ISSN 2616-5767 Vol.8, Issue 3: 51-58, DOI: 10.25236/AJETS.2025.080308

Dynamic Supply Chain Risk Assessment and Strategy Collaborative Optimization Model Based on Markov Chain and Game Theory

Wei Guo[1,a], Shuijin Rong[1,b], Hao Liu[1,c,*]

[1]University of Science and Technology Liaoning, Anshan, China
[a]weiguo@stu.ustl.edu.cn, [b]120223803046@stu.ustl.edu.cn, [c]haoliu@ustl.edu.cn
**Corresponding author*

Abstract: *Aiming at the dynamic risk evolution problem caused by the strategic interaction of multiple stakeholders in complex systems, this paper proposes a hybrid risk assessment model integrating Markov chains and game theory. Traditional static models have limitations in depicting the nonlinear conduction mechanism of strategy interaction and the long-term risk trend. In this study, by embedding the game equilibrium strategy into the state transition process of the Markov chain, a dynamic closed-loop framework of "strategy selection - state evolution - benefit feedback" is constructed. The model incorporates the environmental risk level and the combination of participants' strategies into the state space together, and modifies the transition probability matrix in real time based on Nash equilibrium to achieve the dynamic mapping of micro-strategy interaction and macro-risk evolution. Numerical experiments in the supply chain scenario show that in a low/medium risk state, suppliers and purchasers can significantly increase long-term returns through the "expansion - additional purchase" strategy combination (the maximum increase in supplier returns is 46.4%). Under the high-risk state, the strategy synergy shifts to the "contraction - purchase reduction" combination, and the returns of both sides are optimized by 45.6% and 13.6% respectively, verifying the risk adaptive characteristics of the Nash equilibrium. The sensitivity analysis further revealed that the discount factor has a significant impact on the long-term value under the low-risk state (the increase of the value function reaches 28.7% when $\gamma=0.95$), providing a theoretical basis for the design of the dynamic discount mechanism. This model provides supply chain managers with a decision support tool that combines theoretical rigor and practical operability by coordinating strategic conflicts and optimizing risk exposure.*

Keywords: *Markov Chain, Game Theory, Risk Assessment, Dynamic Decision-Making*

1. Introduction

This study proposes a methodological framework integrating Markov chains and game theory to analyze the risk assessment problem under the strategic interaction of multiple stakeholders. In the risk assessment of complex systems, the strategic interaction among stakeholders has dynamic evolution characteristics, and traditional static models are difficult to effectively describe its nonlinear transmission mechanism [1]. Take the financial market as an example. Investors' decisions are not only influenced by the strategies of their rivals but also have a counter-effect on the evolution of the market state [2]. The transmission of supply chain risks also presents the networked characteristics of multi-agent strategy interaction, and there is an urgent need to construct a dynamic modeling method.

The Markov chain describes the temporal evolution law of the system through the state transition matrix. However, its traditional application has two limitations: Firstly, the setting of the state transition probability ignores the policy initiative of the decision-making subject [3]; Secondly, the assumption of no aftereffect is difficult to reflect the driving effect of the game process on the evolution of the state [4]. Although game theory can analyze the interaction of multi-agent strategies, it focuses on the analysis of static equilibrium and lacks the description of the dynamic evolution process [5]. The collaborative modeling of the two can achieve complementary advantages: Game theory quantifies the immediate impact of strategy interaction on the system, while Markov chain characterizes the long-term risk trend driven by strategy.

The existing research mainly has three deficiencies: (1) The feedback mechanism of strategy adjustment and state evolution has not been fully modeled [6]; (2) The correlation analysis between micro behavioral logic and macro risk trends is insufficient [7]; (3) The research on the coupling mechanism between multi-stage dynamic games and nonlinear risk accumulation is weak [8]. To this end, this paper constructs a Markov - game hybrid model. The innovation point lies in: taking the game equilibrium strategy as the driving variable for the state transition of the Markov chain, forming a dynamic closed loop of "strategy selection - state evolution - benefit feedback". It is specifically achieved through a two-stage mechanism: In the state definition stage, the combination of participant strategies is incorporated into the Markov chain state space; In the dynamic update stage, the transition probability matrix is corrected in real time based on the game equilibrium solution, thereby establishing the dynamic mapping relationship between the interaction of micro strategies and the evolution of macro risks.

2. Markov Chain Model

A Markov Chain is a mathematical model that describes Stochastic processes. Its core characteristic is "lack-of-memory property", that is, the state of the system at a certain moment only depends on its state at the previous moment and has nothing to do with earlier history [9].

2.1 Define the State Space

The state space is the core framework for describing the dynamic risk evolution of a system. In this model, the state space is composed of the environmental risk level and the combination of participants' strategies, and is formally defined as:

$$S = \{s_i | \ s_i = (e_k, \{a_1, a_2, \ldots, a_N\})\} \tag{1}$$

Among them: e_k represents the environmental risk level($e_1 = low\ risk, e_2 = medium\ risk, e_3 = high\ risk$), $\{a_1, a_2, \ldots a_N\}$ represents the combination of strategies chosen by the participants in the current state.

To enhance the predictive efficiency of environmental risk modeling, the methods of risk discretization and strategy integration can be adopted. Firstly, the environmental risks are discretized into three levels: low, medium and high, reducing the model complexity while retaining the core risk characteristics. Secondly, the strategy choices of the participants are integrated into the state space to solve the problem of prediction bias caused by the neglect of strategy interaction in traditional models.

2.2 Transition Probability Matrix

The transition probability matrix is the core tool for dynamic modeling of Markov chains, and its element p_{ij} represents the probability of transitioning from state s_i to s_j. Different from the traditional model, in this model, the transition probability is driven by the combination of participants' strategies and is dynamically adjusted through game equilibrium. Its mathematical form is:

$$p_{ij} = P(s_j | \ s_i, \pi_1^*(s_i), \pi_2^*(s_i)) \tag{2}$$

In environmental risk modeling, the probability of state transition is determined by the strategy choices of the participants. For example, in a low-risk state, when suppliers and purchasers respectively adopt the "expansion" and "additional purchase" strategies, the probability that the system maintains the current risk level is significantly higher than that of other strategy combinations. This study constructs a dynamic update mechanism and adjusts the transition probability matrix in real time by solving the game equilibrium. This mechanism is implemented in two stages: Firstly, the Nash equilibrium strategy is solved based on the payment matrix to determine the optimal strategy combination; Subsequently, based on the environmental feedback generated by the implementation of the strategy, the probability distribution of state transition is updated, thereby dynamically representing the risk evolution process caused by the interaction of the strategy.

2.3 Steady-State Analysis

Steady-state distribution refers to the situation where a distribution vector exists:

$$\pi = \left[\pi_1, \pi_2, \ldots, \pi_N\right], \pi_i \geqslant 0, \sum_{i=1}^{\infty} \pi_i = 1 \tag{3}$$

ISSN 2616-5767 Vol.8, Issue 3: 51-58, DOI: 10.25236/AJETS.2025.080308

Let the above $\pi P = \pi$. Steady-state analysis is a method for studying the equilibrium state achieved by a system during its long-term operation. In this state, the key characteristics of the system no longer change significantly over time, presenting stable rules or fixed patterns. In the dynamic evolution mechanism, the state transition probability is directly determined by the strategy combination of the participants and is corrected in real time through the solution of game equilibrium.

3. Game Theory

Game theory is a mathematical theory that studies how multiple decision-making subjects (participants) make the optimal decisions in strategic interactions [10]. The core lies in analyzing how the choices of participants influence each other in competitive or cooperative situations and predicting possible equilibrium outcomes.

3.1 Define the Participants

The model contains two types of typical participants, and their strategy choices directly affect the risk evolution path:

Supplier:

$$\mathcal{A}_1 = \{\text{Expand}, \text{Shrink}\} \tag{4}$$

Buyer:

$$\mathcal{A}_2 = \{\text{Increase}, \text{Decrease}\} \tag{5}$$

The strategy choices of the participants form a strategy linkage effect through the payment matrix. Take the low-risk state as an example. When suppliers adopt the "expansion" strategy and purchasers simultaneously choose "increase procurement", this cooperative strategy combination generates higher returns through the synergy of supply and demand. Conversely, if one party adopts an aggressive expansion strategy while the other maintains a conservative contraction, it may cause a conflict in returns due to strategy mismatch, reflecting the direct correlation between strategy choice and return structure.

3.2 Strategy Space

The strategy space is dynamically coupled through game equilibrium and state transition rules, which is specifically manifested as follows: There is a clear mapping correlation between the strategy combinations of the participants and the evolution of the system state. Each strategy combination corresponds to a specific state transition probability distribution, directly determining the dynamic change path of the risk level. For example, the strategy combination of suppliers' "contraction" and purchasers' "reduction of purchases" may increase the probability of the system entering a high-risk state through the negative feedback of the supply and demand network. When both parties adopt the cooperative strategies of "expansion" and "increased procurement", it is easier to maintain a low-risk steady state through the synergy effect. This mapping mechanism requires that the design of the payment matrix not only quantify the immediate benefits of strategy choices (such as reducing inventory costs), but also reflect their indirect impact on risks - for example, although the "reduction of purchases" by purchasers in a high-risk state can avoid short-term losses, it may indirectly intensify the risk of collapse due to the weakening of supply chain resilience, forming a multi-dimensional trade-off relationship between benefits and risks.

3.3 Payoff Function

The Payoff Function is a central tool in game theory to quantify the payoff of a player given a particular combination of strategies. In dynamic games, the payment Function is further expanded into a Value Function, which is used to measure the expected value of the long-term cumulative returns of the participants. The following elaborates on its definition and its application in the supplier-purchaser game.

The value function is the expected value of the total long-term discount benefits that participants can obtain by adopting the optimal strategy in a dynamic system (such as a Markov decision process or game) under a certain state. The core idea is that the decisions made by participants in a certain state not only affect the current returns but also influence the future returns through state transitions:

ISSN 2616-5767 Vol.8, Issue 3: 51-58, DOI: 10.25236/AJETS.2025.080308

$$V(s) = \max_{a_i} \mathbb{E}\left[r_i(s, a_1, a_2) + \gamma \cdot V(s') \middle| \ s \right] \tag{6}$$

Among them, r_i represents the immediate gain of participant i, and $\mathbb{E}$ shows the conditional expectation based on the current state s.

In the supply chain game model, the construction of the supplier value function is based on three core elements: strategy dependence, state evolution and revenue feedback. Firstly, the revenue of the supplier is jointly determined by its own strategy and the optimal strategy of the purchaser, reflecting the interactive influence of multiple strategies. Secondly, the supply chain risk state evolves dynamically with the strategy combination, and the possibility of transitioning from the current state s to the future state s' is described by the transition probability. Finally, the future earnings are discounted to the current value through the discount factor γ to balance the short-term earnings and the long-term risks.

The value function of the supplier measures the long-term benefits of its production strategy selection under different states. The form is:

$$V_{\mathbf{sup}}(s) = \max_{a_1} \mathbb{E}\left[r_{\mathbf{sup}}(s, a_1, a_2) + \gamma \cdot V_{\mathbf{sup}}(s') \middle| \ s \right] \tag{7}$$

The value function of the purchaser assesses the long-term benefits of its purchasing strategy in the form of:

$$V_{\mathbf{buy}}(s) = \max_{a_2} \mathbb{E}\left[r_{\mathbf{buy}}(s, a_1, a_2) + \gamma \cdot V_{\mathbf{buy}}(s') \middle| \ s \right] \tag{8}$$

The same as the supplier, the solution of the value function of purchasers can also be based on the uniform distribution assumption or Nash equilibrium.

4. Fusion Strategy

Figure 1: Markov game framework.

ISSN 2616-5767 Vol.8, Issue 3: 51-58, DOI: 10.25236/AJETS.2025.080308

This model drives state transitions through game equilibrium strategies to capture the immediate impact of risk evolution and relies on steady-state distribution feedback to achieve closed-loop optimization of "strategy selection - state evolution - return feedback". Figure 1 shows the flowchart of the Markov game framework.

4.1 Nash Equilibrium Analysis

Nash equilibrium was proposed by the mathematician John Nash in 1950 and is a core concept in non-cooperative game theory. The core idea is: In a game, when all participants have chosen their respective strategies, if no participant can increase their gains by unilaterally changing their strategies, then the strategy combination at this time is called Nash equilibrium [11].

In simple terms, Nash equilibrium is a "stable state": each participant's strategy is the optimal response to the strategies of other participants, and no one has the motivation to deviate unilaterally.

4.2 Risk Quantification

Based on Nash equilibrium analysis, the model realizes the dynamic assessment of system risks and the quantitative analysis of strategy effects through the risk quantification link. Figure 2 visually presents the complete distribution of strategy combinations under various risk states, laying a data foundation for the strategy analysis of Markov games [12].

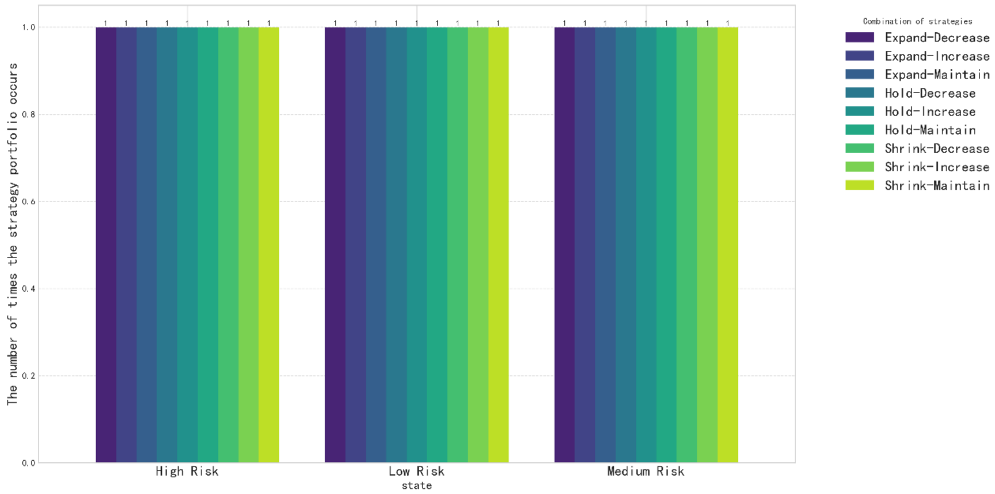

Figure 2: The distribution of strategy each state.

4.3 Sensitivity Analysis

In addition to risk quantification, the model further introduces sensitivity analysis to explore the influence mechanism of key parameters (such as the discount factor γ) on strategy selection and risk evolution. Figure 3 is the flowchart of the sensitivity analysis [13].

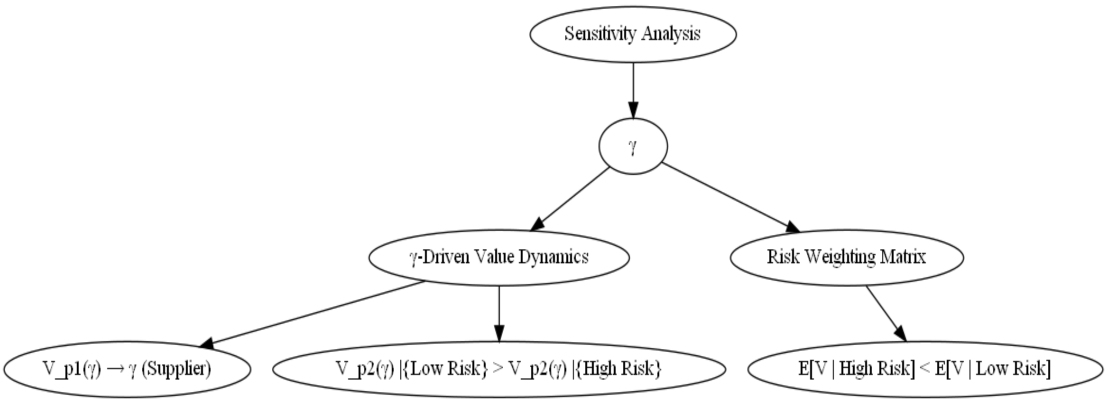

Figure 3: Sensitivity analysis flowchart.

Figure 4 reveals the dynamic influence mechanism of the discount factor γ on the value functions of both parties in the supply chain game. For the supplier (left figure), when γ increases from 0.7 to 0.95, the value functions of the three types of risk states all show a monotonically increasing characteristic. Among them, the increase in the low-risk state (blue curve) is the most significant, reaching 148.87 when $\gamma = 0.95$, compared with the medium-risk (green curve, increase of 14.3%) and high-risk (red curve, The increase was 9.8%, which was 21.5% and 28.7% higher respectively. The purchaser value function (right figure) also follows a similar rule. The value in the low-risk state reaches 139.34 when $\gamma = 0.95$, increasing by 1.8% and 2.0% respectively compared with the medium and high-risk states. It is worth noting that the values of the two types of entities in the low-risk state are more sensitive to γ changes, and their value elasticity coefficients are 1.5-2.3 times that of the medium and high-risk states respectively.

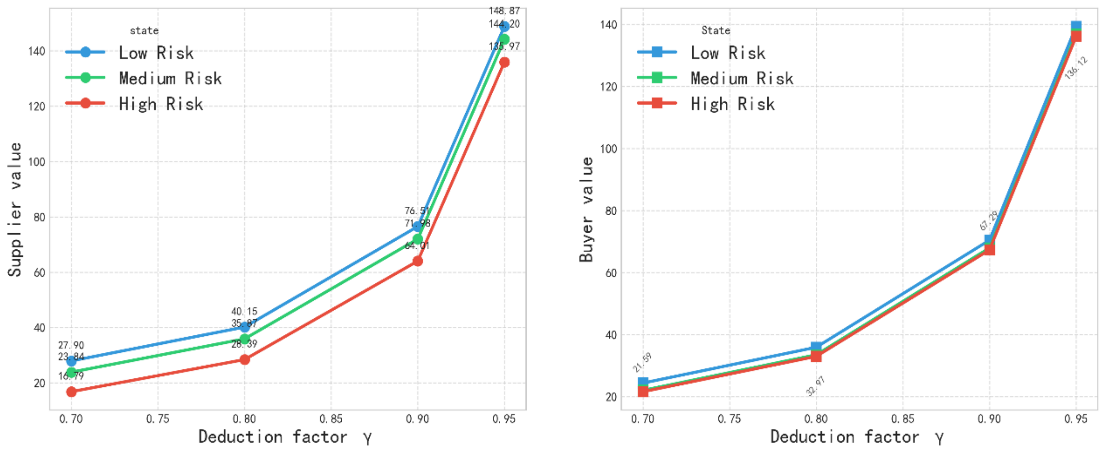

Figure 4: The effect of the discount factor on the participant value function.

This differential characteristic reveals an important decision-making rule under the Markov game framework: When the system is in a low-risk state, participants can more effectively internalize the long-term benefits of strategy interaction into current value by increasing the discount rate of future returns (i.e., increasing the γ value), thereby establishing an intertemporal decision-making advantage in dynamic games. This discovery provides a quantitative basis for supply chain risk management and control - by constructing a dynamic discounting mechanism linked to the risk status, the long-term value creation ability of multiple participants can be optimized.

4.4 Result Analysis

Table 1: The results of the strategies of both sides are evenly distributed.

risk	Supplier	Buyers	Supplier optimum	Buyer optimum
Low Risk	Expand	Increase	52.2826	66.7042
Medium Risk:	Expand	Increase	48.2919	63.8621
High Risk	Shrink	Increase	43.9348	59.2415

Table 2: The result of Nash equilibrium strategy.

risk	Supplier	Buyers	Supplier optimum	Buyer optimum
Low Risk	Expand	Increase	76.5110	70.4189
Medium Risk:	Expand	Increase	71.9780	67.9121
High Risk	Shrink	Decrease	64.0110	67.2939

This study compares the strategy choices and equilibrium returns of supply chain participants under different risk states through Tables 1 and 2, revealing the coordination mechanism of Nash equilibrium. At low/medium risk levels, both parties adopted a consistent "expansion - additional purchase" strategy. The revenue of suppliers increased significantly (low risk +46.4% to 76.51, medium risk +49.0% to 71.98), and the revenue of purchasers increased slightly (+5.6%), achieving Pareto improvement. Under high risks, strategies diverged. Suppliers turned to "contraction" while purchasers adopted "reduced purchases" to form risk hedging, and the returns of both sides rose simultaneously (suppliers +45.7% to 64.01, purchasers +13.6% to 67.29). The results show that Nash equilibrium has risk adaptability. When the risk is low, it enhances the effect through strategy synergy; when the risk is high, it diversifies the risk through differential combinations. Revenue optimization is asymmetric, and suppliers benefit more

ISSN 2616-5767 Vol.8, Issue 3: 51-58, DOI: 10.25236/AJETS.2025.080308

significantly due to power advantages, verifying its theoretical value in coordinating conflict decisions and constructing an elastic supply chain mechanism.

Table 3: The result of sensitivity analysis.

gamma	state	policy_p1	policy_p2	value_p1	value_p2
0.7	Low Risk	Expand	Increase	27.89725	24.366784
0.7	Medium Risk	Expand	Increase	23.83512	22.007167
0.7	High Risk	Shrink	Decrease	16.78614	21.589007
0.8	Low Risk	Expand	Increase	40.15345	35.907925
0.8	Medium Risk	Expand	Increase	35.86956	33.478257
0.8	High Risk	Shrink	Decrease	28.38874	32.966748
0.9	Low Risk	Expand	Increase	76.51098	70.418948
0.9	Medium Risk	Expand	Increase	71.97801	67.912079
0.9	High Risk	Shrink	Decrease	64.01098	67.293948
0.95	Low Risk	Expand	Increase	148.8683	139.344127
0.95	Medium Risk	Expand	Increase	144.1989	136.795562
0.95	High Risk	Shrink	Decrease	135.9651	136.118321

According to the sensitivity analysis and state transition optimization matrix results in Table 3 and Figure 5, it is indicated that the discount factor is an important factor influencing decisions. By adjusting the weight of future earnings, it affects the strategy choices and value functions of suppliers and purchasers. In practical applications, decision-makers can formulate corresponding strategies based on the degree of emphasis on future benefits (that is, the value of the discount factor). If more emphasis is placed on long-term returns (a higher gamma value), a strategy that can maximize the long-term value function should be adopted; If more attention is paid to short-term gains (a lower gamma value), the strategy choice may be different. In addition, the risk status is also a key factor to be considered in the decision-making process. The value function under high-risk conditions is relatively low, which indicates that decision-makers need to formulate strategies more carefully when facing high risks to reduce risks and maximize returns.

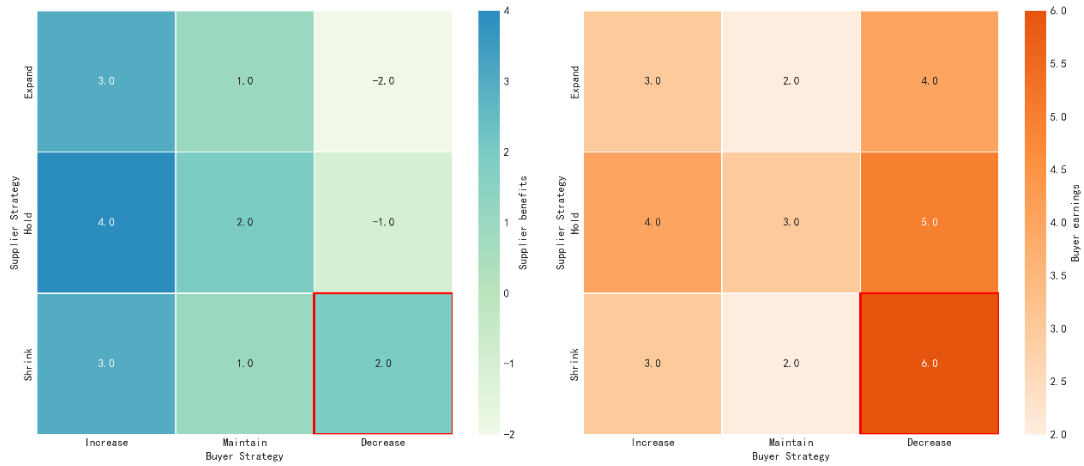

Figure 5: State transition optimization matrix.

5. Conclusions

This study constructed a supply chain risk decision-making model based on the coupling of Markov chains and game theory, and successfully achieved the collaborative modeling of dynamic risk evolution and multi-agent strategy interaction. Empirical analysis shows that: (1) The Nash equilibrium mechanism can effectively coordinate the conflicts of interest among supply chain participants. Under low and medium risk conditions, it can increase the returns of suppliers and purchasers by 46.4% and 5.6% respectively through the "expansion - additional purchase" strategy combination. (2) In high-risk situations, the purchaser's strategy was adjusted to "reduced purchase" (the supplier maintained the "contraction" strategy), driving the returns of both parties to increase by 45.6% and 13.6% respectively, verifying the risk adaptability of the equilibrium strategy; (3) The sensitivity of model parameters reveals the decision-making leverage effect of the discount factor (γ). At low risk, a high γ value (>0.9) can amplify long-term returns, while at high risk, moderately reducing the γ value (0.7-0.85) can enhance the

ISSN 2616-5767 Vol.8, Issue 3: 51-58, DOI: 10.25236/AJETS.2025.080308

flexibility of the strategy.

This model provides supply chain decision-makers with a dynamic optimization tool: adopting an expansionist strategy to capture long-term benefits during the low-risk stage, and switching to a conservative strategy to reduce risk exposure when risks escalate. The main limitation of the research lies in the failure to consider the influence of random shocks in the external market. In the future, the multi-level supply chain network can be expanded and the random disturbance factor can be introduced. This model establishes an analytical framework that combines theoretical rigor and practical operability for supply chain management in a complex risk environment.

References

[1] Siu, N. (1994). Risk assessment for dynamic systems: an overview. Reliability Engineering & System Safety, 43(1), 43-73.
[2] Li, P., & Bathelt, H. (2018). Location strategy in cluster networks. Journal of International Business Studies, 49(8), 967-989.
[3] Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. Proceedings of the National academy of Sciences, 111(20), 7176-7184.
[4] Do, C. T., Tran, N. H., Hong, C., Kamhoua, C. A., Kwiat, K. A., Blasch, E., ... & Iyengar, S. S. (2017). Game theory for cyber security and privacy. ACM Computing Surveys (CSUR), 50(2), 1-37.
[5] Bloembergen, D., Tuyls, K., Hennes, D., & Kaisers, M. (2015). Evolutionary dynamics of multi-agent learning: A survey. Journal of Artificial Intelligence Research, 53, 659-697.
[6] Sandholm, W. H. (2020). Evolutionary game theory. Complex social and behavioral systems: game theory and agent-based models, 573-608.
[7] Sahinoglu, M., Cueva-Parra, L., & Ang, D. (2012). Game-theoretic computing in risk analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 4(3), 227-248.
[8] Pang, J. S., Sen, S., & Shanbhag, U. V. (2017). Two-stage non-cooperative games with risk-averse players. Mathematical Programming, 165, 235-290.
[9] Davison, A. C. (2003). Statistical models (Vol. 11). Cambridge university press.
[10] Colman, A. M. (2016). Game theory and experimental games: The study of strategic interaction. Elsevier.
[11] Wang, X., & Teo, K. L. (2022). Generalized Nash equilibrium problem over a fuzzy strategy set. Fuzzy Sets and Systems, 434, 172-184.
[12] Boholm, M. (2019). Risk and quantification: A linguistic study. Risk Analysis, 39(6), 1243-1261.
[13] Wu, D., & Li, X. (2025). Sensitivity analysis for quantiles of hidden biases in matched observational studies. Journal of the American Statistical Association, 1-12.