

# 人工智能算法决策与企业研发“合谋”

徐 雷, 李 政, 郭晓玲

(辽宁大学经济学部经济学院, 辽宁 沈阳 110036)

**摘要:** 在一个企业研发博弈环境下, 运用强化学习 Q-learning 算法对多智能体进行联合训练, 探讨人工智能决策下企业研发竞争的可能结果。主要结论: 第一, 与纳什均衡结果相比, 人工智能决策下企业研发投入更低而利润水平更高的情况是普遍存在的; 第二, 智能体数量的增加能够降低“合谋”的实现概率; 第三, 研发的外溢度会促进智能体“合谋”, 且在智能体数量较多时更易实现; 第四, 学习率的提高能够降低“合谋”的实现概率, 且在智能体数量较多且研发外溢度较低时作用更显著。基于以上研究结果, 提出加强人工智能算法在企业决策层面应用管理的政策建议, 为数字经济时代的算法监管政策和反垄断规制的完善提供了理论支持。

**关键词:** 人工智能; Q-learning 算法; 研发博弈; 企业创新

**中图分类号:** F406.3 **文献标识码:** A **文章编号:** 1005-0566(2024)06-0214-11

## Algorithms decision-making and corporate R&D “collusion”

XU Lei, LI Zheng, GUO Xiaoling

(School of Economics, Liaoning University, Shenyang 110036, China)

**Abstract:** This paper uses Q-learning algorithm to train agents in a R&D competition environment, and discusses the possible results caused by artificial intelligence decision. The results show that: first, it is common for lower R&D investment and higher profit levels than Nash equilibrium results. Second, the increase in the number of agents can reduce the probability of achieving “collusion”. Third, the spillover degree of R&D will promote the “collusion” of agents, and it is easier to achieve when the number of agents is large. Fourth, the increase in the learning rate can reduce the probability of achieving “collusion”, and it is more obvious when the number of agents is large and the spillover of R&D is low. Based on the above research results, this paper puts forward three suggestions to strengthen the management of artificial intelligence application at the decision-making level of enterprises, which provides theoretical support for the continuous improvement of anti-monopoly regulation in the era of digital economy.

**Key words:** artificial intelligence; Q-learning algorithm; R&D game; enterprise innovation

随着数字经济时代的到来, 企业在重要决策中越来越多地使用了计算机程序辅助。其中的一些程序利用了人工智能机器学习 (Machine Learning) 算法, 能够在自己过去的行动经验中自主学习, 这引起

了学界对这种人工智能决策是否会在没有事前沟通的条件下形成“合谋”的关注, 而这将对现有的反垄断法形成重要挑战。其原因在于, 由机器学习算法实施的“合谋”行为具有两个突出的特点: 一

收稿日期: 2024-02-29 修回日期: 2024-04-18

**基金项目:** 国家社会科学基金重点项目“新时代以创新为引领增强国有经济‘五力’研究”(22AZD032); 国家社会科学基金青年项目“数字经济背景下互联网平台相关市场界定方法重构与反垄断规制研究”(21CJY043)。

**作者简介:** 徐雷 (1983—), 男, 山东泰安人, 辽宁大学经济学部经济学院教授, 经济学博士, 研究方向为产业组织与博弈论、算法合谋与反垄断。

是人工智能行为主体模糊,不具有协同行为的主观意愿;二是具有竞争关系的经营者(算法)之间没有协议甚至没有任何沟通<sup>[1-3]</sup>。这导致算法间的“合谋”成为默契合谋(tacit collusion)<sup>[4-6]</sup>。对默契合谋是否应成为反垄断执法对象的讨论甚至早于算法合谋的出现,一些学者认为只要不存在协议,就不能将事实上的协同结果判定为合谋<sup>[7-8]</sup>,但更多的学者持相反的观点,他们认为反垄断法应将默契合谋界定为垄断协议,从而对其进行规制<sup>[9-13]</sup>。尽管如此,反垄断法对此种默契合谋进行取证和裁量时,仍面临责任定位困难、框架体系失衡、算法控制与算法合谋协议概念局限等诸多难题<sup>[14-17]</sup>。

显然,研发合谋是被反垄断法禁止的。但是,如果企业的研发投入决策是在人工智能辅助下进行的,是否会导致企业形成默契合谋呢?当前,有两条技术路线可以验证算法“合谋”是否存在:第一条技术路线是从真实的市场结果进行判断,如有研究表明人工智能算法能够辅助决策者提升市场运营效率<sup>[18]</sup>,还有研究证实在德国汽油零售市场中人工智能定价算法的采用与加油站的利润率大幅增加相关<sup>[19]</sup>;第二条技术路线是考察人工智能算法决策的仿真结果来分析合谋的可能性,此类文献形成了诸多发现,如在双寡头 Bertrand 动态竞争模型中定价算法会实现接近垄断价格的均衡<sup>[20]</sup>,智能体在价格竞争中能够学会采用惩罚策略从而促进合谋<sup>[21]</sup>,在序贯价格竞争中算法决策可能导致智能体实现一定程度的合谋<sup>[22]</sup>。

本文在已有研究的基础上,采用第二条技术路线考察算法决策能否导致企业研发“合谋”,我们的研究动机如下:第一,现有研究集中于在价格竞争或产量竞争中考察算法决策导致的后果,很少关注到企业研发竞争中的算法决策问题。但对我国企业来说,人工智能技术的运用得到大力支持,我们急需知道是否存在人工智能赋能下的技术进步阻碍效应;第二,与制定价格和制定产量的活动不同,企业的研发投入存在外溢性,这种外溢性对算法决策将产生怎样的影响,这种影响是否与博弈理论的逻辑分析相一致都是需要进行考察

的重要问题;第三,从对强化学习智能体的训练上看,尽管在一个固定环境下的单智能体行为已被证明是最优的<sup>[23]</sup>,但多智能体联合训练的行为逻辑仍然难以给出明确的理论解释<sup>[24]</sup>。因此,本文将在一个研发竞争环境中运用仿真分析考察人工智能算法决策的可能后果,为以上问题提供答案,以此对人工智能参与决策的规制办法的制定和完善提供有力支持。

### 一、企业研发竞争环境设定

本文构建一个研发投入具有溢出效应的企业研发博弈模型,为算法决策提供一个训练环境。假设企业在一个同质产品市场中进行数量竞争,市场中的需求函数为:

$$P = a - Q(Q = \sum_{i=1}^n q_i) \quad (1)$$

企业的研发活动是工艺创新型的,即所开发的技术可以降低生产特定产品的成本。以  $x_i$  表示企业  $i$  的研发量( $i = 1, 2, \dots, n$ ),并以  $c_i(x_i, x_{-i})$  表示企业  $i$  的单位生产成本,这里的  $x_{-i}$  表示除  $i$  外其他企业研发投入向量,也即企业的单位生产成本不仅取决于自身的研发投入,也受到其他企业研发投入的影响。正式地,令:

$$c_i(x_i, x_{-i}) \equiv \alpha - x_i - \beta \sum_{j \neq i} x_j \\ i = 1, 2, \dots, n; j \neq i; \beta > 0 \quad (2)$$

我们假设(2)式中  $\beta > 0$ ,这就刻画了企业研发活动所形成的正的外溢效应。另外,研发对企业来说是有成本的,我们将企业  $i$  投入的研究水平  $x_i$  的成本表示为  $TC_i(x_i)$ ,并有:

$$TC_i(x_i) = \frac{x_i^2}{\gamma} \quad (3)$$

#### (一)非研发合作

在此博弈中,企业首先选择研发投入水平,从而确定成本函数,之后再选择产量水平。为求解博弈均衡,首先求解企业在产量竞争中的均衡解,然后采用逆向归纳法求解企业的研发投入水平。企业的产量竞争发生在研发投入之后,此时,  $n$  个企业竞争的均衡产量和利润水平即为古诺均衡,由下式给出:

$$q_i = \frac{a - nc_i + \sum_{j \neq i} c_j}{n+1} \quad (4)$$

$$\pi_i = \frac{(a - nc_i + \sum_{j \neq i} c_j)^2}{(n+1)^2} \quad (5)$$

我们再分析企业的研发投入:每个企业在其他企业研发水平给定的情况下,以不合作的态度选定其研发水平,此时要求解一个研发水平的纳什均衡。为此,将式(2)代入式(5),再减去企业的研发成本,可以得到企业的利润函数为:

$$\pi_i = \frac{1}{(n+1)^2} \left\{ \alpha - a - [n - (n-1)\beta] x_i + (1-2\beta) \sum_{j \neq i} x_j \right\}^2 - \frac{(x_i)^2}{\gamma} \quad (6)$$

对一个给定的研发水平  $x_j$ , 企业  $i$  选择  $x_i$  以最大化其利润  $\pi_i$ , 一阶条件为:

$$\frac{\partial \pi_i}{\partial x_i} = \frac{-2(n - (n-1)\beta)}{(n+1)^2} \left\{ \alpha - a - [n - (n-1)\beta] x_i + (1-2\beta) \sum_{j \neq i} x_j \right\} - \frac{2x_i}{\gamma} = 0 \quad (7)$$

企业利润最大化需要满足二阶条件  $\partial^2 \pi / \partial x_i^2 < 0$ , 也即:

$$\frac{n - (n+1)/\sqrt{\gamma}}{n-1} < \beta < \frac{n + (n+1)/\sqrt{\gamma}}{n-1} \quad (8)$$

由于利润函数在各个企业之间是对称的,我们可以求一个对称的纳什均衡,在该均衡下,有  $x_1 = x_2 = \dots = x_n \equiv x^{nc}$ , 这里的  $x^{nc}$  即是在非合作条件下,每个企业均衡的研发投入水平,它由下式给出:

$$x^{nc} = \frac{(a - \alpha)(2\beta + 2n - 2\beta n)}{\left( \frac{2}{\gamma} - \frac{(2\beta + 2n - 2\beta n)(\beta n - \beta + 1)}{(n+1)^2} \right) (n+1)^2} \quad (9)$$

## (二) 合作研发

在合作研发的情况下,企业共同选择能够使联合利润最大化的研发投入水平。即企业联合选

择各自的研发投入水平  $x_i$ , 以最大化联合利润:

$$\max_{x_i} \sum_{i=1}^n \pi_i \quad (10)$$

式(10)中,  $\pi_i$  由式(6)给出,  $i = 1, 2, \dots, n$ 。一阶条件为:

$$0 = \frac{\partial \sum_{i=1}^n \pi_i}{\partial x_i} = \frac{\partial \pi_1}{\partial x_i} + \frac{\partial \pi_2}{\partial x_i} + \dots + \frac{\partial \pi_i}{\partial x_i} + \dots + \frac{\partial \pi_n}{\partial x_i} \quad (11)$$

式(11)中,  $\partial \pi_i / \partial x_i$  项衡量企业  $i$  从研发投入  $x_i$  的微小改进获取的边际利润,其他项衡量企业  $j$  ( $j \neq i$ ) 的利润的边际增加,这源于企业  $i$  的研发努力增加带来的溢出效应。因此,式(11)可写为:

$$0 = \frac{-2[n - (n-1)\beta]}{(n+1)^2} \left\{ \alpha - a - [n - (n-1)\beta] x_i + (1-2\beta) \sum_{j \neq i} x_j \right\} - \frac{2x_i}{\gamma} + \frac{2(1-2\beta)}{(n+1)^2} \sum_{j \neq i} [\alpha - a - (n - (n-1)\beta) x_j + (1-2\beta) \sum_{k \neq j} x_k] \quad (12)$$

由于利润函数在各个企业之间是对称的,可以求解一个对称的合作均衡。在此均衡中,  $x_1 = x_2 = \dots = x_n \equiv x^c$ , 其中,  $x^c$  是每个企业在合作均衡中的研发投入水平:

$$x^c = \frac{\gamma(a - \alpha - \alpha\beta + \alpha\beta + \alpha\beta n - \alpha\beta n)}{2\gamma\beta^2 n - \gamma\beta^2 n^2 - \gamma\beta^2 - 2\gamma\beta n + 2\gamma\beta + n^2 + 2n - \gamma + 1} \quad (13)$$

## (三) 对模型参数的数值设定

令  $a = 100$ ,  $\alpha = 50$ ,  $\gamma = 2$ , 依据(8)式给出的二阶条件,我们设定厂商数量  $n$  分别取 2, 3 和 4, 研发外溢度  $\beta$  分别取 0.2, 0.3, 0.4 和 0.5。在这样的参数设定下,我们计算了不同情况下厂商在非合作(NE)与合作(CE)均衡状态中的研发投入水平(R&D)、产量水平( $Q$ )和利润水平( $\pi$ ),表1展示了计算结果。可以看出,在合作状态下,企业的研发投入和产量水平比非合作状态下更小,利润则更高。但随着研发外溢程度( $\beta$ )的提高,二者的差距不断减小,在  $\beta = 0.5$  时,合作与非合作状态均衡完全相同。

表1 对非合作研发与研发状态下企业研发、产量和利润水平的比较

| 企业数量 | 观测指标  | $\beta=0.2$ |        | $\beta=0.3$ |        | $\beta=0.4$ |        | $\beta=0.5$ |        |
|------|-------|-------------|--------|-------------|--------|-------------|--------|-------------|--------|
|      |       | NE          | CE     | NE          | CE     | NE          | CE     | NE          | CE     |
| 2    | R&D   | 38.46       | 19.61  | 37.12       | 23.13  | 35.40       | 27.56  | 33.33       | 33.33  |
|      | Q     | 32.05       | 24.51  | 32.75       | 26.69  | 33.19       | 29.53  | 33.33       | 33.33  |
|      | $\pi$ | 287.64      | 408.50 | 383.76      | 444.84 | 474.78      | 492.13 | 555.56      | 555.56 |
| 3    | R&D   | 29.82       | 11.59  | 28.85       | 14.71  | 27.23       | 18.91  | 25.00       | 25.00  |
|      | Q     | 22.94       | 16.56  | 24.04       | 18.38  | 24.75       | 21.01  | 25.00       | 25.00  |
|      | $\pi$ | 81.54       | 206.95 | 161.80      | 229.78 | 242.01      | 262.61 | 312.5       | 312.5  |
| 4    | R&D   | 24.08       | 8.05   | 23.45       | 10.69  | 22.08       | 14.36  | 20.00       | 20.00  |
|      | Q     | 17.71       | 12.58  | 18.91       | 14.06  | 19.72       | 16.32  | 20.00       | 20.00  |
|      | $\pi$ | 23.57       | 125.75 | 82.68       | 140.61 | 144.92      | 163.19 | 200.00      | 20.00  |

那么,如果企业间重复进行这一研发博弈,会出现什么结果呢?实际上,现代博弈理论已经证明,对于一个仅有唯一纯策略纳什均衡的博弈,它的有限次重复博弈的纳什均衡不会发生改变。对于我们所设定的企业研发博弈,纯策略纳什均衡仅有一个,那么,机器学习算法博弈的结果是否也会向纳什均衡收敛,研发活动的外溢度对博弈结果是否存在影响,这是我们需要回答的问题。下面,对本文所选用的机器学习算法进行简要介绍,之后展示并分析训练结果以回答这些问题。

## 二、算法选择与参数设定

### (一) 算法选择

强化学习是机器学习算法的一种<sup>①</sup>,它是目标导向的计算方式,在这种计算方式下,计算机通过与未知的动态环境进行互动来学习如何完成一项任务,即如何根据情况采取行动,从而实现数值奖励信号最大化。计算机完成任务的过程不会接受人类施加的动作指令,也不需要显著式编程(explicitly programmed),而是通过一系列自行尝试和自主决策去发现回报最高的行动方案。

尽管在强化学习领域有诸多学习算法,但本文基于以下几点原因选用Q-learning算法对智能体进行训练。第一,希望所使用的算法是在实践中被广泛运用的,但企业却极少公布他们所使用的算法。然而,在计算科学领域,Q-learning算法已经得到广泛应用,鉴于该领域理论与实践具有

紧密联系,我们可以相信Q-learning算法也会运用到企业正在使用或即将使用的决策软件中。第二,Q-learning算法的编写和运行比较简单,运用matlab或python等软件可以通过几个参数的设定较为简易地完成代码编写,并且这些参数在经济学和博弈论的范畴内能够给予明确的解释<sup>[25]</sup>。第三,Q-learning算法与更复杂的算法共享相同的架构,这些算法在很多领域取得了重要的成功<sup>[26-28]</sup>,这使我们相信此算法在企业决策领域也具有很大的应用潜力。基于以上原因,本文选择了Q-learning算法。此算法当然也存在缺点,比较明显的是其运算速度慢,这是由其迭代计算方式导致的。

### (二) 参数设定

与所有的强化学习算法相同,Q-learning算法也是基于已有经验选择行动,这就需要在继续学习和利用已经获得的知识之间进行权衡<sup>②</sup>。本文采用的权衡方式是得到广泛使用的 $\epsilon$ -贪婪( $\epsilon$ -greedy)探索模型<sup>③</sup>。依据这个模型,学习算法以 $\epsilon$ 的概率在所有行动中进行均匀化的随机选择(探索模式),以 $1-\epsilon$ 的概率选择当前的最优行动(优选模式)。随着学习次数的增加,智能体对环境和收益的了解不断增强,采用探索模式的概率应该不断下降。因此,设定一个 $\epsilon$ 的消减率为 $\xi$ , $\epsilon$ 随时间的变化由该消减率定义,即有:

$$\epsilon(t) = \epsilon \cdot (1 - \xi)^t \quad (14)$$

① 机器学习算法可分为有监督学习(supervised learning)、无监督学习(unsupervised learning)和强化学习(reinforcement learning)3种类型。

② 由于Q-learning每次计算只能更新一个矩阵要素,而收敛到真实的Q矩阵要求每个要素需要被更新多次,这导致学习过程非常耗时,并且状态和行动空间越大,所需的迭代次数和计算时间就会越高。

③ 实际上,在Q-learning算法中还有更加复杂的探索模式的计算方法,如Boltzman探索模型也在Q-learning智能体训练中得到较为广泛的应用,但本文根据所研究的内容选择 $\epsilon$ -贪婪模型,能够使智能体的训练过程更加简洁和便于理解。

式(14)表明,在初始期,智能体采用探索模式的概率为  $\varepsilon$ ,随着学习的不断进行,这一概率会不断减小,直至趋近于 0。但在实际的智能体训练过程中,往往还要为  $\varepsilon$  设定一个最低的  $\underline{\varepsilon}$  值,我们将其记为  $\underline{\varepsilon}$ ,随着训练过程的进展使得  $\varepsilon$  低于  $\underline{\varepsilon}$  时,智能体将以  $\underline{\varepsilon}$  作为采用探索模式的概率,也即保证智能体始终存在探索的机会。

本文考察人工智能在研发博弈中的行为选择,与现有文献对价格和产量博弈的诸多探讨不同的是,企业的研发投入存在外溢性,因此本文首先基于外溢度  $\beta$  的不同设定对智能体进行分别训练。由于  $\beta$  在大于 0.5 时企业合作均衡下的研发投入水平和产量水平大于非合作均衡,这与现实状态不符,我们仅考虑  $\beta$  小于 0.5 的情况。另外,若  $\beta$  小于 0.2,当企业数量大于等于 4 时企业利润最大化的二阶条件(8)式无法得到满足,因此,我们排除  $\beta$  小于 0.2 的情况。因此, $\beta$  值被分别设定为 0.2、0.3 和 0.4,同时,企业数量分别设定为 2、3 和 4 家。另外,借鉴现有文献,我们令  $\delta = 0.9$ ,  $\varepsilon = 0.9$ ,  $\underline{\varepsilon} = 0.01$ , 学习率  $\theta$  则分别设定为 0.25、0.50 和 0.75 三种不同情况。

依据  $\beta$ ,  $n$  和  $\theta$  的不同组合可有 27 种智能体训练环境,由于随机数对智能体训练结果产生影响,我们在每种环境下对智能体进行 100 次训练并考察均值。由于 Q-learning 算法要求智能体的策略集为离散,同时为了简化博弈过程节省训练时间,我们将智能体的可选行动限定为 3 个:低研发投入(L),即合谋结果的研发投入水平;中研发投入(M),即纳什均衡的研发投入水平与合谋结果的研发投入水平的均值;高研发投入(H),即纳什均衡结果的研发投入水平。为了在每次训练中能够让智能体有足够的机会尝试每一种策略选择和策略组合,我们将每次训练的片段数(episode)设定为策略组合数量的 100 倍。在每次训练中,如果最后一个片段的得益与最后 50 个片段的平均得益相差小于 1%,则认为训练结果收敛而可以接受,否则不接受此次训练结果,最终的行动结果即为最后 50 个片段的平均结果。

为了更直观地考察智能体的行为选择更接近

于非合作均衡还是合作均衡,我们设定了“合谋指数(CI)”,其形式为:

$$CI = \frac{V_{AI} - V_{NC}}{V_C - V_{NC}} \quad (15)$$

具体地,我们设定了两个合谋指数,即研发投入合谋指数( $CI_{rd}$ )和利润水平合谋指数( $CI_{pf}$ )。 $V_{AI}$  表示所考察变量(研发投入或利润水平)的智能体训练结果值,  $V_{NC}$  表示纳什均衡值,  $V_C$  表示合谋结果值。因此,该指数取值在  $[0, 1]$  区间内,越接近 0 意味着人工智能决策越接近于纳什均衡,越接近 1 意味着人工智能决策越接近于合谋。下面,我们对智能体训练结果进行展示。

### 三、智能体训练结果

#### (一) $\beta = 0.2$ 的情况

从表 2 可以看出,在  $\beta = 0.2$  时,智能体研发决策的结果普遍偏离了理论模型中的纳什均衡,即形成了一定程度的“合谋”,这与现有文献的研究结论相一致。下面,我们对训练结果进行具体介绍,并阐述初步的发现。

第一,随着智能体数量的增加,合谋指数下降,NE 实现次数具有增加趋势,CE 实现次数显著下降。从合谋指数上看,不论是  $CI_{rd}$  还是  $CI_{pf}$ ,随着智能体数量的增加,合谋指数都出现了下降趋势。从均衡实现次数上看:①就 NE 实现次数来讲,在学习率为  $\theta = 0.50$  和  $\theta = 0.75$  时,上升趋势明显,但在习率为  $\theta = 0.25$  时,则未能发现此种趋势;②就 CE 实现次数来讲,随着智能体数量的增加,合谋结果实现的次数不断下降。

第二,随着学习率  $\theta$  的提高,合谋指数下降,NE 实现次数显著提高,CE 实现次数显著下降。从合谋指数上看, $CI_{rd}$  和  $CI_{pf}$  的变化具有同样的趋势,在  $N = 2$  时,合谋指数随  $\theta$  的提高先下降再上升,而在  $N = 3$  和  $N = 4$  时,合谋指数随  $\theta$  的提高不断下降。从均衡实现次数上看:①就 NE 实现次数来讲,除在  $N = 2$  时, $\theta$  的提高并未引起非合作均衡实现次数的增加外,在  $N = 3$  和  $N = 4$  两种情况下,随  $\theta$  的提高,非合作均衡实现次数都有显著增加;②就 CE 实现次数来讲,未能发现其随  $\theta$  改变出现明显变化。

表2  $\beta=0.2$  时智能体研发决策联合训练结果

| N | $\theta$ | 观测指标 | 均值     | 标准误     | 95% 置信区间 |        | CI      | 均衡实现次数 |    |
|---|----------|------|--------|---------|----------|--------|---------|--------|----|
|   |          |      |        |         | 下界       | 上界     |         | NE     | CE |
| 2 | 0.25     | 研发投入 | 28.57  | 0.5742  | 27.43    | 29.70  | 0.524 9 | 12     | 12 |
|   |          | 利润水平 | 373.48 | 3.882 3 | 365.78   | 381.18 | 0.710 2 |        |    |
|   | 0.50     | 研发投入 | 28.99  | 0.570 3 | 27.86    | 30.12  | 0.502 4 | 12     | 13 |
|   |          | 利润水平 | 371.58 | 3.859 4 | 363.92   | 379.23 | 0.694 5 |        |    |
|   | 0.75     | 研发投入 | 28.33  | 0.541 6 | 27.26    | 29.40  | 0.537 4 | 9      | 13 |
|   |          | 利润水平 | 376.74 | 3.552 9 | 369.69   | 383.79 | 0.737 2 |        |    |
| 3 | 0.25     | 研发投入 | 20.49  | 0.401 1 | 19.70    | 21.29  | 0.511 6 | 2      | 2  |
|   |          | 利润水平 | 176.67 | 2.902 5 | 170.92   | 182.43 | 0.758 6 |        |    |
|   | 0.50     | 研发投入 | 20.71  | 0.477 0 | 19.76    | 21.65  | 0.499 9 | 7      | 3  |
|   |          | 利润水平 | 171.15 | 3.504 7 | 164.19   | 178.10 | 0.714 5 |        |    |
|   | 0.75     | 研发投入 | 21.71  | 0.482 9 | 20.75    | 22.67  | 0.444 9 | 18     | 1  |
|   |          | 利润水平 | 163.90 | 4.269 3 | 155.43   | 172.37 | 0.656 7 |        |    |
| 4 | 0.25     | 研发投入 | 17.05  | 0.367 5 | 16.32    | 17.78  | 0.438 6 | 10     | 0  |
|   |          | 利润水平 | 92.15  | 2.964 9 | 86.27    | 98.03  | 0.671 2 |        |    |
|   | 0.50     | 研发投入 | 19.73  | 0.457 4 | 18.83    | 20.64  | 0.271 2 | 42     | 0  |
|   |          | 利润水平 | 65.96  | 3.866 7 | 58.29    | 73.64  | 0.414 9 |        |    |
|   | 0.75     | 研发投入 | 23.24  | 0.251 7 | 22.74    | 23.74  | 0.052 5 | 89     | 0  |
|   |          | 利润水平 | 33.89  | 2.255 5 | 29.41    | 38.36  | 0.101 0 |        |    |

为了更直观地展示智能体训练结果,我们以  $N=4$  为例,绘制不同行动组合实现频次图(图1)。该图横轴表示智能体1和智能体2的行动组合(即  $A_1/A_2$ ),纵轴表示智能体3和智能体4的行动组合(即  $A_3/A_4$ )。智能体有3种行动选择,即低研发投入( $L$ )、中研发投入( $M$ )和高研发投入( $H$ )。因此,每两个智能体有9种行动组合,4个智能体共有81种行动组合,图1中每个单元格表示一种行动组合。此图以颜色的深浅表示行动组合的实现次数,颜色越暗,表示实现次数越少,反之则越多。从图1中我们可以发现,随着  $\theta$  的提高,非合作均衡实现次数显著提高,但其他策略组合的实现次数并未展示出明显的规律性。

## (二) $\beta=0.3$ 的情况

从表3可以看出,在  $\beta=0.3$  时,智能体研发决

策的结果仍然普遍偏离纳什均衡结果,与  $\beta=0.2$  的情况相比,训练结果收敛到  $NE$  的次数更少。

第一,智能体数量变化产生的影响。从合谋指数上看:①就  $CI_{rd}$  来讲,在  $\theta=0.25$  和  $\theta=0.75$  时,随智能体数量增加呈现出先上升后下降的趋势,但在  $\theta=0.50$  时,则是先下降再上升;②就  $CI_{pf}$  来讲,在  $\theta=0.25$  和  $\theta=0.75$  时,随智能体数量增加先上升后下降,在  $\theta=0.50$  时,随智能体数量增加而不断上升。从均衡实现次数上看:①就  $NE$  实现次数来讲,在  $\theta=0.25$  和  $\theta=0.50$  时,随智能体数量增加,  $NE$  实现次数存在一定提升趋势,但在  $\theta=0.75$  时,  $N=2$ ,  $N=3$  和  $N=4$  时分别为12, 1和24,整体上呈现出波动上升的趋势;②就  $CE$  实现次数来讲,随智能体数量增加出现较为明显的下降趋势。

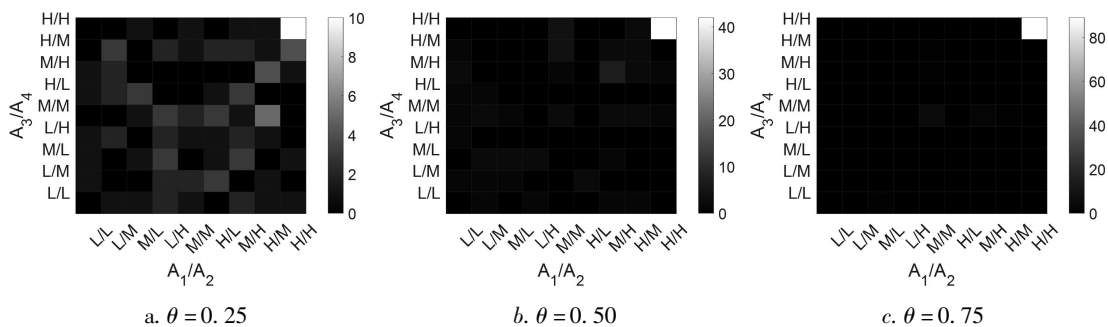
图1  $\beta=0.2$  时四智能体联合训练不同行动组合实现频次

表 3  $\beta = 0.3$  时智能体研发决策联合训练结果

| $N$ | $\theta$ | 观测指标 | 均值     | 标准误     | 95% 置信区间 |        | $CI$    | 均衡实现次数 |      |
|-----|----------|------|--------|---------|----------|--------|---------|--------|------|
|     |          |      |        |         | 下界       | 上界     |         | $NE$   | $CE$ |
| 2   | 0.25     | 研发投入 | 30.09  | 0.367 0 | 29.36    | 30.82  | 0.502 5 | 11     | 6    |
|     |          | 利润水平 | 425.44 | 1.755 6 | 421.96   | 428.93 | 0.682 5 |        |      |
|     | 0.50     | 研发投入 | 30.37  | 0.425 5 | 29.53    | 31.21  | 0.482 5 | 15     | 12   |
|     |          | 利润水平 | 422.84 | 1.966 8 | 418.93   | 426.74 | 0.639 8 |        |      |
|     | 0.75     | 研发投入 | 29.99  | 0.418 6 | 29.15    | 30.82  | 0.510 0 | 12     | 12   |
|     |          | 利润水平 | 424.45 | 1.877 2 | 420.73   | 428.18 | 0.666 2 |        |      |
| 3   | 0.25     | 研发投入 | 21.00  | 0.346 6 | 20.31    | 21.69  | 0.555 0 | 3      | 9    |
|     |          | 利润水平 | 211.72 | 1.587 8 | 208.57   | 214.87 | 0.734 3 |        |      |
|     | 0.50     | 研发投入 | 22.53  | 0.303 1 | 21.93    | 23.14  | 0.446 7 | 2      | 0    |
|     |          | 利润水平 | 205.62 | 1.553 1 | 202.53   | 208.70 | 0.644 5 |        |      |
|     | 0.75     | 研发投入 | 21.38  | 0.223 0 | 20.94    | 21.82  | 0.528 3 | 1      | 5    |
|     |          | 利润水平 | 211.09 | 1.418 6 | 208.27   | 213.90 | 0.725 0 |        |      |
| 4   | 0.25     | 研发投入 | 17.23  | 0.24    | 16.76    | 17.70  | 0.487 5 | 2      | 0    |
|     |          | 利润水平 | 123.55 | 1.16    | 121.26   | 125.85 | 0.705 5 |        |      |
|     | 0.50     | 研发投入 | 17.23  | 0.27    | 16.70    | 17.76  | 0.487 5 | 2      | 1    |
|     |          | 利润水平 | 122.63 | 1.27    | 120.10   | 125.15 | 0.689 6 |        |      |
|     | 0.75     | 研发投入 | 18.27  | 0.37    | 17.53    | 19.00  | 0.406 3 | 24     | 1    |
|     |          | 利润水平 | 114.98 | 2.05    | 110.91   | 119.04 | 0.557 6 |        |      |

第二,学习率  $\theta$  提高产生的影响。从合谋指数上看,  $CI_{rd}$  与  $CI_{pf}$  具有基本一致的变动趋势,即在  $N=2$  和  $N=3$  时,随学习率增加先下降后上升;在  $N=4$  时,随学习率增加而不断下降。从均衡实现次数上看:①就  $NE$  实现次数来讲,在  $N=2$  和  $N=3$  时随  $\theta$  变化的趋势并不明显;在  $N=4$  时,当  $\theta=0.25, \theta=0.50$  和  $\theta=0.75$  时,实现次数分别为 2, 2 和 24,呈现出较为明显的增加趋势。②  $CE$  实现次数并未随学习率的变化产生具较为明显的变化

趋势。

图 2 绘制了  $\beta=0.3$  时四智能体联合训练不同行动组合实现频次。从图 2 中可以看出,当学习率  $\theta=0.75$  时,  $NE$  出现的次数显著的增加。但在  $\theta=0.25$  和  $\theta=0.50$  两种情况中,并未发现具有明显更高实现次数的行动组合,说明智能体行动选择的收敛性受到较强随机因素的影响。与  $\beta=0.2$  时相比,仅在  $\theta=0.75$  时,  $NE$  才具有明显的优势。

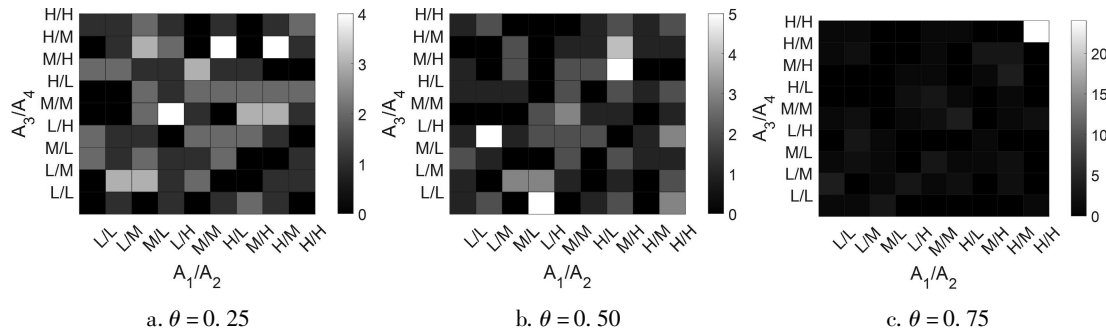


图 2  $\beta=0.3$  时四智能体联合训练不同行动组合实现频次

(三)  $\beta=0.4$  的情况

表 4 展示了在  $\beta=0.4$  时智能体研发决策的联合训练结果。与  $\beta=0.2$  和  $\beta=0.3$  两种情况相比较,  $\beta=0.4$  时的  $CI$  值整体上有所提升,  $CE$  实现次数也有所增加,这表明  $\beta$  的提高增强了智能体的“合作”倾向。

第一,智能体数量变化产生的影响。从合谋

指数上看,  $CI_{rd}$  随智能体数量的增加先下降后上升;  $CI_{pf}$  在  $\theta=0.25$  时,随智能体数量的增加不断上升,但这种趋势在  $\theta=0.50$  和  $\theta=0.75$  时变为先下降后上升。从均衡实现次数上看,  $NE$  实现次数随智能体数量的增加呈下降趋势,这与  $\beta=0.2$  和  $\beta=0.3$  时完全相反;  $CE$  实现次数则随智能体数量增加不断下降。

表4  $\beta=0.4$  时智能体研发决策联合训练结果

| N | $\theta$ | 观测指标 | 均值     | 标准误     | 95% 置信区间 |        | CI      | 均衡实现次数 |    |
|---|----------|------|--------|---------|----------|--------|---------|--------|----|
|   |          |      |        |         | 下界       | 上界     |         | NE     | CE |
| 2 | 0.25     | 研发投入 | 31.30  | 0.229 9 | 30.85    | 31.76  | 0.522 5 | 9      | 14 |
|   |          | 利润水平 | 486.03 | 0.500 3 | 485.04   | 487.02 | 0.648 4 |        |    |
|   | 0.50     | 研发投入 | 31.30  | 0.20    | 30.91    | 31.70  | 0.522 5 | 6      | 9  |
|   |          | 利润水平 | 486.27 | 0.437 7 | 485.40   | 487.14 | 0.662 4 |        |    |
|   | 0.75     | 研发投入 | 31.15  | 0.231 5 | 30.69    | 31.61  | 0.542 5 | 8      | 15 |
|   |          | 利润水平 | 486.35 | 0.494 5 | 485.36   | 487.33 | 0.666 6 |        |    |
| 3 | 0.25     | 研发投入 | 22.90  | 0.196 4 | 22.51    | 23.29  | 0.520 0 | 4      | 4  |
|   |          | 利润水平 | 255.58 | 0.490 6 | 254.60   | 256.55 | 0.658 6 |        |    |
|   | 0.50     | 研发投入 | 23.06  | 0.217 2 | 22.63    | 23.49  | 0.521 7 | 5      | 5  |
|   |          | 利润水平 | 255.16 | 0.5477  | 254.07   | 256.25 | 0.638 3 |        |    |
|   | 0.75     | 研发投入 | 23.17  | 0.197 3 | 22.78    | 23.56  | 0.488 3 | 5      | 2  |
|   |          | 利润水平 | 254.96 | 0.501 8 | 253.96   | 255.95 | 0.628 5 |        |    |
| 4 | 0.25     | 研发投入 | 18.05  | 0.160 8 | 17.73    | 18.37  | 0.522 5 | 1      | 4  |
|   |          | 利润水平 | 157.01 | 0.372 9 | 156.27   | 157.75 | 0.661 9 |        |    |
|   | 0.50     | 研发投入 | 17.96  | 0.171 8 | 17.62    | 18.30  | 0.5338  | 1      | 4  |
|   |          | 利润水平 | 157.23 | 0.392 9 | 156.45   | 158.10 | 0.673 8 |        |    |
|   | 0.75     | 研发投入 | 18.00  | 0.141 8 | 17.72    | 18.28  | 0.528 8 | 1      | 0  |
|   |          | 利润水平 | 157.39 | 0.327 4 | 156.74   | 158.04 | 0.682 6 |        |    |

第二,学习率  $\theta$  提高产生的影响。从合谋指数上看,  $CI_{rd}$  在  $N=2$  时,随  $\theta$  的提高不断提高;在  $N=3$  和  $N=4$  时,随  $\theta$  的提高而先上升后下降;  $CI_{pf}$  在  $N=2$  和  $N=4$  时,随  $\theta$  的提高不断提高;但在  $N=3$  时,随  $\theta$  的提高而不断下降。从均衡实现次数上

看,  $NE$  实现次数在  $N=2$  时随  $\theta$  提高先下降后上升,在  $N=3$  时随  $\theta$  提高逐渐上升,在  $N=4$  时保持 1 未变。  $CE$  实现次数在  $N=2$  时随  $\theta$  提高先下降后上升,在  $N=3$  时随  $\theta$  提高先上升后下降,在  $N=4$  时随  $\theta$  提高先保持未变,后下降为 0。

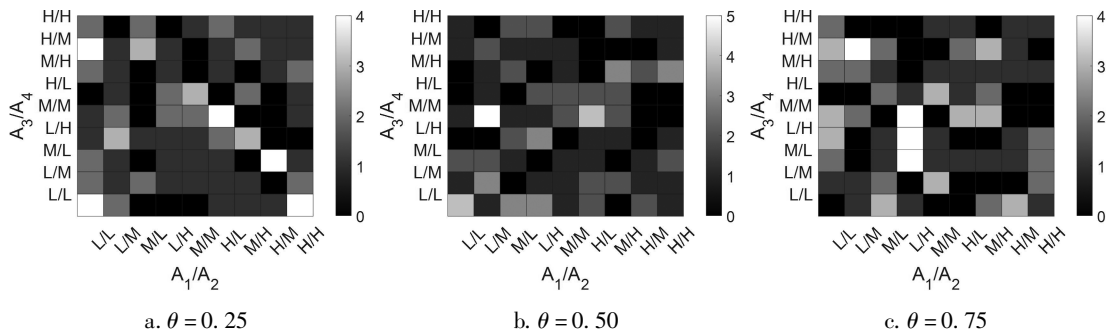
图3  $\beta=0.4$  时四智能体联合训练不同行动组合实现频次

图3绘制了  $\beta=0.4$  时四智能体联合训练不同行动组合实现频次。从3图中能够观察到  $CE$  的实现次数在  $\theta=0.75$  时出现了明显下降,其他行动组合的出现频率具有较强随机性。我们将图1、图2和图3进行整体比较,能够发现随着  $\beta$  值的提高,  $NE$  的实现次数有明显减少,而  $CE$  的实现次数有明显的增加,同时智能体行动组合收敛方向愈加的发散。

#### 四、对智能体决策影响因素的统计分析

##### (一) 整体回归分析

为了在统计学意义上考察  $\beta$ ,  $N$  和  $\theta$  对智能体

决策的影响,我们将共计 2 700 次的智能体训练结果作为统计样本,设定计量模型如下:

$$DV_i = \alpha_0 + \alpha_1\beta_i + \alpha_2N_i + \alpha_3\theta_i + \varepsilon_i \quad (16)$$

式(16)中,  $DV$  为被解释变量,分为四种设定,即研发投入合谋指数 ( $CI_{rd}$ ), 利润合谋指数 ( $CI_{pf}$ ), 是否实现了非合作均衡 ( $INE$ ) 以及是否实现了合作均衡 ( $ICE$ )。其中,  $INE$  和  $ICE$  为 0-1 变量,因此我们采用 logit 模型进行回归分析。解释变量即为研发外溢程度 ( $\beta$ ), 智能体数量 ( $N$ ) 和学习率 ( $\theta$ )。  $\varepsilon$  为随机扰动项。

表 5  $\beta, N$  和  $\theta$  对智能体决策的影响

| 变量       | (1) $CI_{rd}$             | (2) $CI_{pf}$             | (3) $INE$                | (4) $ICE$                |
|----------|---------------------------|---------------------------|--------------------------|--------------------------|
| $\beta$  | 0.005 0 ***<br>(0.000 6)  | 0.002 6 ***<br>(0.000 7)  | 0.902 6 ***<br>(0.008 8) | 1.0148<br>(0.010 9)      |
| $N$      | -0.051 0 ***<br>(0.006 4) | -0.052 8 ***<br>(0.007 2) | 1.610 6 ***<br>(0.147 6) | 0.280 9 ***<br>(0.040 3) |
| $\theta$ | -0.001 2 ***<br>(0.000 2) | -0.001 8 ***<br>(0.000 3) | 1.028 3 ***<br>(0.003 5) | 0.999 1<br>(0.004 3)     |
| $Cons$   | 0.540 7 ***<br>(0.029 9)  | 0.816 6 ***<br>(0.032 0)  | 0.111 6 ***<br>(0.042 3) | 1.125 6<br>(0.612 1)     |
| 样本数      | 2 700                     | 2 700                     | 2 700                    | 2 700                    |
| $R^2$    | 0.054 8                   | 0.041 2                   | 0.134 4                  | 0.097 5                  |

注:①括号里为估计系数的异方差稳健标准误;②\*, \*\*, \*\*\* 分别表示回归系数在 10%、5% 和 1% 水平上显著;③第(3)列和第(4)列各单元格中第一行数字报告的是几率比。

表 5 报告了计量模型( I )的估计结果。从第(1)列可以看出研发外溢度  $\beta$  对  $CI_{rd}$  具有显著的正向影响,  $\beta$  提高 0.01 导致  $CI_{rd}$  提高 0.005; 智能体数量  $N$  对  $CI_{rd}$  具有显著的负向影响, 智能体数量每增加 1 个将导致  $CI_{rd}$  下降 0.051; 学习率  $\theta$  对  $CI_{rd}$  产生显著的负向影响,  $\theta$  提高 0.01 将导致  $CI_{rd}$  下降 0.001 2。从第(2)列可以看出  $\beta$  对  $CI_{pf}$  具有显著的正向影响,  $\beta$  提高 0.01 导致  $CI_{pf}$  提高 0.002 6; 智能体数量  $N$  对  $CI_{pf}$  具有显著的负向影响, 智能体数量每增加 1 个将导致  $CI_{pf}$  下降 0.0528; 学习率  $\theta$  对  $CI_{pf}$  产生显著的负向影响,  $\theta$  提高 0.01 将导致  $CI_{pf}$  下降 0.001 8。第(3)列展示了被解释变量为  $INE$  的估计结果, 其中  $\beta$  的影响显著为负, 它每提高 0.01 将导致非合作均衡实现的概率下降 9.74%;  $N$  的影响显著为正, 每增加 1 个智能体将导致非合作均衡实现的概率提高 61.06%;  $\theta$  的影响显著为正, 它每提高 0.01 将导致非合作均衡实现的概率提高 2.83%。第(4)列展示了被解释变量为  $ICE$  的估计结果, 其中  $\beta$  的影响为正向但不显著; 智能体数量的影响显著为负, 每增加 1 个智能体将导致非合作均衡实现的概率下降 71.91%; 学习率的影响为负向, 但并不显著。

(二) 依参数取值分组后智能体决策结果与各参数的相关性分析

此部分, 我们在不同的参数取值组合下对各参数与智能体决策结果进行逐一的相关性分析, 这能够让我们看到各参数发挥作用的条件区间, 从而帮助我们更全面地考察各参数对智能体决策结果的协同影响。

1. 智能体决策结果与  $\beta$  的相关性

表 6 展示了在  $N$  和  $\theta$  的不同组合下, 智能体

决策结果与  $\beta$  的相关性分析。首先,  $CI_{rd}$  除在  $N=2$  且  $\theta=0.25$  时与  $\beta$  负相关外, 在其他的  $N$  和  $\theta$  组合下均与  $\beta$  正相关, 但仅在  $N=4$  时, 相关关系显著。第二,  $CI_{pf}$  与  $\beta$  在  $N=2$  和  $N=3$  时负相关, 且在  $N=2$  且  $\theta=0.25$ ,  $N=3$  且  $\theta=0.25$  以及  $N=3$  且  $\theta=0.50$  三种情况下负相关关系显著。但这种负相关关系在  $N=4$  且  $\theta=0.50$  和  $N=4$  且  $\theta=0.75$  两种情况下转为正相关, 在  $N=4$  且  $\theta=0.75$  时正相关关系显著。第三,  $INE$  与  $\beta$  在  $N=3$  且  $\theta=0.25$  时正相关, 在其他情况下均为负相关, 在  $N=3$  且  $\theta=0.75$  和  $N=4$  时负相关关系显著。第四,  $ICE$  与  $\beta$  在  $N=2$  且  $\theta=0.50$  时为负相关, 在  $N=4$  且  $\theta=0.75$  相关系数为 0, 在其他情况下均为正相关, 在  $N=4$  且  $\theta=0.25$  和  $N=4$  且  $\theta=0.50$  两种情况下相关关系显著。

表 6 智能体决策结果与  $\beta$  的相关性

|                    | $CI_{rd}$ 与 $\beta$ | $CI_{pf}$ 与 $\beta$ | $INE$ 与 $\beta$ | $ICE$ 与 $\beta$ | 样本数 |
|--------------------|---------------------|---------------------|-----------------|-----------------|-----|
| $n=2, \theta=0.25$ | -0.003 4            | -0.084 4            | -0.039 7        | 0.026 5         | 100 |
| $n=2, \theta=0.50$ | 0.028 6             | -0.043 9            | -0.078 3        | -0.051 5        | 100 |
| $n=2, \theta=0.75$ | 0.007 1             | -0.097 4 *          | -0.013 8        | 0.024 0         | 100 |
| $n=3, \theta=0.25$ | 0.014 7             | -0.172 2 ***        | 0.047 9         | 0.037 5         | 100 |
| $n=3, \theta=0.50$ | 0.002 9             | -0.119 8 **         | -0.038 7        | 0.050 7         | 100 |
| $n=3, \theta=0.75$ | 0.072 8             | -0.042 5            | -0.195 6 ***    | 0.025 3         | 100 |
| $n=4, \theta=0.25$ | 0.163 1 ***         | -0.016 1            | -0.180 5 ***    | 0.142 4 **      | 100 |
| $n=4, \theta=0.50$ | 0.403 1 ***         | 0.344 2             | -0.468 8 ***    | 0.127 6 **      | 100 |
| $n=4, \theta=0.75$ | 0.657 0 ***         | 0.657 5 ***         | -0.740 1 ***    | 0.000 0         | 100 |

注: \*, \*\*, \*\*\* 分别表示相关系数在 10%、5% 和 1% 水平上显著。下同。

2. 智能体决策结果与  $N$  的相关性

表 7 展示了在  $\beta$  和  $\theta$  的不同组合下, 智能体决策结果与  $N$  的相关性分析。首先,  $CI_{rd}$  与  $N$  在  $\beta=0.3$  且  $\theta=0.50$ ,  $\beta=0.4$  且  $\theta=0.25$  和  $\beta=0.4$  且  $\theta=0.50$  等三种情况下正相关, 在其他情况下负相关, 在  $\beta=0.2$  且  $\theta=0.25$ ,  $\beta=0.2$  且  $\theta=0.50$ ,  $\beta=0.2$  且  $\theta=0.75$  和  $\beta=0.3$  且  $\theta=0.75$  等四种情况下负相关关系显著。第二,  $CI_{pf}$  与  $N$  在  $\beta=0.3$  且  $\theta=0.25$ ,  $\beta=0.3$  且  $\theta=0.50$ , 以及  $\beta=0.4$  时正相关, 在其他情况下负相关, 在  $\beta=0.2$  且  $\theta=0.50$ ,  $\beta=0.2$  且  $\theta=0.75$  和  $\beta=0.3$  且  $\theta=0.75$  时负相关关系显著。第三,  $INE$  与  $N$  在  $\beta=0.2$  且  $\theta=0.50$ ,  $\beta=0.2$  且  $\theta=0.75$ ,  $\beta=0.3$  且  $\theta=0.75$  等三种情况下具有显著的正相关关系, 在其他情况下负相关, 在  $\beta=0.3$  且  $\theta=0.25$ ,  $\beta=0.3$  且  $\theta=0.50$  以及  $\beta=0.4$  时负相关关系显著。第四,  $ICE$  在九种情况下均与  $N$  成负相关关系, 但在  $\beta=0.4$  且  $\theta=$

0.50 时不显著。整体上看,随着智能体数量的增加,研发合谋的倾向会下降,企业的研发投入会增加。

表7 智能体决策结果与  $n$  的相关性

|                          | $CI_{rd}$ 与 $n$ | $CI_{pf}$ 与 $n$ | INE 与 $n$    | ICE 与 $n$    | 样本数 |
|--------------------------|-----------------|-----------------|--------------|--------------|-----|
| $\beta=0.2, \theta=0.25$ | -0.137 7 **     | -0.056 2        | -0.030 1     | -0.232 3 *** | 100 |
| $\beta=0.2, \theta=0.50$ | -0.312 2 ***    | -0.322 4 ***    | 0.304 3 ***  | -0.236 2 *** | 100 |
| $\beta=0.2, \theta=0.75$ | -0.617 7 ***    | -0.643 6 ***    | 0.670 7 ***  | -0.251 6 *** | 100 |
| $\beta=0.3, \theta=0.25$ | -0.026 2        | 0.038 9         | -0.163 5 *** | -0.112 4 *   | 100 |
| $\beta=0.3, \theta=0.50$ | 0.008 3         | 0.078 1         | -0.217 9 *** | -0.220 6 *** | 100 |
| $\beta=0.3, \theta=0.75$ | -0.153 2 ***    | -0.146 6 **     | 0.149 9 ***  | -0.189 1 *** | 100 |
| $\beta=0.4, \theta=0.25$ | 0.000 0         | 0.022 4         | -0.154 8 *** | -0.156 6 *** | 100 |
| $\beta=0.4, \theta=0.50$ | 0.018 7         | 0.019 2         | -0.104 2 *   | -0.086 0     | 100 |
| $\beta=0.4, \theta=0.75$ | -0.023 1        | 0.027 3         | -0.135 5 **  | -0.264 9 *** | 100 |

### 3. 智能体决策结果与 $\theta$ 的相关性

表8展示了在 $\beta$ 和 $N$ 的不同组合下,智能体决策结果与 $\theta$ 的相关性分析。首先, $CI_{rd}$ 与 $\theta$ 在 $N=2$ 和 $\beta=0.4$ 且 $N=4$ 时为正相关但不显著,在其他情况下为负相关,在 $\beta=0.2$ 且 $N=3$ , $\beta=0.2$ 且 $N=4$ 和 $\beta=0.3$ 且 $N=4$ 等三种情况下负相关关系显著。第二, $CI_{pf}$ 与 $\theta$ 在 $\beta=0.2$ 且 $N=2$ , $\beta=0.4$ 且 $N=2$ 和 $\beta=0.4$ 且 $N=4$ 等三种情况下为正相关但不显著,在其他情况下为负相关,在 $\beta=0.2$ 且 $N=3$ , $\beta=0.2$ 且 $N=4$ 和 $\beta=0.3$ 且 $N=4$ 等三种情况下负相关关系显著。第三,NE实现次数与 $\theta$ 在 $\beta=0.2$ 且 $N=3$ , $\beta=0.2$ 且 $N=4$ 和 $\beta=0.3$ 且 $N=4$ 等三种情况下显著正相关,在 $\beta=0.3$ 且 $N=2$ , $\beta=0.4$ 且 $N=3$ 和 $\beta=0.4$ 且 $N=4$ 三种情况下正相关但不显著,在其他情况下负相关但不显著。第四,CE实现次数与 $\theta$ 在 $\beta=0.4$ 且 $N=4$ 时显著负相关,在其他情况下的相关性均不显著。

表8 智能体决策结果与  $\theta$  的相关性

|                  | $CI_{rd}$ 与 $\theta$ | $CI_{pf}$ 与 $\theta$ | INE 与 $\theta$ | ICE 与 $\theta$ | 样本数 |
|------------------|----------------------|----------------------|----------------|----------------|-----|
| $\beta=0.2, n=2$ | 0.017 2              | 0.035 4              | -0.039 1       | 0.012 3        | 100 |
| $\beta=0.2, n=3$ | -0.108 8 *           | -0.144 0 **          | 0.228 2 ***    | -0.029 2       | 100 |
| $\beta=0.2, n=4$ | -0.566 9 ***         | -0.610 3 ***         | 0.646 2 ***    | 0.000 0        | 100 |
| $\beta=0.3, n=2$ | 0.010 6              | -0.021 7             | 0.012 3        | 0.081 6        | 100 |
| $\beta=0.3, n=3$ | -0.047 0             | -0.016 7             | -0.058 3       | -0.077 4       | 100 |
| $\beta=0.3, n=4$ | -0.141 9 **          | -0.221 0 ***         | 0.308 7 ***    | 0.050 2        | 100 |
| $\beta=0.4, n=2$ | 0.029 1              | 0.027 0              | -0.015 3       | 0.012 3        | 100 |
| $\beta=0.4, n=3$ | -0.052 9             | -0.049 5             | 0.019 4        | -0.043 4       | 100 |
| $\beta=0.4, n=4$ | 0.012 5              | 0.042 4              | 0.000 0        | -0.101 4 *     | 100 |

## 五、总结与政策建议

本文运用强化学习Q-learning算法仿真分析了人工智能决策下的企业研发竞争问题,基于研究结果可以得到以下结论:第一,尽管算法决策结果向“合谋”收敛的概率并不高,但与纳什均衡结果相比,研发投入更低而利润水平更高的情况是

普遍出现的;第二,智能体数量的增加能够降低“合谋”实现的概率,且这种影响具有较高的全局性;第三,研发的外溢程度会促进智能体“合谋”,但这种促进作用是局部的,在智能体数量较多时更易实现;第四,学习率的提高能够降低“合谋”实现的概率,但这种作用也仅在局部发生,在智能体数量较多且研发外溢性较低时作用更显著。

本文可能的贡献有以下几点:第一,由于企业研发行为存在外溢性,本文分析了外溢程度的变化对智能体决策产生的影响,这是对现有理论研究的一个重要补充;第二,本文深入挖掘了影响算法合谋的3个重要因素,以及各因素影响作用发挥的协同性,一定程度上揭示了算法“合谋”的内在逻辑性;第三,本文研究了运用机器学习算法进行研发投入决策的可能结果,阐明人工智能的无限制应用可能会对企业创新产生负向影响,证明了高科技成果的应用可能阻碍科技进步这一悖论的存在,这对数字经济时代如何进一步推动企业加大研发力度,提升研发效率具有较强的启示作用,对算法合谋的规制政策制定也具有较大的借鉴意义。

依据研究结论,本文提出政策建议。第一,机器学习算法决策在竞争强度不同的行业的应用应差异化推进。由于参与博弈的智能体数量能够显著地降低“合谋”发生的概率,因此,在竞争性较强的行业,可放宽算法决策的准入门槛。而对于规模经济性较强,企业数量较少,竞争性较弱的行业,对算法决策应更加审慎,从而有效避免具有极大隐蔽性而易于逃避反垄断法监管的智能体“合谋”的发生。第二,进一步加强知识产权保护,降低由于知识产权保护不足导致的过度外溢性。知识产权保护不足所导致的过度知识外溢会使企业减少研发投入,这种影响机制在算法决策中同样存在。本文的研究表明,在算法研发决策中,过度的知识外溢不仅引起研发投入下降,还促进了智能体“合谋”从而进一步降低研发投入。因此,应加强知识产权保护,避免知识过度外溢,减少研发竞争中算法“合谋”的发生。第三,对应用于企业决策的人工智能算法加强反垄断监管。本文研究结论发现,在智能体训练中的参数设置(如学习率

$\theta$ )也会影响算法“合谋”的发生概率。因此,应加强对应用于企业决策的人工智能算法的反垄断监管,对智能体的学习机制、参数设定等在不同的应用场景下进行系统研究,制定科学的算法监管办法。

#### 参考文献:

- [1] 孟昌,曲寒瑛. 算法合谋及其规制研究进展[J]. 经济动态, 2021(6):128-143.
- [2] 苏敏,夏杰长. 数字经济中竞争性垄断算法合谋的治理困境[J]. 财经问题研究, 2021(11):37-46.
- [3] 乔榛,刘瑞峰. 大数据算法的价格合谋问题研究[J]. 学术交流, 2022(1):82-95,191.
- [4] SCHWALBE U. Algorithms, machine learning, and collusion[J]. Journal of competition law & economics, 2019, 14(4):568-607.
- [5] GAUTIER A, ITTOO A, CLEYNENBREUGEL P V. AI algorithms, price discrimination and collusion: a technological, economic and legal perspective[J]. European journal of law and economics, 2020,50:405-435.
- [6] EZRACHI A, STUCKE M E. Sustainable and unchallenged algorithmic tacit collusion[J]. Northwestern journal of technology & intellectual property, 2020,17(2):217-260.
- [7] TURNER D. The definition of agreement under the Sherman Act: conscious parallelism and refusals to deal[J]. Harvard law review, 1962,75(4):655-706.
- [8] 沈亮亮. 算法在市场竞争者中的应用与法律难题:从大数据杀熟谈起[J]. 太原学院学报(社会科学版), 2019(3):26-31.
- [9] POSNER R A. Oligopoly and the antitrust laws: a suggested approach[J]. Stanford law review, 1969,21(6):1562-1606.
- [10] KAPLOW L. Competition policy and price fixing[M]. Princeton: Princeton university press, 2013:1-15.
- [11] 李振利,李毅. 论算法共谋的反垄断规制路径[J]. 学术交流, 2018(7):73-82.
- [12] 施春风. 定价算法在网络交易中的反垄断法律规制[J]. 河北法学, 2018(11):111-119.
- [13] VELJANOVSKI C. Pricing algorithms as collusive devices[EB/OL]. (2020-10-22) [2024-01-10]. [https://www.researchgate.net/publication/343438462\\_Pricing\\_Algorithms\\_as\\_Collusive\\_Devices](https://www.researchgate.net/publication/343438462_Pricing_Algorithms_as_Collusive_Devices).
- [14] CALVANO E, CALZOLARI G, DENICOLO V, et al. Algorithmic pricing what implications for competition policy[J]. Review of industrial organization, 2019, 55(1):155-171.
- [15] 唐要家,尹钰锋. 算法合谋的反垄断规制及工具创新研究[J]. 产经评论, 2020(3):5-16.
- [16] 尹振涛,陈媛先,徐建军. 平台经济的典型特征、垄断分析与反垄断监管[J]. 南开管理评论, 2022,25(3):213-226.
- [17] 刘学. 人工智能时代算法合谋反垄断的适法困境与逻辑依归[J]. 西北民族大学学报, 2022(1):87-94.
- [18] MILGROM P R, TADELIS S. How artificial intelligence and machine learning can impact market design[EB/OL]. (2018-02-20) [2024-01-10]. <https://www.nber.org/papers/w24282>.
- [19] ASSAD S, CLARK R, EESHOV D, et al. Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market[EB/OL]. (2020-08-31) [2023-12-25]. [https://www.econ.queensu.ca/sites/econ.queensu.ca/files/wpaper/qed\\_wp\\_1438.pdf](https://www.econ.queensu.ca/sites/econ.queensu.ca/files/wpaper/qed_wp_1438.pdf).
- [20] SALCEDO B. Pricing algorithms and tacit collusion[EB/OL]. (2015-10-15) [2023-12-20]. <https://api.semanticscholar.org/CorpusID:18195281>.
- [21] CALVANO E, CALZOLARI G, DENICOLÒ V, et al. Artificial intelligence, algorithmic pricing, and collusion[J]. American economic review, 2020,110(10):3267-97.
- [22] KLEIN T. Autonomous algorithmic collusion: Q-learning under sequential pricing[J]. Rand journal of economics, 2021,52(3):538-599.
- [23] WATKINS C J, DAYAN P. Q-learning[J]. Machine learning, 1992(8):279-292.
- [24] WALTMAN L, KAYMAK U. Q-Learning agents in a Cournot oligopoly model[J]. Journal of economic dynamics and control, 2008,32(10):3275-93.
- [25] UK Competition and Markets Authority. Pricing algorithms[EB/OL]. (2018-10-08) [2024-01-10]. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/746353/Algorithms\\_econ\\_report.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/746353/Algorithms_econ_report.pdf).
- [26] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015,518(7540):529-33.
- [27] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of go with deep neural networks and tree search[J]. Nature, 2016,529(7587):484-89.
- [28] SILVER D, HUBERT T, SCHRITTWIESER J, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play[J]. Science, 2018,362(6419):1140-44.

(本文责编:默黎)