

# Mechanism Designs of Boxed Pig Game and Experiment and Analysis Based on Q-Learning

Yicheng Gong<sup>1,2</sup>, Qing Liu<sup>1\*</sup>, Yanli Xu<sup>1</sup>, Yuqiang Feng<sup>1,2</sup>

<sup>1</sup>. School of Science, Wuhan University of Science and Technology, Wuhan 430065, China;

<sup>2</sup>. Hubei Province Key Laboratory of Systems Science in Metallurgical Process (Wuhan University of Science and Technology), Wuhan 430081, China

## Abstract

The equilibrium situation of the traditional boxed pig game is that the piglet "waits" while the big pig "presses the button" but gets a lower return. How to avoid the negative influence of this "unfairness" and improve the social efficiency has become a topic worth discussing. In order to improve the efficiency of social fairness, this paper designs three new mechanisms for the boxed pig game: (1) no physical cost, (2) buttons with the same side of the trough, (3) with storage containers. And then analyzes the equilibrium under different mechanisms theoretically. The classical Q-learning algorithm is used to simulate the players to do experiments. The results show that: Mechanism (1) can completely avoid piglets' "free riding"; Mechanism (2) cannot avoid pig lets' "free-riding"; Mechanism (3) the possibility of piglets' "free riding" is reduced by 70%. Mechanism (3) is a fairer and more efficient mechanism.

## Keywords

boxed pig game; free riding; mechanism design; reinforcement learning; Q-learning

## 1. Introduction

Free-riding exists widely in social life. But free-riding have many adverse effects. It leads to market failures; inadequate supply of public goods; reduce production efficiency; undermine cooperation and reduce the overall efficiency of society. How can the harm of "free riding" be reduced or even eliminated? How to achieve the desired effect? This is worth studying and discussing. Game theory is the method that players find the optimal response strategy, so we use the boxed pig game to analyze problems.

In the boxed pig game, due to the obvious power gap between the two players, the piglet always chooses to "wait" while the big pig chooses to "press the button". The equilibrium result is the piglet free riding, so the game is also called the free-riding game. It was first described by Nash J. in 1950<sup>[1]</sup>. Such a situation is unfair to the powerful "big pigs", which will affect their enthusiasm for action and reduce the efficiency of society in the long run.

The boxed pig game model is widely used to solve the practical problem of the "free riding" phenomenon. In the field of environmental governance, Yang Kun<sup>[2]</sup> (2019) studied the game relationship between polluting enterprises and local governments by using the boxed pig game model. Chen Guisheng<sup>[3]</sup> applied the boxed pig game to the game between local governments and civil society to find solutions to pollution problems and break the dilemma of the boxed pig game in 2019. In the field of enterprise management, Yang and Shen<sup>[4]</sup> (2016), Yu<sup>[5]</sup> (2019), and Huang Zena<sup>[6]</sup> (2020) discussed the boxed pig game between large enterprises and small and medium-sized enterprises, so as to find a way to stimulate the innovation of small and medium-sized enterprises. In the field of higher education, the boxed pig game has been applied to scientific research team management<sup>[7][8]</sup>, library management<sup>[9][10]</sup>, cooperative learning in universities<sup>[11]</sup>, talent incentive in universities<sup>[12]</sup> and so on. These scholars have explored ways to improve the inefficient equilibrium of piglet "free-riding" in different fields and games. But these studies focus on the construction and analysis of theoretical models and lack the test of game experiments.

---

AHPCAI2022@2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence

EMAIL: \*Corresponding author: liuqing@wust.edu.cn (Qing Liu)



© 2022 Copyright for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).  
CEUR Workshop Proceedings (CEUR-WS.org)

In reality, it is very difficult to conduct game experiments to test the effect of game theory. Noted that reinforcement learning, which has emerged in recent years, aims to train agents to learn through trial and error, and to seek the optimal coping strategy. Therefore, this paper tries to use reinforcement learning agents to do game experiments. It is reasonable for three reasons: first, the agent's goal is to maximize its expected return as set by the programmer, so the agent has no desire to hide his real preference. Second, the game experiment on agents only needs to debug learning strategies and procedures, which is much cheaper than the experiment on real people. Third, train agents to do game experiments where data can be collected easily. Therefore, this paper uses reinforcement learning method to simulate players to do game experiments, in order to test the practical effect of theoretical analysis of game mechanisms.

In the fast-paced modern era, people increasingly pursue efficiency, how to avoid the phenomenon of "free riding" and improve efficiency is widespread concern of people, but there is no attention to the reinforcement learning of free riding games. Aiming how to change the inefficient equilibrium of "free-riding", this paper designs three kinds of boxed pig game mechanisms, and uses the agent in multi-agent reinforcement learning to simulate the players in the boxed pig game to do game experiments, so as to explore the experimental equilibrium and effect of the boxed pig game model under different mechanisms.

## 2 Introduction to Basic Knowledge

### 2.1 Traditional boxed Pig game

Suppose there are two pigs in the pen: a big pig and a little pig. On one side of the pen there is a trough, and on the other side, there is a button that controls the supply of pig food. When pressed, 10 units of pig food will be placed in the trough. The big pig and the piglet have two strategies: press the button and wait. Pressing the button costs 2 units of stamina. If only one party chooses to press the button, it loses the chance to get to the slot first. There are differences in the ability of the large pig and the small pig, and the ability of the big one to eat is stronger than that of the small one. If the piglet came to the trough first to eat, the ratio of the food eaten by the big pig to the piglet was 6:4. If they eat at the trough at the same time, the benefit ratio of the big pig and the piglet is 7:3. If the big pig comes to the trough first, the payoff ratio between the big pig and the piglet is 9:1<sup>[13]</sup>. The food intake of the big pig and the piglet is regarded as the benefit, and the physical energy consumed is regarded as the cost. The income matrix of the boxed pig game can be obtained, as shown in Table 1:

**Table 1** The return matrix of the traditional boxed pig game

		Little pig	
		Pressing the button	Waiting
Big pig	Pressing the button	(5, 1)	(4, 4)
	Waiting	(9, -1)	(0, 0)

In Table 1, the first component of each payoff vector is the payoff of the large pig, and the second component is the payoff of the piglet. According to the payoff matrix, "waiting" is the dominant strategy of the piglet. When the piglet choose "waiting", the optimal response strategy of the big pigs is "pressing the button", and the Nash equilibrium of the game is (the big pig presses the button, the piglet waits).

A repeated game is a dynamic game consisting of a stage game repeated many times. It can be divided into finite repeated game and infinite repeated game. When the stage game has only a unique Nash equilibrium, the equilibrium result of the stage game will not be changed by the finite number of repeated game<sup>[14]</sup>. The subgame refined Nash equilibrium, which is completely different from the stage game, is produced when the game is repeated an infinite number of times. If the above-boxed pig game is regarded as a stage game, the equilibrium situation of the boxed pig game with finite repeats is that the big pig keeps pressing the button and the piglet keeps waiting. Infinite repetitions of the boxed pig game will lead to different equilibria.

## 2.2 Mechanism Design

Mechanism design is to achieve a specific equilibrium by designing game mechanisms. The designed mechanism needs to satisfy two constraints: individual rationality constraint and incentive compatibility constraint. The individual rationality constraint means that the expected return obtained by the agent under the designed mechanism is not lower than the return obtained by the agent if he does not accept the mechanism. The incentive compatibility constraint means that the designer of the agent selection mechanism expects the payoff of the action he chooses to be no less than the expected payoff of other actions he chooses. According to the display principle, the configuration result achieved by any mechanism can be achieved by a direct mechanism. Therefore, direct mechanism design is often carried out during mechanism design.

## 2.3 Q-learning algorithm

Q-learning algorithm is a model-free reinforcement learning algorithm based on value function<sup>[15]</sup>. Using the time difference method, we can learn the off-line strategy, and use the Behrman equation to solve the optimal strategy of Markov process.  $Q(s, a)$  state action value function, used to evaluate the expected revenue of action A in a certain state S. The environment rewards R immediately based on the action. The Q-learning algorithm will build a Q-value table to store the Q-value of all state action pairs, and then select the action with the maximum profit according to the Q-value. Game theory always relies on the assumption of rational man, which is divorced from the reality, while Q-learning algorithm can better analyze the game problems in reality through continuous trial and error learning.

## 3. Three New Mechanisms of the Boxed Pig Game

According to Section 2.1, the equilibrium of the traditional boxed pig game is (big pig presses the button, little pig waits), that is, due to the natural strength gap of big pig and little pig and the established game background, little pig gets free-riding, and big pig has low income. In real life, this will cause the pig to have negative thoughts over time, which will reduce the overall income. From the social point of view, how to improve the efficiency and fairness of the boxed pig game has become an important topic. In order to try to improve the single stable situation of "free-riding", this paper designs three direct game mechanisms : (1) no physical cost; (2) with storage container; (3) button and trough on the same side. In the boxed pig game, the big pig and the piglet have no other food sources and will not gain anything by not participating in the game, so participating in the game satisfies the individual rational constraint. The incentive compatibility constraint should be analyzed in a specific mechanism.

### 3.1 Game mechanism with no physical cost

The assumption is that the act of "pressing the button" does not require physical exertion, meaning that big pig and little pig do not have to pay additional costs to obtain food. In this case, the payoff matrix of the boxed pig game without physical cost is shown in Table 2:

**Table 2** The return matrix of boxed pig game with no physical cost

		Little pig	
		Pressing the button	Waiting
Big pig	Pressing the button	(7, 3)	(6, 4)
	Waiting	(9, 1)	(0, 0)

The expected income of "pressing the button" and "waiting" of big pig is 6.5 and 4.5 respectively. The expected payoff of "pressing the button" is no less than that of "waiting". Therefore, this mechanism satisfies the constraint that motivates pigs to "pressing the button".

Thus, the boxed pig game with no physical cost has two pure Nash equilibria :(big pig presses the button, little pig waits) and (big pig waits, little pig presses the button). A mixed strategy Nash equilibrium: big pig choose "pressing the button" with probability 0.5 and "waiting" with probability 0.5; The piglet choose "pressing the button" with a probability of 0.75 and "waiting" with a probability of 0.25.

### 3.2 Game mechanism with storage container

Assuming that the big pig and little pig have separate food storage containers, the pig that reaches the trough first can store a portion of the food (assuming  $C$  units of food,  $C \geq 1$ ). In this case, if the piglet reach the trough first, the payoff ratio between the big pig and the piglet is  $6 - C : 4 + C$ ; If the big pig reaches the trough first, the big pig will monopolise all the food ( $9 + C \geq 10$ ); If both big pig and little pig arrive at the trough at the same time, the pay-off ratio is still 7:3. In this case, the payoff matrix of boxed pig game with storage container is shown in Table 3:

**Table 3** The return matrix of boxed pig game with storage containers

		Little pig	
		Pressing the button	Waiting
Big pig	Pressing the button	(5, 1)	(4-c, 4+c)
	Waiting	(10, -2)	(0, 0)

At this point, the expected benefit of "pressing the button" for the big pig is  $4.5 - 0.5C$ , and the expected benefit of "pressing the button" for the piglet is 0.5, which is lower than the expected benefit of "waiting" ( $2 + 0.5C$ ). This mechanism cannot effectively motivate piglets to "pressing the button". According to the analysis, the Nash equilibrium of the boxed pig game with storage containers is (the big pig presses the button, the piglet waits).

### 3.3 Game mechanism with button and trough on the same side

If the button is on the same side as the trough and the two pigs are resting on the other side, the pig that presses the button will reach the trough first, and then the other pig will be attracted to the food. It costs 2 units to "pressing the button". If the piglet "pressing the button" and gets to the trough first, the payoff for the pigs are 6:2. If the big pig "pressing the button", the pigs payoff ratio are 7:1; If both pigs arrive at the trough at the same time, the payoff ratio is 5:1. In this case, the payoff matrix of the boxed pig game with button and trough on the same side is shown in Table 4:

**Table 4** The return matrix of boxed pig game with the same side of the button and the trough

		Little pig	
		Pressing the button	Waiting
Big pig	Pressing the button	(5, 1)	(7, 1)
	Waiting	(6, 2)	(0, 0)

For the big pig, the expected income of "pressing the button" is 6, and the expected income of "waiting" is 3. The expected payoff of "pressing the button" is 1.5 higher than that of "waiting" is 0.5, so this mechanism satisfies the constraint that motivates pigs to "pressing the button".

According to the game theory analysis, there are two Nash equilibria in the boxed pig game with the button and the same side of the trough: (the big pig presses the button, the piglet waits), (the big pig waits, the piglet presses the button), and a mixed strategy Nash equilibrium: the big pig will choose "pressing the button" with probability 1; The piglet chose "pressing the button" with a probability of 0.875 and "waiting" with a probability of 0.125.

In the traditional boxed pig game, the piglet has been "free riding", while the big pig can only "pressing the button" because of its rationality. However, in real life, players are often bounded

rationality, players out of "envy", "exhaustion", "threat" and other reasons, have a negative and lazy idea, so as to make irrational decisions, resulting in the overall income reduction. Therefore, whether the game practice is consistent with the theoretical analysis and how effective the new game mechanism designed to improve the free-riding phenomenon remain to be tested. In this paper, the classical Q-learning algorithm of reinforcement learning is used to simulate the game experiment of real players.

## 4. Game Simulation Experiment of Boxed Pig Game Based on Q-Learning

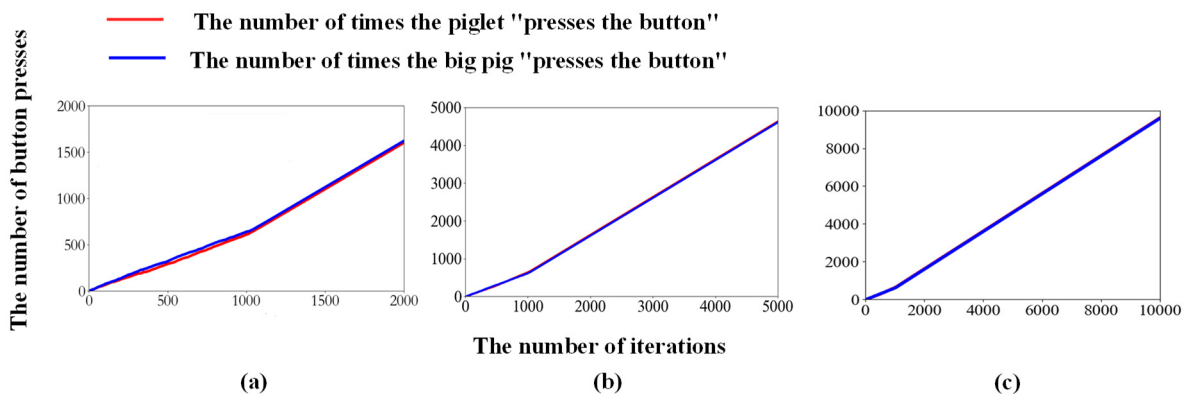
### 4.1 Simulation experiment design of boxed pig game based on Q-learning

In this paper, each boxed pig game was simulated for 2000, 5000 and 10000 rounds. In the simulation experiment of Q-learning, players cannot determine when the game will end, so it can be regarded as an infinite number of repeated games. The three game mechanisms designed in this paper are simulated respectively, and the changing rules of agent strategy and the avoidance effect of the three mechanisms on "free-riding" behavior are analyzed.

### 4.2 Boxed pig game experiment with no physical cost

Under the design of boxed pig game mechanism without physical cost, 2000 rounds, 5000 rounds, and 10000 rounds of game experiments were carried out respectively. The changing rules of the number of "pressing the button" for the big pig and the piglet are shown in Figure 1 (a), (b) and (c) respectively.

According to Figure 1(a), after 2000 rounds of simulation experiments, the big pig and the piglet tend to "press the button" at the same time in the later period after the first 1000 rounds of exploration, and the frequency of "pressing the button" is



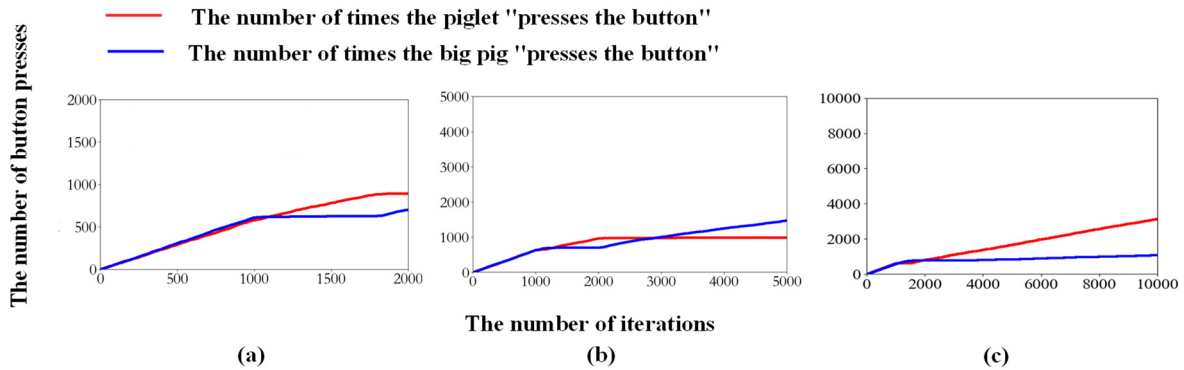
**Figure 1.** The game experiment of the boxed pig with no physical cost

more than 80%. See Figure 1(b), after 5000 rounds of simulation experiments, the strategy of the big pig and the piglet stabilized at "pressing the button" after 1000 rounds, and the frequency of "pressing the button" remained above 90%. Similar to the 2000 simulation, the big pig and the piglet tended to "press the button" after 1000 rounds of learning. See Figure 1(c), after 10000 rounds of simulation experiments, the big pig and the piglet will be stable at "pressing the button" after 1000 times of training, and the frequency of "pressing the button" will remain above 96% in the end.

Under the mechanism of boxed pig game without physical cost, the results of 2000, 5000, and 10000 rounds of simulation experiments are basically the same: the big pig and the piglet learn their optimal coping strategy through 1000 rounds of training, that is, "pressing the button". After 1000 rounds, players "press the button" nearly 100% of the time, with the situation (big pig press the button, little pig press the button). This mechanism avoids piglet free-riding altogether. Big pigs and little pigs don't get a higher payoff by "pressing the button" at the same time, they just keep going back and forth between the trough and the button.

### 4.3 Boxed pig game experiment with storage container

Under the game mechanism design of boxed pig game with storage containers, 2000 rounds, 5000 rounds, and 10000 rounds of game experiments were carried out respectively. The changing rules of the number of "pressing the button" for the big pig and the piglet are shown in Figure 2 (a), (b), and (c), respectively.



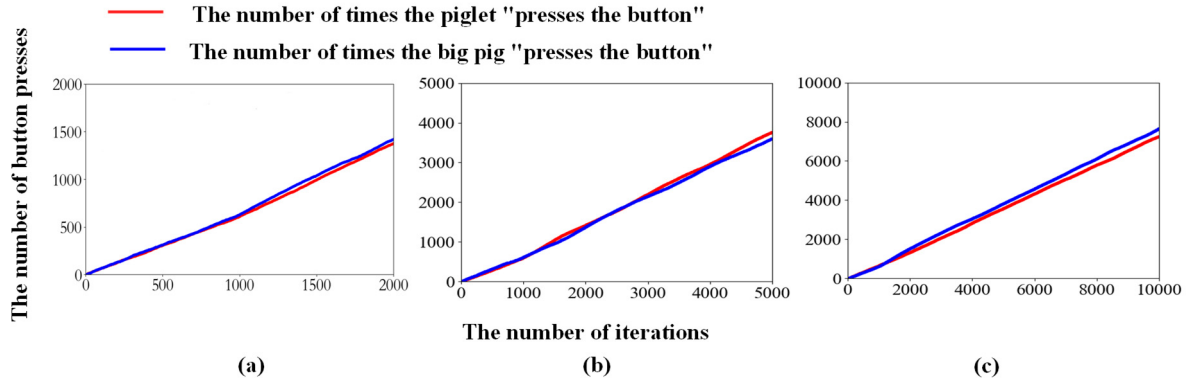
**Figure 2.** Game experiment of boxed pig with storage container

According to Figure 2(a), after 2000 rounds of simulation experiments, the frequency of "pressing the button" is about 60% for the big pig and the piglet in the first 1000 rounds of random exploration strategy. After 1000 rounds, the big pig tried the pure strategy of "waiting", and the piglets played the mixed strategy, and the frequency of "pressing the button" decreased to 53%. After 1800 rounds, the pigs tried the pure strategy of "waiting" and the big pigs tried the mixed strategy, reducing the frequency of "pressing the button" to 36%. It can be seen from Figure 2(b) that after 5000 rounds of simulation experiments, the big pig and the piglet constantly adjusted their strategies in 5000 rounds of simulation. In the initial random exploration strategy, the frequency of "pressing the button" was about 70%. When the pigs then tried the pure "waiting" strategy, the bigger pig used the mixed strategy, and the frequency of "pressing the button" decreased to 42 percent. When the big pig tried the pure "waiting" strategy, the piglet played the mixed strategy, and the frequency of "pressing the button" decreased to 28%. It can be seen from Figure 2(c) that 10000 rounds of simulation experiments were conducted, which showed that the strategy of the big pig and the piglet finally stabilized (the big pig pressed the button and the piglet waited) by constantly adjusting the strategy.

Under the intelligent pig game mechanism with storage containers, it is not difficult to find that the stability of the piglet strategy of the big pig is significantly delayed after 2000, 5000, and 10000 rounds of experiments, but it will eventually be stable until the piglet "waiting" for the big pig to "pressing the button". The effect of this mechanism on improving piglet "free riding" is not obvious.

### 4.4 Boxed pig game experiment with button and trough on the same side

Under the game mechanism design of boxed pig game with the same edge of the button and the trough, 2000 rounds, 5000 rounds, and 10000 rounds of game experiments were conducted respectively. The changing rules of "pressing the button" for big pigs and piglets were shown in Figure. 3 respectively.



**Figure 3.** Boxed pig game experiment with button and trough on the same side

According to Figure 3(a), after 2000 rounds of simulation experiments, the frequency of “pressing the button” of the big pig and the piglet can reach 70%. This suggests that the big pig adjust their strategy through the first 1000 rounds of learning, and later are more likely to "press the button" at the same time. According to Figure 3(b), after 5000 rounds of simulation experiments, the frequency of “pressing the button” of the big pig and piglets can reach 70%. As you can see, the big pigs adjusted their strategies during the simulation, but were still actively "pressing the button". It can be seen from Figure 3(c) that after 10,000 simulation experiments, the big pig and the piglet tend to "press the button" to increase their profits in the whole simulation process.

In the game mechanism of boxed pig game with the same edge of the button and the trough, the results of 2000, 5000, and 10000 rounds of simulation experiments are basically the same: the number of the big pig and the piglet choosing to "press the button" increases evenly, reaching a higher level, and neither of them will continue to "wait". After strategy learning, the mechanism can avoid long-term "free riding" and encourage pigs to "press buttons". It also gives the piglet the chance to "free ride" in order to get a bigger profit. Bounded rational players adjust their strategies only according to their expected returns so that they can achieve a better overall level.

## 5 Conclusion

Under the background of the traditional mechanism game, the equilibrium result is that the piglet chooses to "free ride", which leads to the lower income of the big pig. In view of this unfairness, this paper designs three new game mechanisms for the traditional boxed pig game: the boxed pig game without physical cost, the boxed pig game with storage container, and the boxed pig game with button and trough on the same side. The equilibrium under the three mechanisms is analyzed theoretically. Considering that the analysis of game theory relies on the assumption of complete rationality, while in real life, players are often bounded rationality, this paper chooses reinforcement learning simulation game experiments that can be learned through trial and error. Game experiments based on Q-learning algorithm show that:

(1) The game mechanism without physical cost completely avoids piglet "free riding", and the optimal coping strategy of big pigs and piglets is "press the button". The big pig and the piglet go back and forth between the button and the trough at the same time, which will not generate higher returns, and will form an inefficient equilibrium.

(2) The game mechanism with storage containers can not avoid piglet "free-riding", and the stability of the strategy of the big pig and the piglet is significantly delayed. A container to store food in does not bridge the gap in power, and "waiting" is still a better strategy for the pigs.

(3) The button and trough on the same side of the mechanism reduced the frequency of "free-riding" by at least 70%, and the frequency of “pressing the button” was more than 70% in both big pig and piglet. This mechanism not only avoids long-term "free riding", but also makes pigs actively "press the button". And give the opportunity, so that pigs get more income. In real life, if you can make the distance between the button and the trough closer, and make the delayed return of the effort more timely, you will get a fairer and more efficient situation.

## 6 Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (72031009) and Hubei Province Key Laboratory of Systems Science in Metallurgical Process(Y202105)

## 7 Reference

- [1] Nash J F. Equilibrium Points in N-Person Games[J]. Proceedings of the National Academy of Sciences, 1950, 36(1):48-49.
- [2] Yang Kun. Influence of carbon trading price on haze control from the perspective of game theory [J]. Science and Technology Innovation Review, 2019, 16(08):244-247.
- [3] Chen G S. Local government and Civil society in the paradox of environmental governance: a smart pig game model [J]. Journal of Sichuan University (Philosophy and Social Sciences Edition),2019(02):85-93.
- [4] Yang T, Shen J. System analysis of innovation mechanism of small and medium-sized enterprises - Based on the perspective of game theory [J]. Technical Economics and Management Research,2016(01):50-53.
- [5] Yu M. Research on the innovation and R&D behavior of large enterprises and small and medium-sized enterprises: Based on the thinking of smart pig game model [J]. Chinese Business Theory,2019(13):147-148.
- [6] Huang Z N. The enlightenment of "smart pig Game" on stimulating innovation of small and medium-sized enterprises [J]. Journal of Guangdong polytechnic of light industry,2020,19(04):25-28.
- [7] Zhu X, Jiang D. Research on management of university scientific research team based on axiom system of intelligent pig game [J]. Economic Research Guide,2013(09):295-297.
- [8] Ouyang W M. Analysis on the Predicament of College Scientific Research Team Construction from the Perspective of Game Theory[J]. DEStech Transactions on Social Science, Education and Human Science, 2017.
- [9] Yi Zhou. Boxed Pig Game and Cooperation between Large and small libraries in information resource sharing [J]. Library and Information Service, 2006:131-133+138.
- [10] Feng Qin. Research on the Management Strategy of Subject Librarian System in Small and medium-sized University Libraries at the initial Stage -- Based on the analysis of "Smart Pig Game" model [J]. Library and Information Service, 2010, 54(9):4.
- [11] Ke Ren. Research on cooperative learning in colleges and universities from the perspective of game theory [J]. Cultural and Educational Materials,2014(32):139-140.
- [12] Sheng Li. Countermeasure analysis of university talent incentive mechanism based on the perspective of smart pig game [J]. Value Engineering,2015, 34(20):210-211.
- [13] Shin H S, Rasmusen E. Games and Information: An Introduction to Games Theory [J]. The Economic Journal, 1989, 99(397):864.
- [14] Fudenberg D, Maskin E. On the dispensability of public randomization in discounted repeated games[J]. Journal of Economic Theory, 1991.
- [15] Watkins C, Dayan P. Technical Note: Q-Learning[J]. Machine Learning, 1992, 8(3-4):279-292.