要约束，甚至遭到有意忽视。部分平台可能刻意利用算法的不透明性，对高粘性、低比价倾向的用户实施隐蔽的溢价策略。与此同时，绝大多数电商平台既未建立专门的算法伦理审查机制，也缺乏统一、可操作的公平性评估标准。定价算法通常由技术团队与运营部门联合开发，全程没有跨部门的独立审核与常态化的公平性监测，更未配套与绩效挂钩的问责体系。这种“重效率、轻伦理”的内部治理空白，使得技术层面的潜在偏见难以在组织内部被识别、纠正或阻断，最终将技术风险转化为实际的市场伤害。

3.2.3. 制度层面：监管体系滞后与救济渠道不畅

外部制度环境的不健全，为平台实施歧视性定价提供了可乘之机。一方面，现行法律缺乏对算法定价的专门规制。尽管《价格法》《消费者权益保护法》均明令禁止价格欺诈与价格歧视，但未明确划分

“合理个性化定价”与“违法大数据杀熟”的法律边界，也未规定平台在算法透明度、解释义务和公平性验证等方面的具体责任[9]。这导致执法机关面对新型算法侵权行为时，常陷入“无法可依”或“法律适用模糊”的困境。另一方面，传统监管手段与技术发展严重脱节。以人工抽查、事后举报为主的监管模式，难以应对算法实时动态调价、高度个性化且隐蔽性强的运行特征。目前算法备案、第三方审计、实时价格异常监测等智能化监管工具尚未完善，监管响应存在显著时滞。更为关键的是，消费者在遭遇算法不公时，既无法获取平台内部的定价逻辑，也难以证明自身遭受了不合理的差别待遇，维权面临重重困难。加之集体诉讼机制不完善、公益诉讼适用范围有限，市场难以对平台失范形成有效反制，制度约束力被进一步削弱。

综上，技术、组织与制度三层因素相互交织，共同构成算法公平性风险的生成闭环：技术漏洞提供操作空间，内部治理缺位将隐患转化为现实风险，外部制度滞后则使失范行为缺乏有效约束。唯有三者协同发力，构建覆盖算法全生命周期的治理体系，方能遏制风险，推动平台经济向更公平、透明的方向发展。

4. 算法公平性对电商生态系统的影响效应

4.1. 消费者信任崩塌与品牌忠诚流失

消费者是电商生态活力的核心，其对算法公平的感知直接影响平台发展。首先，算法定价的不透明性严重侵蚀了交易的信任基础。“大数据杀熟”等歧视性定价打破了消费者对“同质同价”的基本预期，一旦用户意识到平台利用其浏览轨迹、消费能力等私密数据攫取剩余价值，原本基于便利性建立的信任便迅速崩塌。

其次，不公平定价会引发强烈的心理抗拒效应。根据公平理论，消费者在交易中不仅关注绝对价格，更在意与他人相比的相对价格。当算法精准识别并对其推送更高价格时，这种“价格背叛”会激发负面情绪，不仅导致购买中断，还可能引发差评、投诉或社交媒体发酵，进而触发舆论危机。而信任是品牌忠诚的基石，公平性问题引发的信任裂痕难以通过简单整改修复，往往促使消费者转向更透明、合规的平台。加之负面口碑在社交网络中的快速扩散，用户流失范围进一步扩大。

4.2. 平台短期逐利与长期合规声誉风险的失衡

当算法公平性约束缺失时，平台虽可通过剥夺消费者剩余、诱导消费等方式快速获取短期超额利润，但长期将面临严峻风险。

在合规层面，数字经济监管体系持续完善，算法备案、事前审查等制度逐步落地，现有法律对价格歧视、不正当竞争的规制，使得算法不公平定价的违法成本不断攀升，平台易面临行政处罚、集体诉讼及巨额赔偿。在声誉层面，算法不公一旦曝光，极易引发信任危机，不仅导致用户流失、市场份额萎缩，还可能影响供应商与合作方的信心，压缩发展空间。尤其在 ESG 评价体系日益受重视的背景下，算法公平性已成为衡量平台社会责任的关键指标。负面事件造成的品牌损伤具有长期性和累积性，修复难度极大。因此，短期利润的获取，实质是透支品牌价值与用户基础，其长期损失远非短期收益所能弥补。

4.3. 市场机制扭曲与结构性失衡加剧

算法公平性风险不仅影响个体平台，还可能破坏市场机制的正常运行，引发多重市场失衡问题。

一是价格竞争机制失灵。在传统市场中，价格作为公开信号引导资源配置。而在“千人千面”的算法定价模式下，价格转变为私密且动态的信息，消费者之间、消费者与商家之间的信息不对称问题被进一步放大。头部平台凭借数据与算法优势，使价格脱离成本与供需关系，导致市场价格信号失真，市场

自我调节功能被削弱。

二是数字鸿沟加剧。算法往往通过设备类型、地理位置等信息识别数字素养较低的群体，如老年人、农村用户，并对其实施更高定价。这类用户因缺乏规避能力成为“技术性剥削”对象，而高素养用户却能有效规避，这种差异持续拉大消费福利差距，与数字经济的普惠初衷相背离。

三是形成市场准入壁垒。头部平台构建起“数据垄断-算法优势-利润积累-数据再获取”的闭环，新进入的中小平台因缺乏数据资源与算法研发资金，既无法实现优质的个性化定价，也无力参与价格竞争。这最终导致电商市场集中度提升、竞争程度降低，中小平台的创新空间被挤压，市场的资源配置效率与发展活力下降。

5. 算法公平性的多维协同治理机制构建

5.1. 协同治理的逻辑基础

当前，政府仍沿袭传统的“二元分散治理”模式对数字经济市场进行监管，即强调市场自由竞争、主体间相互制衡，政府则秉持包容审慎原则，主要采取事后监管方式^[10]。然而，这一模式在应对电商平台个性化定价问题时，正日益显现出“被动式、碎片化监管”的治理困境：缺乏整体性与协同性，难以识别和规制隐藏在技术黑箱中的公平性风险，不仅治理效能受到限制，还可能削弱公众对数字创新的信任。因此，政府需革新治理范式，坚持主导作用，吸纳平台、消费者及第三方社会力量，构建多元参与、协同联动的治理机制。

相较于传统单向规制模式，“多元共治”的协同治理模式优势显著。它通过搭建信息共享与利益协调机制，实现治理资源的优化配置，有效弥补单一主体的能力短板，破解主体间的信息壁垒。同时实现“事前预防、事中监测、事后追责”的全流程治理，既能守住算法公平底线，又能兼顾电商行业的技术创新需求。

该模式具备现实可行性。首先，各方在维护市场信任、保障交易公平上有共同利益基础。其次，政府有统筹协调的制度优势，能通过立法引导、标准制定与监管激励，搭建多方对话与协作平台。此外，随着算法透明度要求提升和消费者权利意识增强，多元主体参与治理的意愿与能力也在持续提升。

5.2. 政府主导的纵向约束机制

政府作为公共利益的守护者，在协同治理中应发挥“划线”与“撑腰”的双重作用。一方面，应在现有法律框架下健全算法定价配套制度体系。尽管《电子商务法》《个人信息保护法》等已确立公平交易等原则，但对个性化定价中“合理差异”与“不公平歧视”的边界仍缺乏可操作标准。为此，政府相关部门可依据法定权限，制定算法定价合规指引或发布负面清单，明确禁止基于敏感属性或利用信息不对称实施的剥削性定价，同时允许基于成本、促销等正当理由的价格差异化，划清合法商业实践与违规行为的界限。此外，可推行算法备案与公平性自评估制度，要求大型平台就定价算法的核心逻辑、数据来源及用户画像维度向监管部门报备，提升监管可追溯性。

另一方面，政府应创新监管方式，以“柔性监管”激发平台合规内生动力。例如，引入监管沙盒机制，在限定范围、设置价格波动上限并保障用户退出权的前提下，允许平台测试新型定价模型，监管部门同步评估其公平影响，实现创新与风险的动态平衡。同时，强化事后救济，降低消费者维权门槛。可通过完善集体诉讼适用条件、建立算法定价争议快速调解通道、推广小额纠纷在线裁决等方式，畅通权利救济渠道，为普通用户提供必要的举证支持或法律援助，切实形成“有规可依、有责可追、有救可及”的治理闭环。

5.3. 平台驱动的横向自律机制

平台作为算法的设计者与部署者，是治理链条中最关键的责任主体。面对日益严格的监管预期与消费者信任危机，平台需将公平性内化为组织基因，而非被动应对。

首先，应推行“嵌入式公平设计”，在算法开发初期即引入公平性约束指标，如价格差异率阈值、用户画像敏感变量屏蔽机制等，从源头减少偏见生成[11]。其次，建立独立的算法伦理委员会，由法律、技术、消费者代表等多方组成，对重大定价策略进行前置审查与定期审计。该委员会应拥有对重大决策的“一票否决权”，负责对算法上线前的潜在风险进行前置评估，并在运行过程中实施定期的合规审计。再次，实施算法透明度报告制度，定期向社会公开非涉密的算法逻辑说明、公平性评估结果信息。这种“有限透明”既能满足公众知情权，又可避免商业机密泄露，有助于重建平台公信力。

5.4. 社会参与的底层驱动机制

算法公平治理不能仅依赖政府规制与平台自律，还需激活广泛的社会力量。消费者、独立研究机构、行业组织等多元主体，共同构成协同治理的底层支撑。其中消费者不应仅是算法的被动接受者，而应成为治理网络中的积极行动者。为此，需通过制度完善提升其议价与监督能力。一方面，落实《个人信息保护法》赋予的算法解释权与选择权。平台应提供清晰易懂的个性化定价说明，并设置“一键关闭”选项，允许用户选择统一定价模式。这种“退出权”既是对个体自主的尊重，也是对平台滥用算法的有力制约。另一方面，消费者也应自动平衡“让渡隐私”与“获取便利”的关系，通过有条件开放权限、货比三家、理性消费提升反制能力[12]。发现违法行为时及时留存证据，通过12315平台或法律渠道维权，为协同治理贡献力量。

此外，应大力培育第三方评估生态。鼓励高校、研究机构、非营利组织开发算法公平性测评工具，开展“黑箱测试”与横向比较，发布平台算法公平指数。此类独立评估可为监管决策提供依据，也为消费者提供择优参考。同时，行业协会也应发挥自律功能，制定《电商算法伦理公约》，明确禁止“大数据杀熟”等行为，并建立声誉惩戒机制。对违反公约的成员，可采取公开谴责、限制行业评优、暂停会员资格等措施，形成“软法”约束。通过消费者赋权与社会监督的双重加持，可构建自下而上的反向制衡力量，弥补政府监管的盲区。

5.5. 多元主体的协同整合与运行路径

政府、平台与社会三方治理机制若孤立运行，难以形成合力。唯有通过制度化的协同接口，方能将纵向约束、横向自律与底层制衡有机融合，构建闭环、动态、可持续的算法公平治理体系。

5.5.1. 跨主体信息共享机制

协同治理的前提是信息互通。当前，政府监管部门难以实时获取平台定价数据，第三方机构缺乏系统性的评测样本，消费者更是处于信息孤岛的末端，使得算法公平性风险难以及时识别、精准定位。

为此，可构建“算法定价治理信息平台”，在保障商业秘密与个人隐私的前提下，推动三类数据的有限互通：一是平台备案数据，包括定价算法的核心逻辑、特征变量维度等；二是第三方评测数据，由高校、研究机构通过黑箱测试形成的算法公平性指数与合规评级；三是消费者感知数据，即投诉举报的热点分布、高频争议场景等。通过建立跨主体信息共享机制，既为政府监管提供决策依据，也为第三方研究提供样本支持，同时倒逼平台提升算法透明度，逐步打破信息垄断形成的算法黑箱。

5.5.2. 动态闭环反馈机制

当前，消费者个体投诉常难以及时回应，第三方评测报告也多被束之高阁，而政府监管往往依赖于

事后舆情的被动触发。这种分散式的响应模式，难以对平台行为形成持续有效的约束。破解这一困局的关键，在于构建一个“社会监督-政府介入-平台响应-第三方核验”的闭环反馈机制[13]。当消费者或第三方机构发现疑似不公平定价行为时，可通过标准化渠道提交证据；监管部门据此启动快速核查程序，必要时约谈相关平台；平台须在限期内完成整改并反馈；最终由第三方机构对整改成效进行验证，并向社会公开披露。由此，形成“社会监督触发、政府介入跟进、平台整改落实、公众监督闭环”的良性循环。

5.5.3. 激励相容机制

协同治理不仅需要约束与惩戒，更需正向激励引导平台主动合规。唯有将公平性内化为市场竞争的积极要素，才能激发平台持续改进的内生动力。

一方面，建立“合规红利”激励机制。对主动实施嵌入式公平设计、定期发布算法透明度报告、配合监管沙盒测试的平台，给予信用加分、监管频次减免等政策倾斜。另一方面，探索“主动披露从宽”制度。对主动发现算法漏洞、及时向监管部门报告并采取有效整改措施的平台，依法从轻、减轻或免除处罚，避免“隐瞒问题-集中曝光-声誉崩塌”的恶性循环。通过“奖优罚劣”的利益引导，将外部监督压力转化为平台主动合规的内生动力，使算法公平从“不得不为”走向“愿为之、乐为之”。

上述机制共同支撑起三层联动架构：政府设定底线规则并提供制度保障，平台履行主体责任并优化算法实践，社会力量提供外部校验与持续监督。三者协同运作，既能保障创新空间，又确保公平底线，使消费者诉求切实转化为治理实效。

6. 结论与展望

电商平台个性化定价作为数字经济时代的典型商业模式，在提升交易匹配效率、优化用户体验的同时，也因算法的“黑箱效应”与平台的逐利导向，衍生出分配不公、程序不公与信息不公三重公平性风险。本文研究表明，这些风险的生成并非单一因素所致，而是技术缺陷、组织目标偏差与制度供给滞后相互交织、层层递进的结果。算法公平性的缺失不仅会侵蚀消费者信任、损害平台长期声誉，更可能扭曲市场价格机制、加剧数字鸿沟，最终危及电商生态系统的健康可持续发展。

破解这一治理困局，需突破传统的“政府-市场”二元监管框架，转向政府、平台、消费者与第三方社会力量协同发力的多维治理模式。政府应发挥主导作用，完善算法定价的制度底线，创新监管工具；平台须将公平性内嵌于算法设计全流程，建立自律机制；消费者需提升数字素养与维权意识；第三方机构则应发挥独立评估与舆论监督功能。通过构建信息共享、闭环反馈与激励相容的协同机制，实现“事前预防-事中监测-事后追责”的全流程治理，方能在保障技术创新活力的同时，守住市场公平的底线。

展望未来，随着生成式人工智能的迭代升级，个性化定价算法将愈发复杂、动态与隐蔽，算法公平性治理将面临新的挑战。后续研究可在以下方向深化拓展：一是探索算法公平性的量化评估指标体系，为监管提供可操作的技术工具；二是追踪不同行业、不同规模平台在个性化定价实践中的差异化公平风险；三是关注消费者“算法回避”行为与“退出权”行使对平台定价策略的制衡效应。唯有理论与实践探索同频共振，方能推动数字市场在公平与效率的平衡中行稳致远。

参考文献

- [1] 周扬欣. 个性化定价对消费者购买意愿的影响: 价格公平感知的中介机制[D]: [硕士学位论文]. 上海: 华东师范大学, 2025.
- [2] OFT (2013) The Economics of Online Personalised Pricing. <https://webarchive.nationalarchives.gov.uk/ukgwa/20130704053952/>

-
- [3] OECD (2018) Personalised Pricing in the Digital Era (OECD Roundtables on Competition Policy Papers No. 222). OECD Publishing.
 - [4] 邹开亮, 徐培杰, 曾文琴. 算法个性化定价的规制困境与纾解进路[J]. 北京科技大学学报(社会科学版), 2026, 42(2): 72-79.
 - [5] 阿瑟·塞西尔·庇古. 福利经济学: 上[M]. 朱泱, 张胜纪, 吴良健, 译. 北京: 商务印书馆, 2006: 290.
 - [6] 孙鑫. 消费者权益保护视域下算法个性化定价行为规制研究[D]: [硕士学位论文]. 重庆: 西南政法大学, 2023.
 - [7] 王锋. 数字经济中个性化定价的法律规制[J]. 北京航空航天大学学报(社会科学版), 2023, 36(5): 199-207.
 - [8] 刘仕贤, 李佳薇. 算法价格歧视下消费者权益保护的技术逻辑、规制反思与路径重构[J]. 价格理论与实践, 2025(3): 63-67.
 - [9] 陈增宝, 张凌寒. 算法技术的法律规制: 治理困境、发展逻辑与优化路径[J]. 中国运用法学, 2024(4): 155-166.
 - [10] 李丹. 算法歧视消费者: 行为机制、损益界定与协同规制[J]. 上海财经大学学报, 2021, 23(2): 17-33.
 - [11] 梁正, 曾雄. “大数据杀熟”的政策应对: 行为定性、监管困境与治理出路[J]. 科技与法律, 2021(2): 8-14.
 - [12] 陶苞朵, 张等文. 遏制大数据“杀熟”: 政府主导的协同治理模式[J]. 常州大学学报(社会科学版), 2022, 23(5): 13-21.
 - [13] 梁宇, 郑易平. 我国政府数据协同治理的困境及应对研究[J]. 情报杂志, 2021, 40(9): 108-114.