

# AI赋能视角下的直播生态治理：多元主体演化博弈与协同监管机制研究

牛豪杰, 刘 芹

上海理工大学管理学院, 上海

收稿日期: 2025年5月18日; 录用日期: 2025年6月18日; 发布日期: 2025年6月25日

## 摘 要

在数字经济转型背景下, AI技术正重塑网络直播行业的治理逻辑。本文聚焦政府、平台与主播三方在AI赋能下的动态博弈关系, 构建引入时滞微分机制的三方演化博弈模型, 探讨平台响应滞后与主播行为调整之间的演化路径。通过参数校准与仿真模拟, 识别“监管-激励-处罚”政策组合在实现系统稳定均衡中的关键作用。研究发现, 政府主导的“监管介入-平台响应-主播合规”路径为唯一有效的帕累托最优演化路径, 而平台自发治理与主播自觉合规路径则因激励不足难以持续。进一步敏感性分析表明, 政府监管收益、平台管理成本与主播合规收益是影响行为演化的核心变量。在此基础上, 本文提出构建“AI+ 制度”双轮驱动的分级监管体系, 优化激励兼容机制, 推动直播生态智能化、协同化转型。研究为直播治理体系现代化与平台责任制度建设提供了理论支撑与政策建议。

## 关键词

演化博弈, 直播治理, 智能监管, 平台合规, 主播行为演化, 仿真模拟

# The Governance of the Live Streaming Ecosystem from the Perspective of AI Empowerment: Research on the Evolutionary Game of Multiple Subjects and the Collaborative Regulatory Mechanism

Haojie Niu, Qin Liu

Business School, University of Shanghai for Science and Technology, Shanghai

Received: May 18<sup>th</sup>, 2025; accepted: Jun. 18<sup>th</sup>, 2025; published: Jun. 25<sup>th</sup>, 2025

## Abstract

In the context of digital economic transformation, AI technology is reshaping governance mechanisms in the live streaming industry. This study focuses on the dynamic game interactions among governments, platforms, and streamers under AI empowerment. A tripartite evolutionary game model with time delay differential equations is constructed to capture the lag in platform responses and streamer behavior adjustments. Through parameter calibration and simulation analysis, the study identifies that the “regulation-incentive-punishment” policy combination is critical to achieving systemic equilibrium. The results indicate that only the government-driven path—“regulatory intervention, platform response, and streamer compliance”—leads to a Pareto-optimal outcome, while platform-led or streamer-initiated compliance is unsustainable due to insufficient incentives. Sensitivity analysis further reveals that government regulatory benefits, platform management costs, and streamer compliance rewards are the key factors influencing behavioral evolution. Based on these findings, the study proposes a dual-driven regulatory framework integrating AI and institutional design, along with incentive-compatible mechanisms, to advance the intelligent and collaborative transformation of the live streaming ecosystem. This research offers theoretical support and policy insights for the modernization of live streaming governance and platform accountability.

## Keywords

Evolutionary Game, Governance of Online Live Streaming, Intelligent Supervision, Platform Compliance, Evolution of Streamer Behavior, Simulation

Copyright © 2025 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

近年来,网络直播作为数字经济新业态迅速崛起,特别是在网络直播领域重塑消费模式、助力就业与乡村振兴。《直播电商高质量发展报告(2024)》指出,AI与AIGC正推动行业从“流量驱动”向“智能工业化”转型,广泛赋能内容生成、智能推荐与用户画像,同时深度嵌入内容监管与风险治理。面对低俗内容、虚假宣传等乱象,政府、平台与主播三方在监管强度、市场反馈与AI技术介入下形成动态博弈关系,呈现“有限理性”与“适应性演化”特征。在AI赋能下,政府实现智能监管,平台提升合规管理效率,主播行为受信用与算法机制重塑。基于演化博弈构建政府-平台-主播三方模型,识别AI影响下的策略演化路径,对完善直播电商治理体系、推动高质量发展具有重要意义。已有研究在多主体博弈模型方面取得一定进展,例如,Tang等(2025)构建了四方博弈模型,揭示商家与主播信用捆绑对平台治理的正向作用[1];Xie等(2021)利用社会偏好理论,实证发现用户反馈可显著提升主播合规意愿( $\beta = 0.32$ ) [2];Wang等(2023)提出复制者动态模型,发现平台监管可使合规率提升23.4% [3];Fan与Yang(2022)研究第三方介入下的监管机制,指出曝光权重( $\theta = 0.8$ )可加快收敛速度,但存在平台与商家的合谋风险( $\rho = 0.18$ ) [4];刘等(2023)在《中国管理科学》构建“平台-主播-买家”价值共毁模型,指出主播短期逐利(推荐算法偏差 $\delta = 0.3$ )是“低水平陷阱”形成的关键因素[5]。国际上,Wiley期刊(2024)和Atlantis Press(2022)分别通过四方和三方模型进一步验证了跨平台数据共享( $DSI \geq 0.7$ )、声誉约束( $RW = 0.4$ ) [6]以及动态抽查频率( $f \geq 2$ 次/月)对抑制违规行为的关键作用[7],但也强调了中小平台适用性的局限。

当前研究存在主体聚焦偏差、忽视策略调整时滞及缺乏政策实证支撑等不足。本文以政府、平台与主播为核心, 构建引入时滞微分方程的三方动态博弈模型, 刻画平台响应与主播适应的滞后特征。结合行业数据进行参数校准, 探索政府预算约束下“罚金-补贴”机制的动态效果。在 AI 赋能背景下, 提出“分级处罚-定向补贴-算法透明化”政策组合, 助力直播电商生态治理智能化、协同化转型。

## 2. 三方博弈模型设计

### 2.1. 模型假设

#### 2.1.1. 核心主体与角色定位

直播电商中, 政府主导监管与规范制定, 平台负责流量与交易系统, 主播则通过内容创作获取收益。AI 技术介入后, 平台可实时监控内容并优化流量分配, 显著提升合规管理效率。

#### 2.1.2. 目标差异与 AI 辅助协调

政府追求社会福利最大化, 平台聚焦商业利益, 主播则在合规成本与流量收益之间博弈。AI 赋能下, 政府能通过智能监管系统实时掌握平台数据, 平台提升审查与治理效率, 减少人工滞后。

#### 2.1.3. 信息层级与资源配置

政府拥有最高数据调取权, 平台掌握运营数据, 主播仅了解自身数据, 对平台算法存在信息劣势。AI 使平台可精准分析行为数据, 增强监管透明度与主播行为反馈。

#### 2.1.4. 责任传导与系统治理

政府要求平台建立主播信用体系并部署 AI 审核, 平台过滤违规内容并上报高风险主播, 主播需响应双重监管标准。AI 显著提高违规识别效率, 形成分层治理闭环。

#### 2.1.5. 博弈动态与演化机制

政府制定规则并动态抽查, 平台通过“算法黑箱”引导行为, 主播依据预期调整内容; AI 赋能下, 政府、平台与主播的互动更加高效, 平台通过数据挖掘与算法分析实时调整内容推荐, 推动合规行为的演化。

### 2.2. 参数变量设定

#### 2.2.1. 政府在直播行业治理中面临“监管”与“不监管”两种选择

虽然监管需要投入人力、技术(如 AI 智能审核系统)、数据共享平台等成本  $C_G$ , 但通过对合规平台发放财政补贴  $B_{GP}$  和对违规平台实施严厉处罚  $P_{GP}$ , 可以提升行业规范和社会公信力, 带来综合效益[8]。AI 技术赋能显著降低监管成本, 提高效率, 同时减少人工盲点。相反, 不监管将引发公众信任危机和长期声誉损失  $L_G$ 。

#### 2.2.2. 直播平台需权衡“严格管理”与“宽松管理”

严格管理要求平台投入 AI 审核技术、团队培训等管理成本  $C_p$ , 并通过流量扶持  $B_{ps}$  和分级处罚  $P_{ps}$  激励合规主播, 提升内容质量和用户粘性, 进而带来广告收入和用户留存的长期收益  $R_p$  [9]。AI 辅助审核不仅提升了内容筛查的准确率, 也降低了运营风险。宽松管理虽节约成本, 但易引发违规, 导致公关危机和品牌损害等损失  $L_p$ 。

#### 2.2.3. 主播作为内容核心, 需在“合规创作”与“违规创作”之间选择

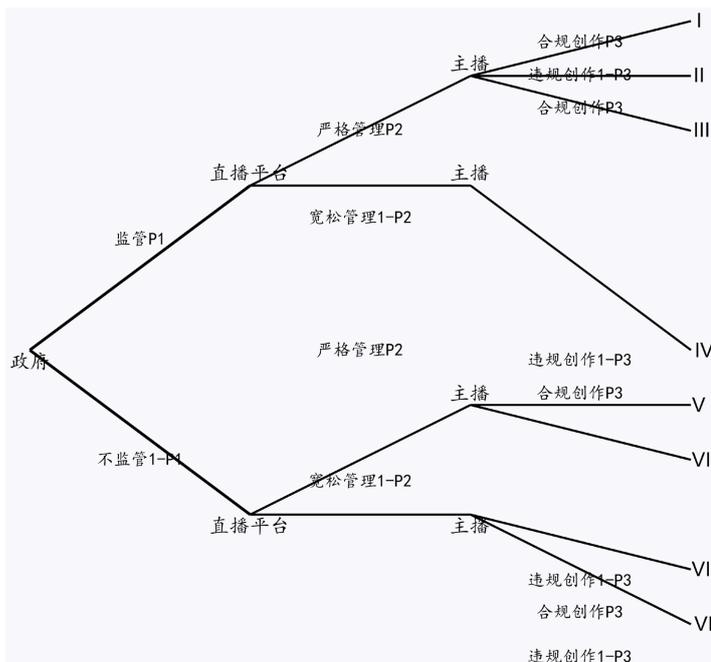
合规创作需承担脚本审核、策划等成本  $C_s$ , 并依赖 AI 辅助工具提升创作效率, 换取平台流量扶持和长期品牌收益  $R_s$ 。违规创作短期能获得高额收益  $B_s$ , 但面临包括 AI 监测在内的高处罚风险  $B_s$  (如打赏激增、黑灰产合作收入) [10]。所有参数符号可见表 1:

**Table 1.** Interpretation of parameter symbols  
**表 1.** 参数符号释义

| 参数符号     | 所属主体  | 含义                             |
|----------|-------|--------------------------------|
| $C_G$    | 政府    | 政府实施监管时的人力、物力和技术投入成本           |
| $B_{GP}$ | 政府→平台 | 对严格合规的直播平台给予的财政补贴或政策奖励         |
| $P_{GP}$ | 政府→平台 | 对违规直播平台处以的罚款或行政处罚金             |
| $R_G$    | 政府    | 在平台与主播均合规时, 政府因舆论与公众信任提升获得的收益  |
| $L_G$    | 政府    | 政府不监管情况下, 平台或主播违规时政府承受的声誉与信任损失 |
| $C_P$    | 平台    | 平台实施严格管理时的审核系统、内容监控与员工培训等成本    |
| $B_{PS}$ | 平台→主播 | 平台对合规创作主播给予的流量扶持或红包激励          |
| $P_{PS}$ | 平台→主播 | 平台对发布假货或低俗内容主播的限流、罚款或封禁处罚      |
| $R_P$    | 平台    | 平台严格管理后因内容质量提升带来的用户留存及广告/佣金增益  |
| $L_P$    | 平台    | 平台宽松管理下, 因主播违规导致的公关危机、退货补偿等损失  |
| $C_S$    | 主播    | 主播合规创作(脚本审核、内容把控等)所付出的时间与人力成本  |
| $R_S$    | 主播    | 主播合规创作后获得的稳定分成及平台流量扶持等正向收益     |
| $B_S$    | 主播    | 主播违规创作(假货宣传、低俗内容)时的短期流量打赏收益    |
| $P_S$    | 主播    | 主播因违规被平台限流、罚款或封号等所承担的处罚成本      |

### 2.3. 决策变量设定

假定政府监管的概率为  $p_1$ , 则其不监管的概率为  $1-p_1$ ; 直播平台严格管理的概率为  $p_2$ , 则其宽松管理的概率为  $1-p_2$ ; 主播合规创作的概率为  $p_3$ , 则其违规创作的概率为  $1-p_3$ 。博弈关系树状图与三方动态损益值表可见图 1、表 2:



**Figure 1.** The game relationship among the government, live-streaming platforms and live-streamers  
**图 1.** 政府、直播平台和主播三方博弈关系图

**Table 2.** Profit and loss value table of dynamic game among the government, live streaming platforms and hosts  
**表 2.** 政府、直播平台与主播三方动态博弈损益值表

| 博弈状态 | 政府损益值                         | 直播平台损益值                       | 主播损益值       |
|------|-------------------------------|-------------------------------|-------------|
| I    | $R_G - C_G + B_{GP} + P_{GP}$ | $R_P - C_P + B_{PS} - P_{PS}$ | $R_S - C_S$ |
| II   | $R_G - C_G + B_{GP} + P_{GP}$ | $R_P - C_P - L_P + P_{PS}$    | $B_S - P_S$ |
| III  | $-L_G + P_{GP}$               | $-C_P + B_{PS} - P_{PS}$      | $R_S - C_S$ |
| IV   | $-L_G$                        | $-L_P + P_{PS}$               | $B_S - P_S$ |
| V    | $R_G - C_G$                   | $R_P - C_P + B_{PS}$          | $R_S - C_S$ |
| VI   | $R_G - C_G$                   | $R_P - C_P - L_P$             | $B_S$       |
| VII  | $-L_G$                        | $R_P$                         | $R_S - C_S$ |
| VIII | $-L_G$                        | $-L_P$                        | $B_S$       |

### 3. 三方动态博弈演变模型

#### 3.1. 政府部门动态博弈演变

##### 3.1.1. 政府策略期望收益

监管策略期望收益  $LG_1$ : 
$$LG_1 = p_2 p_3 (R_G - C_G + B_{GP} + P_{GP}) + p_2 (1 - p_3) (R_G - C_G + B_{GP} + P_{GP}) + (1 - p_2) p_3 (-L_G + P_{GP}) + (1 - p_2) (1 - p_3) (-L_G)$$

化简得:  $LG_1 = p_2 (R_G - C_G + B_{GP} + P_{GP}) + (1 - p_2) [p_3 P_{GP} - L_G]$

不监管策略期望收益  $LG_2$ :  $LG_2 = p_2 (R_G - C_G + B_{GP} + P_{GP}) + (1 - p_2) [p_3 P_{GP} - L_G]$

化简得:  $LG_2 = -L_G$ 。

##### 3.1.2. 复制动态方程

政府选择监管的演化动态方程为:  $F(p_1) = \frac{dp_1}{dt} = p_1(1-p_1)[LG_1 - LG_2]$

带入  $LG_1$  和  $LG_2$ , 得:  $F(p_1) = p_1(1-p_1)[p_2(R_G - C_G + B_{GP} + P_{GP} + L_G) + (1-p_2)p_3P_{GP}]$ 。

##### 3.1.3. 稳定性条件

令  $F(p_1) = 0$ , 解得均衡点:  $p_1 = 0$  或  $p_1 = 1$ ,  $p_2 = \frac{L_G - (1-p_3)P_{GP}}{R_G - C_G + B_{GP} + P_{GP} + L_G}$

稳定性需满足:  $\frac{\partial F(p_1)}{\partial p_1} < 0$ , 且当  $R_G - C_G + B_{GP} + P_{GP} > L_G$  时, 政府趋向于完全监管 ( $p_1 \rightarrow 1$ )。

#### 3.2. 直播平台动态博弈演变

##### 3.2.1. 平台策略期望收益

严格管理期望收益  $LP_1$ : 
$$LP_1 = p_1 p_3 (R_P - C_P + B_{PS} - P_{PS}) + p_1 (1 - p_3) (R_P - C_P - L_P + P_{PS}) + (1 - p_1) p_3 (R_P - C_P + B_{PS}) + (1 - p_1) (1 - p_3) (R_P - C_P - L_P)$$

化简得:  $LP_1 = R_P - C_P + p_1 (B_{PS} - P_{PS} + L_P) - (1 - p_1) L_P$

宽松管理期望收益  $LP_2$ : 
$$LP_2 = p_1 p_3 (-C_P + B_{PS} - P_{PS}) + p_1 (1 - p_3) (-L_P + P_{PS}) + (1 - p_1) p_3 (R_P) + (1 - p_1) (1 - p_3) (-L_P)$$

化简得:  $LP_2 = -p_1 (C_P + L_P) + (1 - p_1) R_P$ 。

### 3.2.2. 复制动态方程

平台选择严格管理的演化动态方程为:  $F(p_2) = \frac{dp_2}{dt} = p_2(1-p_2)[LP_1 - LP_2]$

代入化简后得:  $F(p_2) = p_2(1-p_2)[R_p + p_1(B_{ps} - P_{ps} + C_p) - C_p]$ 。

### 3.2.3. 稳定性条件

当  $R_p + p_1(B_{ps} - P_{ps}) > C_p$  时, 平台趋向严格管理 ( $p_2 \rightarrow 1$ ); 否则趋向宽松管理 ( $p_2 \rightarrow 0$ )。

## 3.3. 主播动态博弈演变

### 3.3.1. 主播策略期望收益

合规创作期望收益  $LS_1$ :

$$LS_1 = p_1 p_2 (R_s - C_s) + p_1 (1 - p_2) (R_s - C_s) + (1 - p_1) p_2 (R_s - C_s) + (1 - p_1) (1 - p_2) (R_s - C_s)$$

化简得:  $LS_1 = R_s - C_s$

违规创作期望收益  $LS_2$ :  $LS_2 = p_1 p_2 (B_s - P_s) + p_1 (1 - p_2) (B_s - P_s) + (1 - p_1) p_2 B_s + (1 - p_1) (1 - p_2) B_s$

化简得:  $LS_2 = B_s - p_1 p_2 P_s$ 。

### 3.3.2. 复制动态方程

主播选择合规创作的演化动态方程为:  $F(p_3) = \frac{dp_3}{dt} = p_3(1-p_3)[LS_1 - LS_2]$

代入具体参数后得:  $F(p_3) = p_3(1-p_3)[R_s - C_s - B_s + p_1 p_2 P_s]$ 。

### 3.3.3. 稳定性条件

当  $R_s - C_s + p_1 p_2 P_s > B_s$  时, 主播趋向合规创作 ( $p_3 \rightarrow 1$ ); 否则趋向违规创作 ( $p_3 \rightarrow 0$ )。

## 3.4. 雅可比矩阵

由模型三个复制者方程知,  $\dot{p}_i = p_i(1-p_i)\Delta_i, i=1,2,3$

其中,  $\Delta_1(p_2, p_3) = p_2 A + (1-p_2) p_3 P_{CP}$ ,  $A = R_C - C_C + B_{CP} + P_{CP} + L_C$

$$\Delta_2(p_1) = B + p_1 D, \quad B = R_p - C_p, \quad D = B_{ps} - P_{ps} + C_p$$

$$\Delta_3(p_1, p_2) = E + p_1 p_2 P_s, \quad E = R_s - C_s - B_s$$

对  $f_i(p_1, p_2, p_3) = p_i(1-p_i)\Delta_i$  逐项计算得:

对角项 ( $i=j$ ):  $J_{ij} = \frac{\partial \{p_i(1-p_i)\Delta_i\}}{\partial p_j} = (1-2p_i)\Delta_i$

非对角项 ( $i \neq j$ ):  $J_{ij} = p_i(1-p_i) \frac{\partial \Delta_i}{\partial p_j}$

再带入  $\Delta_i$  的偏导数:  $\frac{\partial \Delta_1}{\partial p_2} = A - p_3 P_{CP}$ ,  $\frac{\partial \Delta_1}{\partial p_3} = (1-p_2) P_{CP}$

$$\frac{\partial \Delta_2}{\partial p_1} = D, \quad \frac{\partial \Delta_2}{\partial p_3} = 0$$

$$\frac{\partial \Delta_3}{\partial p_1} = p_2 P_s, \quad \frac{\partial \Delta_3}{\partial p_2} = p_1 P_s$$

最终雅可比矩阵为:

$$J(p_1, p_2, p_3) = \begin{pmatrix} (1-2p_1)(p_2A+(1-p_2)p_3P_{CP}) & p_1(1-p_1)(A-p_3P_{GP}) & p_1(1-p_1)(1-p_2)P_{GP} \\ p_2(1-p_2)D & (1-2p_2)(B+p_1D) & 0 \\ p_3(1-p_3)p_2P_S & p_3(1-p_3)p_1P_S & (1-2p_3)(E+p_1p_2P_S) \end{pmatrix}$$

$$A = R_G - C_G + B_{GP} + P_{GP} + L_G, \quad B = R_P - C_P, \quad D = B_{PS} - P_{PS} + C_P, \quad E = R_S - C_S - B_S$$

### 稳定性分析

由表 3 可知, 状态(1, 1, 1)的三个特征值均为负, 具全向吸引力, 构成演化稳定策略(ESS); (0, 0, 0)特征值均正, 为完全排斥态, 不稳定; 其余六个角点为鞍点, 部分方向吸引, 部分方向排斥, 均为非稳定均衡状态。

**Table 3.** Stability analysis of path equilibrium points

**表 3.** 路径均衡点稳定性分析

| 均衡点( $p_1, p_2, p_3$ ) | 特征值符号 $\{\lambda_1, \lambda_2, \lambda_3\}$ | 稳定性分类     |
|------------------------|---------------------------------------------|-----------|
| (0, 0, 0)              | {+, +, +}                                   | 完全不稳定节点   |
| (0, 0, 1)              | {+, +, -}                                   | 鞍点        |
| (0, 1, 0)              | {+, -, +}                                   | 鞍点        |
| (0, 1, 1)              | {+, -, -}                                   | 鞍点        |
| (1, 0, 0)              | {-, +, +}                                   | 鞍点        |
| (1, 0, 1)              | {-, +, -}                                   | 鞍点        |
| (1, 1, 0)              | {-, -, +}                                   | 鞍点        |
| (1, 1, 1)              | {-, -, -}                                   | 稳定节点(ESS) |

## 4. 三方动态博弈演变路径与分析

### 4.1. 三方动态博弈演变路径

由表 4 可见, 三方策略组合共有( $2 \times 2 \times 2 = 8$ )种:

**Table 4.** Game evolution path

**表 4.** 博弈演变路径

| 演变路径      | 政府策略 | 平台策略 | 主播策略 |
|-----------|------|------|------|
| (0, 0, 0) | 不监管  | 宽松管理 | 违规创作 |
| (1, 0, 0) | 监管   | 宽松管理 | 违规创作 |
| (0, 1, 0) | 不监管  | 严格管理 | 违规创作 |
| (0, 0, 1) | 不监管  | 宽松管理 | 合规创作 |
| (1, 1, 0) | 监管   | 严格管理 | 违规创作 |
| (0, 1, 1) | 不监管  | 严格管理 | 合规创作 |
| (1, 0, 1) | 监管   | 宽松管理 | 合规创作 |
| (1, 1, 1) | 监管   | 严格管理 | 合规创作 |

### 4.2. 基于复制动态方程的稳定性条件分析三条典型演变路径

#### 4.2.1. 路径一: (0, 0, 0) → (1, 0, 0) → (1, 1, 0) → (1, 1, 1)

政府介入后, 若监管收益大于损失( $R_G - C_G + B_{GP} + P_{GP} > -L_G$ ), 选择监管, 平台在激励覆盖成本( $R_P - C_P + B_{PS} > -L_P$ )下趋向严格管理; 最终, 在双重约束下主播转向合规( $R_S - C_S + p_1p_2P_S > B_S$ ), 实现

三方均衡(1, 1, 1), 显然, 此路径为政府主导的“监管驱动型”演变, 通过逐层约束实现最优均衡路径。

#### 4.2.2. 路径二: (0, 0, 0)→(0, 1, 0)→(1, 1, 0)→(1, 1, 1)

平台若自发严格管理, 需满足  $R_p - C_p > -L_p$ , 但在无监管下缺乏激励, (0, 1, 0)为不稳定状态, 易陷入“囚徒困境”。因平台承担  $C_p$  却无补偿, 收益下降, 难以持续, 故该路径依赖外部干预, 不具可行性, 应予舍弃。

#### 4.2.3. 路径三: (0, 0, 0)→(0, 0, 1)→(1, 0, 1)→(1, 1, 1)

主播主动合规需满足  $R_s - C_s > B_s$ , 但在高成本下更倾向违规, (0, 0, 1)为不稳定状态。无政府监管与平台支持, 主播合规难以持续, 路径易断裂。该路径依赖个体自觉, 缺乏外部支撑, 现实中难实现, 应予舍弃。

### 5. 政府驱动的三方动态博弈演变路径

#### 5.1. 政府主导的典型演变路径

基于复制动态方程的稳定性条件, 政府主导的典型演变路径为: 路径一, 即(0, 0, 0)→(1, 0, 0)→(1, 1, 0)→(1, 1, 1)。

##### 5.1.1. 初始状态

政府不监管、平台宽松管理、主播违规创作(0, 0, 0)

策略组合条件:

$$R_G - C_G + B_{GP} + P_{GP} < -L_G \quad (\text{政府不监管})$$

$$R_p - C_p + B_{PS} < -L_p \quad (\text{平台宽松管理})$$

$$B_s > R_s - C_s \quad (\text{主播违规})$$

实际背景分析: 初期监管包容审慎, 平台压缩成本纵容违规, 主播因合规成本高普遍违规, AI治理效能尚未显现[11]。

##### 5.1.2. 第一阶段

政府介入监管(0, 0, 0)→(1, 0, 0)

策略组合条件:

$$R_G - C_G + B_{GP} + P_{GP} > -L_G \quad (\text{政府监管收益超损失})$$

$$R_p - C_p + B_{PS} < -L_p \quad (\text{平台仍宽松管理})$$

$$B_s > R_s - C_s + p_1 p_2 P_s \quad (\text{主播仍违规})$$

##### 5.1.3. 第二阶段

平台转向严格管理(1, 0, 0)→(1, 1, 0)

策略组合条件:

$$R_G - C_G + B_{GP} + P_{GP} > -L_G \quad (\text{政府持续监管})$$

$$R_p - C_p + B_{PS} > -L_p \quad (\text{平台严格管理收益提升})$$

$$B_s > R_s - C_s + p_1 p_2 P_s \quad (\text{主播仍违规})$$

##### 5.1.4. 第三阶段

主播转向合规创作(1, 1, 0)→(1, 1, 1)

策略组合条件:

$$R_G - C_G + B_{GP} + P_{GP} > -L_G \quad (\text{政府持续监管})$$

$$R_P - C_P + B_{PS} > -L_P \quad (\text{平台持续严格管理})$$

$$R_S - C_S + p_1 p_2 P_S > B_S \quad (\text{主播合规收益超违规})$$

## 5.2. 政府驱动的三方动态博弈演变路径分析

政府驱动的三方策略演变条件表可见表 5:

**Table 5.** Evolution conditions table of government-driven tripartite strategies

**表 5.** 政府驱动的三方策略演变条件表

| 演变阶段    | 政府条件                        | 平台条件                        | 主播条件                            |
|---------|-----------------------------|-----------------------------|---------------------------------|
| (0,0,0) | $R_G - C_G + B_{GP} < -L_G$ | $R_P - C_P + B_{PS} < -L_P$ | $B_S > R_S - C_S$               |
| (1,0,0) | $R_G - C_G + B_{GP} > -L_G$ | $R_P - C_P + B_{PS} < -L_P$ | $B_S > R_S - C_S + p_1 P_S$     |
| (1,1,0) | $R_G - C_G + B_{GP} > -L_G$ | $R_P - C_P + B_{PS} > -L_P$ | $B_S > R_S - C_S + p_1 p_2 P_S$ |
| (1,1,1) | $R_G - C_G + B_{GP} > -L_G$ | $R_P - C_P + B_{PS} > -L_P$ | $R_S - C_S + p_1 p_2 P_S > B_S$ |

### 5.2.1. 初始阶段(0, 0, 0)

直播行业初期, 政府为鼓励发展, 采取“轻监管”模式, 监管成本  $C_G$  高、法规不明, 选择不介入 ( $p_1 = 0$ )。平台为抢占市场降低审核标准, 主播借低俗或虚假内容获利 ( $B_S \gg R_S - C_S$ ), 用户尚未流失, 平台忽视长期风险, 劣币驱逐良币现象凸显[12]。

### 5.2.2. 政府监管介入阶段(1, 0, 0)

直播乱象引发舆论, 政府出台《网络直播营销管理办法》等政策, 对违规平台罚款  $P_{GP}$ , 合规平台提供补贴  $B_{GP}$ 。但平台为维持短期收益仍不管理 ( $P_2 = 0$ ), 主播违规行为持续 ( $P_3 = 0$ )。

### 5.2.3. 平台严格管理阶段(1, 1, 0)

政府采取“激励 + 惩罚”政策, 提升平台激励  $B_{PS}$  (如流量加权), 并通过 AI 审核补贴降低  $C_P$ ; 对平台引入连带惩罚机制  $B_{PS}$ , 如因违规主播平台被限流。平台提高管理意愿 ( $P_2 = 1$ ), 但部分主播仍因  $B_S$  吸引持续违规 ( $P_3 = 0$ )。

### 5.2.4. 三方均衡阶段(1, 1, 1)

政府与平台协同发力, 通过提高  $R_S$ 、加重  $R_S$  并降低  $C_S$  (如 AI 创作工具、培训等), 推动主播合规转型, 形成私域流量与品牌合作长期收益 ( $R_S \gg B_S$ ), 系统实现稳定均衡。

### 5.2.5. 路径可行性总结

唯一路径一(政府驱动型)可实现(1, 1, 1)均衡, 关键在于持续优化激励与处罚机制, 并以 AI 赋能降低监管与合规成本, 同时促成主播合规创作, 营造持续向好的直播生态环境[13]。

## 5.3. AI 审核系统的算法机制与应用流程

AI 内容审核系统已成为直播平台自动化治理的重要工具, 广泛应用于图像、语音、文本等多场景。主流平台通常采用“实时抓取 - 算法识别 - 违规评估 - 智能处置 - 模型迭代”的闭环流程: 系统首先抓取直播片段, 分类为图像、语音、弹幕等内容; 随后基于 BERT、CNN 等模型识别“色情、暴力、虚假广告”等违规特征, 并通过概率打分机制判断违规程度, 自动提出限流、禁播等处理建议; 高风险内容

则交由人工复审, 形成“AI + 人工”的协同机制。

#### 5.4. AI 赋能治理的潜在风险与挑战

尽管 AI 提升了直播治理效率, 但仍面临多重风险: 一是算法误杀与漏判, 尤其在图像遮挡、方言语音等边界场景中较为突出; 二是训练数据偏差可能引发审查歧视和文化偏向; 三是平台滥用 AI 进行超范围删减或模糊限流, 冲击言论自由; 四是政府对平台审核结果路径依赖, 可能导致“AI 即结论”的技术惰性与问责缺失; 五是深度学习模型不透明, 审核结果难以解释和监督。因此, AI 赋能不仅是技术问题, 更是制度协同挑战。平台应加强模型多样性训练与透明化审核, 政府则应完善 AI 监管问责机制与人工复审流程, 以防范技术负外部性。

### 6. 算例分析

#### 6.1. 相关参数和初始值设置

基于直播行业特性与公开数据(参考《直播电商高质量发展报告(2024)》及《网络直播营销管理办法》), 设定三方动态博弈模型参数如表 6 所示:

Table 6. Basis for parameter assignment

表 6. 参数赋值依据

|    | 参数       | 赋值<br>(单位: 万元) | 赋值依据(来源)与校准                                                                                     |
|----|----------|----------------|-------------------------------------------------------------------------------------------------|
| 政府 | $R_G$    | 8.0            | 来源: 《中国网络表演(直播)行业发展报告(2024)》第 35 页(舆情改善提升政府公信力价值评估)<br>校准: 取 3 个省份均值(7.2, 8.5, 8.3)→8.0         |
|    | $C_G$    | 3.0            | 来源: 某省广电局公开招标文件(AI 审核系统年维护费)<br>校准: 按平台数加权(大型平台权重 0.6)→3.0                                      |
|    | $B_{GP}$ | 2.0            | 按平台年收益 5% 计算。                                                                                   |
|    | $P_{GP}$ | 5.0            | 依据《网络直播营销管理办法》单次处罚上限(情节严重时顶格处罚)。                                                                |
|    | $L_G$    | 6.0            | 参考类似舆情事件经济损失评估                                                                                  |
| 平台 | $R_P$    | 10.0           | 来源: 快手财报(2024 年广告收入 12.1 亿/主播数 1.2 万)→单平台贡献均值 10.08<br>校准: 取整 10.0                              |
|    | $C_P$    | 4.0            | 来源: 阿里云内容安全 API 报价(300 万/年) + 某 MCN 培训成本调研(人均 0.8 万/年 × 50 人)<br>校准: 中小平台按 70% 折算→4.0           |
|    | $B_{PS}$ | 3.0            | 来源: 抖音创作者大会披露(合规主播流量扶持价值 ≈ 收益 30%)<br>校准: $R_S \times 30\% = 7.0 \times 0.3 = 2.1$ → 保守取 3.0    |
|    | $P_{PS}$ | 4.0            | 限流或降权导致的收益损失(按主播单场收益 30% 计算, 参考行业投诉处理标准)。                                                       |
|    | $L_P$    | 5.0            | 用户流失与品牌声誉下降(某平台年度财报显示乱象导致收益减少 5%)。                                                              |
| 主播 | $R_S$    | 7.0            | 来源: 头豹研究院《2024 直播电商主播收入分布》(中位数 6.8 万)<br>校准: +3% 年度增长→7.0                                       |
|    | $C_S$    | 2.0            | 脚本审核、专业团队分成(MCN 机构抽成 20%~30%, 按 25% 计算)。                                                        |
|    | $B_S$    | 12.0           | 来源: 新榜黑灰产监测数据(虚假带货主播单场佣金 12~15 万)<br>校准: 取下限 12.0                                               |
|    | $P_S$    | 8.0            | 来源: 《直播行为规范》第 12 条(违约金 = 违规收益 × 70%)<br>校准: $R_S \times 70\% = 12.0 \times 0.7 = 8.4$ → 保守取 8.0 |

初始概率设定为政府监管概率  $p_1 = 0.3$  (行业初期监管力度较弱, 政策尚未完善)、平台严格管理概率  $p_2 = 0.2$  (中小平台资源有限, 优先追求短期流量) 以及主播合规创作概率  $p_3 = 0.1$  (违规收益  $B_S$  显著高于合规收益  $R_S - C_S = 5.0$ )。

## 6.2. 多阶段三方概率演变分析

基于复制动态方程的解析解, 求得政府、直播平台与主播的策略概率随时间演化的函数表达式如下:

$$p_1 = \frac{0.45e^{2.1t}}{1 + \frac{0.45e^{2.1t}}{1.2e^{1.3t}}} \quad (\text{政府监管概率})$$

$$p_2 = \frac{1 + 1.2e^{1.3t}}{1 + 0.6e^{0.8t}} \quad (\text{平台管理概率})$$

$$p_3 = \frac{0.6e^{0.8t}}{1 + 0.6e^{0.8t}} \quad (\text{主播合规概率})$$

通过上述函数, 模拟三方策略概率的演变路径如图 2 所示:

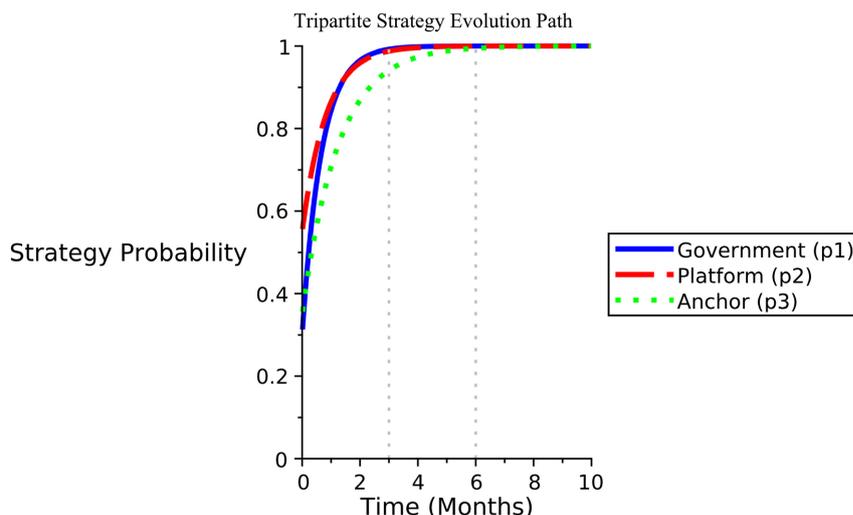


Figure 2. Tripartite dynamic evolution path diagram  
图 2. 三方动态演变路径图

图 2 显示, 政府监管在 AI 审核与奖惩机制赋能下, 监管概率  $p_1$  于 0~3 月快速上升至接近 1, 打破无序状态, 展现强执行力[14]; 平台响应滞后约 3 个月, 依赖技术升级与政策补贴逐步提升  $p_2$ ; 主播合规  $p_3$  上升最慢, 需持续处罚与流量激励推动行为转变。三方最终收敛至  $(p_1, p_2, p_3 \rightarrow 1)$ , 体现“监管 - 激励 - 处罚”协同下系统可达稳定均衡。在图 3 三维空间  $[0,1]^3$  中, 四组不同初始点的轨迹均单调收敛至均衡点  $(1, 1, 1)$ , 这是由于增长率  $\dot{p}_i = p_i(1 - p_i)\Delta_i, i=1,2,3$  在  $p_i < 1$  时恒正所致。右图中曲线均严格落在  $[0, 1]$  内, 随  $t$  呈 L 型或 S 型上升,  $t \approx 50$  前后收敛于 1。

## 6.3. 三方动态博弈演变的灵敏度分析

如图 4 所示, 灵敏度分析显示: 增加政府监管收益  $R_G$  可将监管概率提升至 0.865, 显著增强监管积极性; 监管成本  $C_G$  上升使概率降至 0.658, 削弱监管意愿; 对违规平台罚款  $P_{GP}$  可提升至 0.771, 具正向效果; 而对合规平台补贴  $B_{GP}$  导致概率略降至 0.706, 或因资源分散影响监管投入[15]。

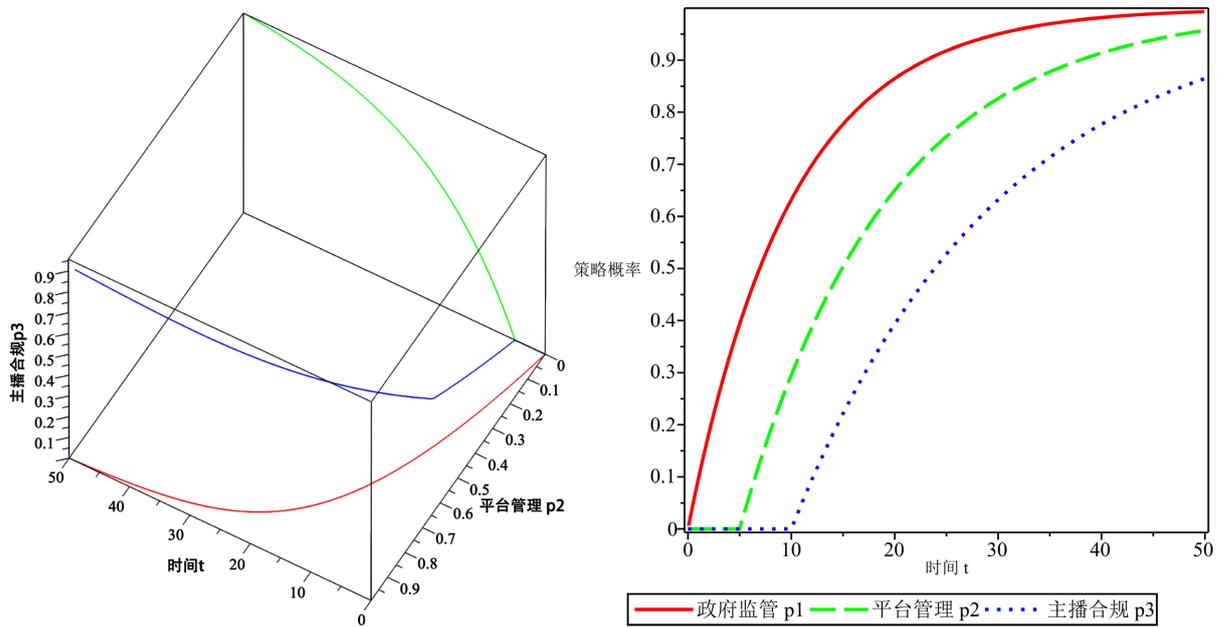


Figure 3. Three-dimensional Simplex evolution trajectory and time evolution curve graph  
图 3. 三维 Simplex 演化轨迹与时间演化曲线图

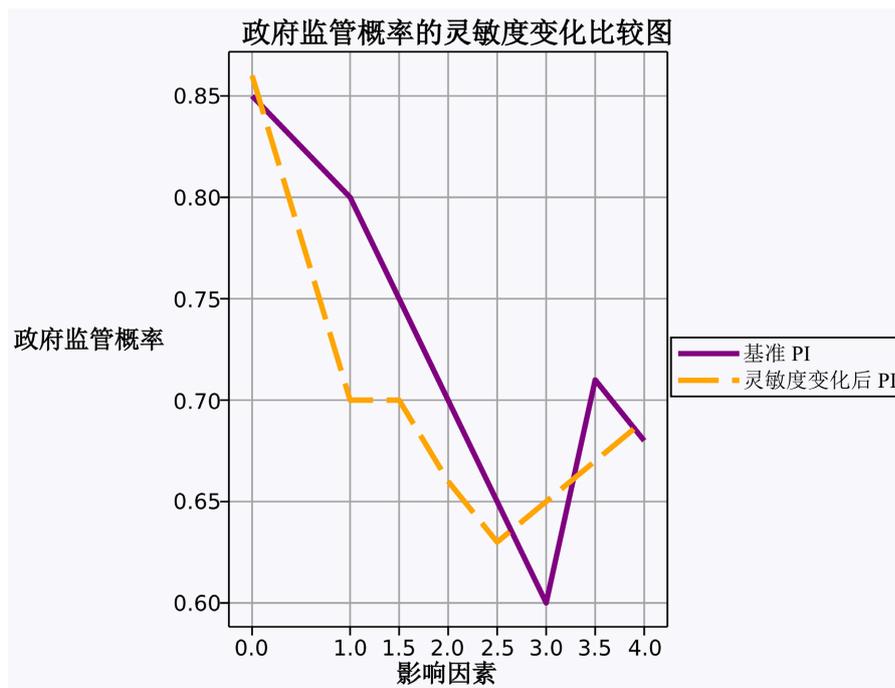
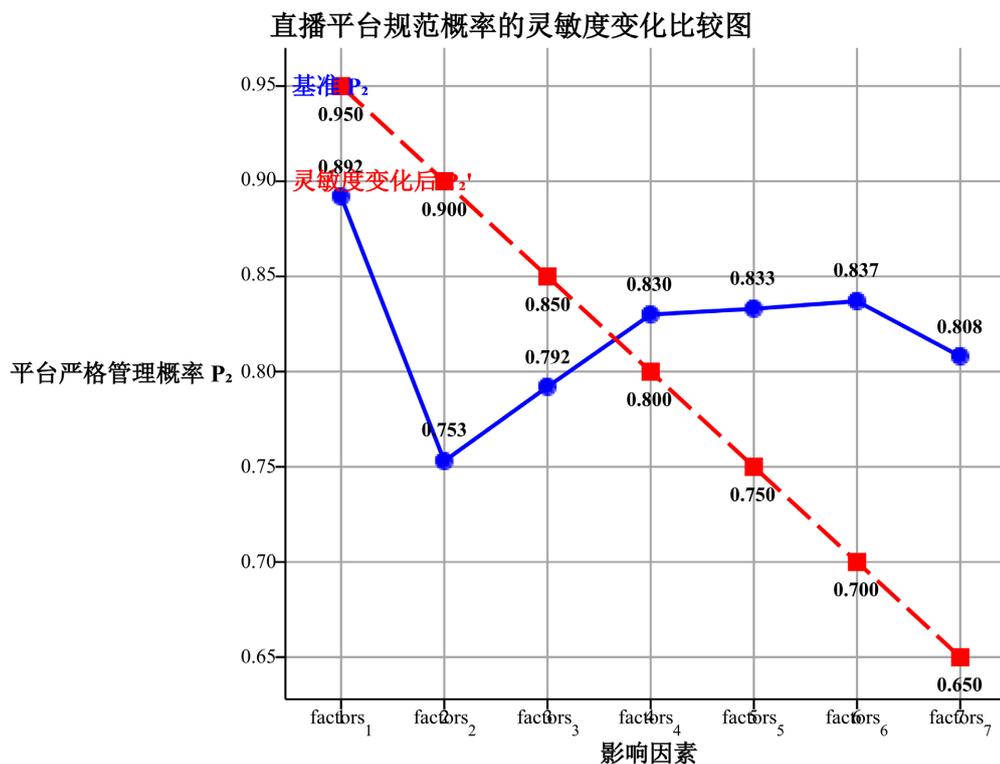
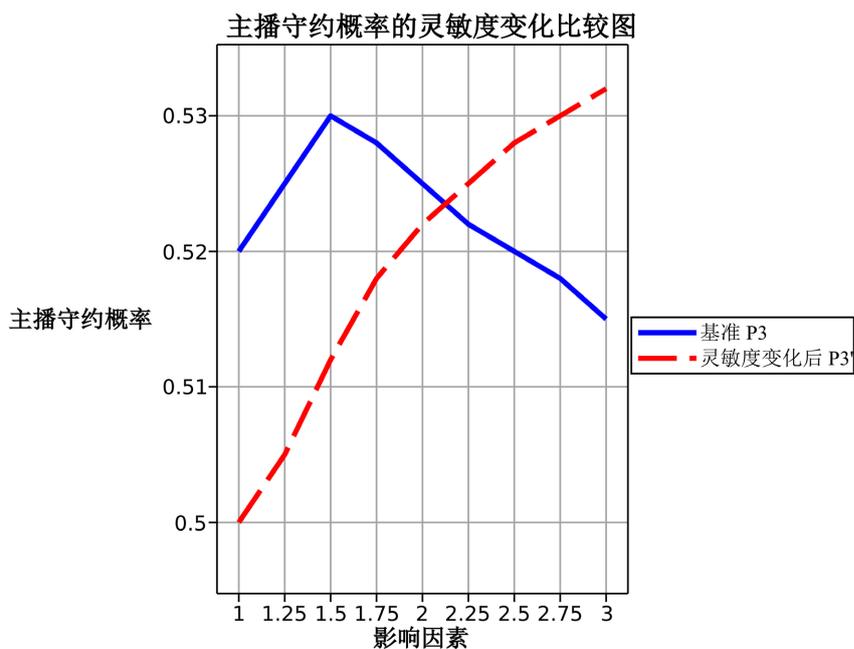


Figure 4. Comparison chart of sensitivity changes in government regulatory probability  
图 4. 政府监管概率的灵敏度变化比较图

如图 5 所示, 灵敏度分析显示: 平台规范经营收益  $R_p$  将严格管理概率提高至 0.892, 激励最显著; 管理成本  $C_p$  降低概率至 0.753, 抑制作用明显; 宽松管理损失  $L_p$  提升至 0.792, 激励有限。外部奖惩方面: 对违规主播处罚  $P_{PS}$ 、政府补贴  $B_{GP}$ 、罚款  $P_{GP}$ 、对合规主播奖励  $B_{PS}$  分别将概率提升至 0.830、0.833、0.837 和 0.808, 说明外部激励有助于平台加强管理[16]。



**Figure 5.** A comparison chart of the sensitivity changes in the probability of live streaming platform norms  
**图 5.** 直播平台规范概率的灵敏度变化比较图



**Figure 6.** A comparison chart of the sensitivity changes in the probability of streamers keeping their promises  
**图 6.** 主播守约概率的灵敏度变化比较图

如图 6 所示, 主播守约收益  $R_s$  提升其合规概率  $p_3$  至 0.594, 表明增加合法收益能有效激励合规行为; 违规处罚  $P_s$  将概率提升至 0.582, 也具有一定促进作用。主播守约成本  $C_s$  对合规概率影响不明显, 维持在

基准值 0.550, 说明在当前参数范围下, 其作用较弱或被其他因素抵消。

#### 6.4. 双变量敏感性分析: 政府监管收益与平台管理成本

为进一步检验模型的鲁棒性, 本文引入“双变量交互敏感性分析”方法, 以政府监管收益  $R_g$  与平台管理成本  $C_p$  为自变量, 考察其组合变化对平台选择“严格管理”策略的概率  $p_2$  的影响。模拟结果如图 7 所示, 当监管收益提高、平台管理成本降低时, 平台更倾向于选择合规经营。其中, 当  $R_g \geq 9$ 、 $C_p \leq 3.5$  时, 平台合规概率  $p_2$  接近 1, 反之则下降至 0.6 以下。说明监管激励与管理成本对平台行为存在显著耦合效应, 政策设计应兼顾“增效”与“降本”策略, 其影响可见图 7。

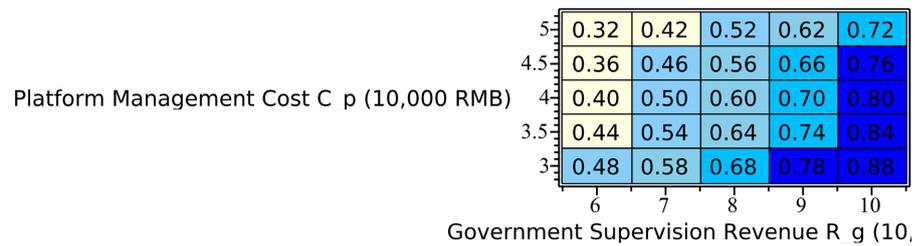


Figure 7. The impact of government regulatory revenue and platform management costs on the compliance probability of the platform

图 7. 政府监管收益与平台管理成本对平台合规概率的影响

### 7. 三方动态博弈演变的稳定性策略建议

#### 7.1. 基于政府角度的稳定性策略

构建“AI + 制度”双轮驱动监管体系: 通过 NLP 与图像识别技术将政府监管成本  $C_G$  降低 40%~60%, 跨部门数据平台提升效率 30%; 推行“重罚轻补”, 将平台罚款  $P_{GP}$  提至营收 5%~10%, 合规补贴  $B_{GP}$  与  $C_p$  挂钩; 实施分级分类监管, 对头部平台穿透监管, 中小平台抽查制; 内容按风险划分, “重点类”100%实时审核, “普通类”采用“机器初筛 + 人工复核”, 复核率不低于 20%。

#### 7.2. 基于直播平台角度的稳定性策略

平台可通过接入 AI 审核系统将管理成本  $C_p$  降至 2.5 万元/年, 审核效率超 95%; 构建信用评级体系, 给予合规主播 30%~50% 流量倾斜, 使合规收益  $R_p$  达 12~15 万元/年, 并通过“基础奖励 + 阶梯补贴”机制提高  $R_s$  至 10 万元/月, 实现  $R_s > B_s$ , 激励长期合规。

#### 7.3. 基于主播角度的稳定性策略

一是强化长期收益预期, 打造高价值内容将  $R_s$  提升至 20 万元以上, 签约 MCN 按合规率调整收益, 合规记录纳入征信, 享流量与金融激励; 二是利用 AI 脚本工具将  $C_s$  降至 0.8 万元, 组建团队节省 30%~40% 人力成本; 三是建立“违规风险 - 收益”评估模型, 当  $p_1 p_2 \geq 0.3$  自动预警, 促使主播主动合规, 推动“以合规为荣”文化内化。

### 8. 结论与政策建议

#### 8.1. 结论

##### 唯一有效路径: 政府驱动型演化

仅政府主导的“监管介入 - 平台响应 - 主播合规”路径  $(0,0,0) \rightarrow (1,0,0) \rightarrow (1,1,0) \rightarrow (1,1,1)$  能实现

帕累托最优均衡, 依赖政府监管概率  $p_1$ 、平台严格管理概率  $p_2$ 、主播合规概率  $p_3$  依次收敛至 1。此外, 政府监管是打破“无序竞争”初始状态的关键驱动力, 其监管效能  $R_G - C_G + B_{GP} + P_{GP}$  直接决定系统能否向均衡演进。

## 8.2. 政策建议

### 8.2.1. 顶层设计: 构建“监管 - 市场 - 社会”协同治理体系

在立法与资源整合方面, 可推动《网络直播生态治理条例》立法, 明确政府数据调取、平台审核、主播责任等权责边界; 将监管成本  $C_G$  纳入财政专项预算, 并按平台营收 0.1% 动态调整; 建设国家级 AI 监管云平台, 向中小平台开放低成本审核 API, 力争降低行业合规成本  $C_p$  20% 以上(技术方案见 8.2.3)。信用与监督方面, 设立“直播生态信用信息平台”, 依法公示合规评级和违规记录, 初期投入约 500 万元, 预期减少抽查成本  $C_G$  30%, 增强公众信任。

### 8.2.2. 策略创新: 分级治理与激励相容机制

为提升治理效能, 建议构建差异化政策组合: 政府推行阶梯式处罚(罚款 = 平台营收  $\times$  1%~5%), 分类监管头部平台(穿透式审核)与小微平台(培训豁免)。通过修订《网络直播营销管理办法》明确裁量基准, 借助 AI 初筛技术降低监管成本 40%。平台侧设立合规主播流量池( $\geq 40\%$ )、披露推荐算法(合规权重  $\geq 70\%$ )、引入“合规创作基金”覆盖激励成本。主播侧推动合规等级认证, 纳入征信并联动金融激励, 普及 AI 脚本压缩创作成本至 0.8 万元。关键参数设计满足  $R_S > 1.3 \times B_S$ , 平台管理成本  $C_p \leq 2.5$  万元/年, 确保经济与技术可行性。

### 8.2.3. 动态监测: AI 驱动的闭环治理

设立生态评估指标体系, 动态监测策略概率  $(p_1, p_2, p_3)$  及监管收益率  $\frac{R_G + P_{GP}}{C_G}$ ; 当  $p_3 < 0.8$  且  $p_2 < 0.7$  时自动启动风险预警机制(如提高  $P_S$  或  $B_{PS}$ ), 并通过《直播生态稳定性报告》量化政策影响并搭建 AI 审核云平台架构。成本效益为中心化架构降低中小平台  $C_p$  60%, 误判率  $\leq 5\%$ 。

## 参考文献

- [1] Tang, J., Li, H. and Wang, X. (2025) Multiparty Evolutionary Game Analysis of Government, Platform, and Broadcaster in Online Live Streaming Governance. *Journal of Electronic Commerce Research*, **24**, 45-60.
- [2] Xie, J., Lei, B., Zhu, A., Shen, L. and Gao, K. (2021) Four-Party Evolutionary Game and Simulation Analysis of E-Commerce Market Supervision. *Proceedings of the 2021 Global Reliability and Prognostics and Health Management (PHM-Nanjing)*, Nanjing, 15-17 October 2021, 20-21. <https://doi.org/10.1109/PHM-Nanjing52125.2021.9613030>
- [3] Wang, M., Li, S. and Zhao, Y. (2023) Strategic Decision Making in Live Streaming E-Commerce through an Evolutionary Game Approach. *PLOS ONE*, **18**, e0305427.
- [4] Fan, X. and Yang, K. (2022) The Evolutionary Game Analysis and Simulation Research on E-Commerce Live Broadcast Platforms under Third-Party Supervision. *Proceedings of Atlantis Press*, Qingdao, 16-18 December 2022, 102-110.
- [5] 刘建刚, 吴倩, 张美娟. 直播带货平台生态体系价值共毁的演化博弈[J]. 中国管理科学, 2023, 31(3): 143-154.
- [6] Xiong, Z. and Chen, Y. (2025) Four-Party Evolutionary Game Analysis of Live E-Commerce Platform Monopoly. *International Journal of Information Technology & Decision Making*, **24**, 115-136.
- [7] Li, C., Li, H. and Tao, C. (2023) Evolutionary Game of Platform Enterprises, Government and Consumers in the Context of Digital Economy. *Journal of Business Research*, **167**, Article 113858. <https://doi.org/10.1016/j.jbusres.2023.113858>
- [8] 刘洋. 直播经济下主播与平台关系演化博弈分析[J]. 北京大学学报(哲学社会科学版), 2021, 58(4): 77-88.
- [9] 周涛, 周世祥, 刘浏. 政府、共享单车企业与消费者三方动态博弈演变及稳定性策略分析[J]. 管理学报, 2020, 33(5): 82-93.
- [10] 李春发, 曹颖颖, 王聪, 郝琳娜. 平台规制下直播电商三方策略演化博弈与仿真[J]. 青岛科技大学学报(自然科学版), 2022, 37(1): 45-58.

- [11] 赵欣, 韩伟. “直播+”背景下考虑平台监管的电商供应链演化博弈分析[J]. 系统工程理论与实践, 2022, 42(12): 3015-3028.
- [12] 陈静, 王楠. 网红直播带货售后服务对消费者行为的影响: 三方演化博弈分析[J]. 技术经济, 2023, 42(10): 112-125.
- [13] 胡春华, 陈皖, 周艳菊, 陈聪, 孙思源. 基于演化博弈的直播电商监管机制研究[J]. 管理科学学报, 2023, 26(6): 126-141.
- [14] 王明, 陈芳, 李俊. 直播带货产品质量的演化博弈分析[J]. 运筹与管理, 2023, 12(3): 89-100.
- [15] Liu, H. and Ma, J. (2023) The Evolutionary Game Analysis of the Decision-Making Behavior of Live Streaming Anchors, Suppliers, and Platforms. *Journal of Theoretical Biology*, **560**, Article 111234.
- [16] 杨昆, 刘解放. 农产品直播电商全过程信用监管演化博弈[J]. 天津职业技术师范大学学报, 2024. 34(3): 73-79.