Innovative Study on Government Debt Risk Contagion Model Integrating the Copula Model with Multi-Agent Reinforcement Learning Algorithm

Zebin Liu, *Member, CFA,* Xiaoheng Zhang, Hui Lv, and Lei Shen

Abstract—**Amid increasing global economic uncertainty, local government debt burdens have significantly intensified, triggering growing concerns regarding cross-regional debt contagion risks. Traditional approaches, such as Copula-based models, effectively capture static dependencies but lack responsiveness to dynamic policy environments. Conversely, Multi-Agent Reinforcement Learning (MARL) offers strong adaptability for strategic optimization but often falls short in accurately modeling complex dependency structures. To overcome these limitations, this research introduces the Dynamic Copula-Reinforcement Learning Contagion Model (DCRL-CM), effectively integrating the strengths of both methodologies to dynamically model debt contagion pathways. Employing empirical data from 16 cities in Anhui Province, China, the DCRL-CM robustly characterizes the evolution of inter-city debt dependencies under diverse economic and fiscal conditions, allowing for dynamic strategic adjustments. The proposed model demonstrates superior forecasting performance regarding debt risk contagion. Dynamic simulations unveil spatially heterogeneous risk transmission patterns and highlight a distinct core-periphery contagion structure. Compared with traditional static models, the DCRL-CM substantially enhances the accuracy of modeling nonlinear debt contagion dynamics, effectively mitigating systemic risk amplification in core urban areas while facilitating decentralized risk management practices regionally. These results provide a solid analytical foundation for the dynamic management of regional governmental debt risks and offer critical policy insights for strengthening fiscal resilience and advancing sustainable economic development.**

Index Terms—**regional government debt risk, dynamic contagion model; multi-agent reinforcement learning; Copula function**

Manuscript received November 14, 2024; revised June 25, 2025.

This study was financially supported by Humanity and Social Science Research Project of Anhui Educational Committee (2023AH051164,2023AH051512,2023AH051508), Huainan Normal University School-Level Key Project on Quality Engineering Education and Teaching Reform (2023hsjyxm02), and Huainan Normal University 2024 Key School-Level Scientific Research Project (2024XJZD005,2023XJZD018).

Zebin Liu is a lecturer of Huainan Normal University, Huainan, 232038, China (e-mail: jack1776@163.com).

Xiaoheng Zhang is a lecturer of Anhui University of Science and Technology, Huainan, 232038, China (phone: +86-177-5548-9921; e-mail: 1422150820@qq.com)

Hui Lv is a PhD candidate of Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, China (e-mail: 417181314@qq.com).

Lei Shen is a lecturer of Huainan Normal University, Huainan, 232038, China (e-mail: 1048790944 6@qq.com).

I. Introduction

In recent years, the contagion effect associated with regional government debt has emerged as a critical focus in global economic research. Amid escalating uncertainties surrounding global economic growth, numerous regional governments have encountered progressively burdensome debt obligations[1]. This scenario not only endangers local economic stability but also exacerbates systemic contagion risks through fiscal interdependencies among regions, potentially destabilizing broader national economies[2]. Since the outbreak of the COVID-19 pandemic in 2020, nations worldwide have adopted substantial fiscal stimulus packages to mitigate economic disruptions, consequently triggering significant increases in regional government indebtedness[3]. To sustain local economies during this challenging period, regional authorities have notably elevated public expenditure and intensified debt issuance, sharply amplifying their cumulative debt burdens. Concurrently, recessionary pressures triggered by the pandemic, coupled with declining tax revenues, have considerably undermined regional governments' fiscal capabilities and debt-servicing capacities[4, 5]. Against this backdrop, the transmission of debt risks across regions has become increasingly pronounced. If left unaddressed, these vulnerabilities could precipitate nationwide debt crises, further jeopardizing macroeconomic stability. Indeed, since the pandemic's onset, regional governments globally have faced mounting fiscal pressures as increased pandemic-related expenditures and reduced revenues have created a vicious cycle of escalating debt burdens, diminished repayment abilities, and heightened dependency on refinancing, thus fostering structural debt vulnerabilities in numerous regions[6].

To elucidate the dynamic characteristics and transmission pathways associated with regional government debt risk, numerous studies have employed various methodologies to capture dependency structures. Among these methods, the Copula approach has emerged prominently, given its effectiveness in modeling complex risk contagion phenomena due to its ability to accurately represent nonlinear dependencies[7]. By separately defining marginal and joint distributions, Copula methods provide substantial flexibility for constructing intricate dependency frameworks among financial variables. Specifically, techniques such as t-Copula and Vine Copula have demonstrated notable efficacy in capturing tail dependencies, thus making them particularly suitable for analyzing extreme dependence relationships inherent in debt-risk contagion networks[8].Nonetheless, despite its advantages, the traditional Copula methodology inherently captures static dependency structures, limiting its effectiveness in dynamically changing environments, particularly those influenced by economic policy shifts[9, 10]. In rapidly evolving financial conditions, fixed dependency frameworks might inadequately reflect the adaptive complexities inherent in debt contagion pathways [11], thereby restricting their applicability in risk forecasting and management. Consequently, researchers have increasingly sought more adaptive modeling approaches. Multi-agent reinforcement learning (MARL), in particular, has gained considerable attention in financial modeling due to its adaptability and potential for strategic optimization. Reinforcement learning leverages iterative trial-and-error processes combined with policy optimization, thereby offering unique advantages for simulating intricate interactions among financial agents [12, 13]. In multi-agent settings, coordinated decision-making effectively encapsulates the systemic interdependencies, and such frameworks have already been successfully applied in various financial contexts, including investment management, market dynamics, and trading behaviors[14-16]. Recent developments in deep reinforcement learning have further augmented the capabilities of MARL, enabling sophisticated deep-network architectures to devise strategies and optimize decision-making processes efficiently. Thus, MARL represents a particularly promising approach for addressing complex contagion risks associated with multi-regional government debt[17].

While both Copula methods and MARL possess distinct advantages for financial risk modeling, existing research has yet to integrate these methodologies to examine regional government debt contagion. Traditional Copula models typically fall short in capturing real-time dependency shifts caused by dynamic economic policies and external shocks[18]. In contrast, although MARL effectively captures adaptive and strategic agent interactions, it struggles to accurately represent the intricate dependency structures among financial entities[19]. To overcome these limitations, this study proposes a novel integrated framework that merges Copula modeling with MARL, thus forming an adaptive and dynamic contagion model. This integrated approach successfully identifies contagion pathways of regional government debt risk within dynamic and uncertain environments. By simultaneously modeling sophisticated dependency structures and dynamically simulating policy-driven regional responses, the proposed methodology provides a more comprehensive and realistic depiction of the complex interactions and path dependencies inherent in debt contagion phenomena.

The primary contribution of this study is the development of an innovative debt contagion model that integrates Copula methods with MARL, addressing critical shortcomings inherent in traditional dynamic risk modeling approaches. By leveraging MARL's adaptive modeling capabilities, the proposed framework effectively analyzes debt contagion under diverse policy scenarios and external economic shocks, yielding valuable theoretical insights and practical recommendations for regional government debt management. This integration not only enhances the theoretical understanding of financial risk contagion but also establishes a novel methodological foundation that can inform the formulation and optimization of regional fiscal policies.

II. THEORETICAL MODEL

Compared to traditional static Copula models and single-government decision-making frameworks, the proposed DCRL-CM model effectively captures the dynamic nature of government debt risks as well as intricate inter-regional interactions under changing economic conditions. By integrating the GARCH(1,1)-t mode [20] for accurate estimation of marginal distributions and employing the Vine Copula decomposition method[21] for constructing high-dimensional dependency structures, the proposed model robustly characterizes nonlinear and tail dependencies among regional risk indicators, particularly during economic downturns and periods of policy uncertainty. Furthermore, by conceptualizing regional governments as intelligent agents operating within a Centralized Training and Decentralized Execution (CTDE) framework [22], the model supports coordinated policy optimization. This design explicitly addresses the dynamic trade-offs between fiscal revenue generation and effective debt risk management, thus providing enhanced policy guidance and adaptability.

A. Assumptions

In this study, we begin by assuming that the government debt risk indicators for each region are represented by sufficiently long and stationary time series. This assumption enables reliable application of the GARCH (1,1)-t model, effectively capturing volatility clustering and heavy-tailed distributions typically observed in financial data. Furthermore, in constructing the high-dimensional joint distribution, we posit that conditional Copula parameters depend on bounded exogenous macroeconomic variables—such as GDP growth and fiscal self-sufficiency—and historical residual values. Ensuring boundedness is critical for maintaining numerical stability during parameter estimation. By employing suitable nonlinear transformations (e.g., logit or hyperbolic tangent), we further guarantee that estimated parameters consistently remain within their theoretically permissible ranges. Additionally, we conceptualize each regional government as an intelligent decision-making agent, assuming that their state and action spaces are either finite or compact, and their corresponding reward functions are continuous and differentiable. These assumptions satisfy the necessary conditions for applying classical Markov game theory and conducting convergence analyses in multi-agent frameworks. Lastly, we adopt a "slowly changing environment" assumption, indicating that dynamic Copula parameters evolve more gradually compared to agents' decision-making processes. Consequently, the environment can be approximated as quasi-static within each short-term analytical interval. Collectively, these assumptions lay a rigorous theoretical foundation essential for model development, algorithm design, empirical analysis, and subsequent theoretical validation.

B. Marginal distribution modeling

Building upon the framework introduced by Han et al.[23], this study employs a time-varying Copula model to examine the dependency structure among local government debt risk indicators. Additionally, macroeconomic variables are incorporated as exogenous drivers, enabling dynamic analysis of the impacts of policy shifts and external shocks on risk contagion pathways. Initially, marginal distributions for diverse debt-risk indicators are estimated separately; subsequently, Copula functions are utilized to model the interdependencies among these indicators. Let $I_i(t)$ denote the observation of the *i-th* debt-risk indicator at time *t* (e.g., outstanding debt balance, growth rate, or interest payment ratio). To effectively capture the volatility clustering and heavy-tailed characteristics commonly found in financial time series data, this research applies the GARCH(1,1)-t model, formulated as follows:

$$\begin{gathered} I_i(t) = \mu_i + \epsilon_i(t), \\ \epsilon_i(t) = \sqrt{h_i(t)} \cdot \eta_i(t), \eta_i(t) \sim t(\nu_i)(0,1), \\ h_i(t) = \omega_i + \alpha_i \epsilon_i^2(t-1) + \beta_i h_i(t-1), \end{gathered} \tag{1}$$

where, μ_i and $\epsilon_i(t)$ are the mean and random noise respectively; $\eta_i(t)$ follows a *Student-t* distribution with ν_i degrees of freedom; $h_i(t)$ is the conditional variance. The parameters $\{\omega_i, \alpha_i, \beta_i\}$ satisfy $\omega_i > 0, \alpha_i \ge 0, \beta_i \ge 0$, and $\alpha_i + \beta_i < 1$ to ensure model stationarity. Once the marginal parameters have been estimated, the corresponding cumulative distribution function $F_i(\cdot)$ is used to transform the residuals into probability variables:

$$u_i(t) = F_i(\epsilon_i(t)) \in (0,1), \tag{2}$$

which serve as the input data for subsequent joint distribution modeling.

C. Construction of a high-dimensional dynamic copula model

Because multiple debt-risk indicators are involved, directly fitting a high-dimensional Copula model can lead to an excessive number of parameters, making estimation difficult. To address this, we employ a Vine Copula decomposition, representing the joint distribution of N indicators, denoted by $H(I_1,...,I_N)$, as a product of multiple bivariate conditional Copulas. Concretely, we define

$$u(t) = \left(u_1(t), \ldots, u_N(t)\right), \tag{3}$$

and based on either a C-vine or D-vine structure, we have

$$\begin{gathered} H(I_1,\ldots,I_N) = \prod_{k=1}^{N-1} \prod_{j=1}^{N-k} c_{j,j} \\ +k \mid s(u_j(t), u_{j+k}(t) | u_S(t)), \end{gathered} \tag{4}$$

where $c_{j,j+k} \mid S(\cdot)$ denotes the Copula density between dimensions *j* and *j+k*, conditional on the variables in the set *S(t)*.

To capture the influence of shifts in economic conditions (e.g., macroeconomic recessions or fiscal policy adjustments) on the dependency structures of regional debt risks, we introduce time-varying mechanisms. Specifically, let $\Theta_{j,j+k|s}(t)$ denote the parameter(s) of each conditional Copula at time *t*. We then specify an update function:

$$\Theta_{j,j+k|S}(t) = \Gamma\left(\Theta_{j,j+k|S}(t-1), Z(t), \gamma\right), \tag{5}$$

where $\Gamma(\cdot)$ is a mapping that adjusts the Copula parameters based on their previous values $\Theta_{j,j+k|s}(t-1)$, current macro variables *Z(t)*, and a set of hyperparameters γ . Through this dynamic update rule, our model can more accurately reflect how economic environment changes affect inter-regional debt risk dependence. The update function $\Gamma(\cdot)$ can be designed as either a linear or nonlinear mapping. To ensure that updated parameters remain within theoretically permissible bounds, we often apply a logit or hyperbolic tangent transformation. For example, one might choose

$$\Lambda(x) = \tanh\left(\frac{x}{2}\right), \tag{6}$$

which maps any real input $x \in \mathbb{R}$ into the interval (-1,1). To further strengthen the theoretical foundation, we provide a rigorous proof of the stability and parameter sensitivity of the update function in the proof of Theorem 1.

Theorem 1 (contraction property and local stability of dynamic copula parameter updates): Let $\Gamma : X \to X$ be an update function defined on a complete normed space ($X, \|\cdot\|$). Suppose Γ satisfies the Lipschitz continuity condition with a constant $L \in [0,1]$, meaning that for any $x, y \in X$:

$$\| \Gamma(x) - \Gamma(y) \| \le L \| x - y \| . \tag{7}$$

Define the update process recursively as $x_{n+1} = \Gamma\left(x_n\right)$. Under these assumptions:

(1) contraction mapping: Γ is a contraction mapping, so there exists a unique fixed point $x^* \in X$ such that $x^* = \Gamma(x^*)$;

(2) convergence: for any initial point $x_0 \in X$, the iterative sequence $\{x_n\}$ converges to $x^\star$. Moreover, the error bound satisfies:

$$\|x_n - x^*\| \le L^n \| x_0 - x^* \| ; \tag{8}$$

(3) parameter stability: if Γ is differentiable, then by the mean value theorem[24], for any *x* and any small perturbation $\sum x$ we have

$$\| \Gamma(x + \Delta x) - \Gamma(x) \| \le \| \Gamma'(x) \| \| \Delta x \| . \tag{9}$$

Under the Lipschitz condition, we require that

$$\| \Gamma'(x) \| \le L, \ \forall x \in X. \tag{10}$$

This ensures that small perturbations in the input result in proportionally small changes in the output, demonstrating that Γ exhibits relatively low parameter sensitivity.

Proof of theorem 1: Given that Γ is Lipschitz continuous [25] on X with constant $L \in (0,1)$, for any initial point $x_0 \in X$, there exists

$$\| x_{n+1} - x_n \| = \| \Gamma(x_n) - \Gamma(x_{n-1}) \| \le L \| x_n - x_{n-1} \| \tag{11}$$

where *L* is the Lipschitz constant. By recursively iterating, we obtain

$$\| x_{n+1} - x_n \| \le L^n \| x_1 - x_0 \| . \tag{12}$$

Moreover, for any *m*>*n*, applying the triangle inequality yields

$$\begin{gathered} \| x_m - x_n \| \le \sum_{k=n}^{m-1} \| x_{k+1} - x_k \| \le \sum_{k=n}^{\infty} L^k \| x_1 - x_0 \| \\ = L^n \cdot \frac{1}{1-L} \| x_1 - x_0 \| . \end{gathered} \tag{13}$$

Since $0<L<1$, it follows that as $n \to \infty$, L^n converges to 0. Consequently, $\| x_m - x_n \|$ tends to 0, which implies that the sequence $\{x_n\}$ is a Cauchy sequence.

As the space X is complete, this Cauchy sequence must converge to some $x^* \in X$. By the continuity of Γ , taking the limit as $n \to \infty$ gives

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \Gamma(x_n) = \Gamma\left(\lim_{n\to\infty} x_n\right) = \Gamma(x^*), \quad (14)$$

demonstrating that x^* is a fixed point of Γ .

To prove the uniqueness of the fixed point, suppose there exists another fixed point $y^* \in X$. Then, by the Lipschitz condition,

$$\| x^* - y^* \| = \| \Gamma(x^*) - \Gamma(y^*) \| \le L \| x^* - y^* \|. \quad (15)$$

Since $L<1$, the only possibility is $\| x^* - y^* \| = 0$, which implies that $x^* = y^*$. Thus, the fixed point is unique. Finally, if Γ is differentiable and satisfies $|\Gamma'(x)| \le L$, then for any small perturbation Δx , the mean value inequality yields

$$\| \Gamma(x+\Delta x) - \Gamma(x) \| \le L \| \Delta x \|. \quad (16)$$

This shows that the mapping Γ is relatively insensitive to small changes in its argument, ensuring that each iterative update remains stable. Moreover, the convergence rate is characterized by

$$\| x_n - x^* \| \le L^n \| x_0 - x^* \|, \quad (17)$$

which implies that the sequence converges at a geometric rate determined by the constant L . This behavior ensures excellent robustness and stability in numerical implementations.

D. Multi-Agent reinforcement learning system

Building upon the high-dimensional dynamic Copula model, we seek to capture the decision interactions among regional governments in debt management by treating each regional government as an agent and constructing a multi-agent reinforcement learning (MARL) system. For each agent *i*, its state vector $S_i(t)$ is composed of the following three components:

(1) local debt risk indicators derived from the GARCH (1,1)-t model (e.g., debt balance, debt growth rate, interest burden, etc.);

(2) Dependence parameters extracted from the dynamic Vine Copula model, such as the correlation coefficients $\rho_{i,j,t}$ with other regions;

(3) Exogenous macroeconomic variables $Z(t)$, for instance, GDP growth rate and fiscal self-sufficiency.

We denote the overall system state as

$$S(t) = \{S_1(t), S_2(t), \ldots, S_d(t)\}. \quad (18)$$

The action $a_i(t)$ taken by agent *i* includes decisions on debt issuance, risk management measures, and fiscal strategies. To balance the government's pursuit of fiscal revenue against the goal of reducing debt risk, we define the following reward function:

$$r_i(t) = \lambda \Pi_i\left(a_i(t), S(t)\right) - (1-\lambda)\Psi_i\left(a_i(t), S(t)\right), \quad (19)$$

where Π_i measures fiscal revenue, Ψ_i captures the risk of debt default or contagion, and $\lambda \in (0,1)$ is a weighting coefficient. During the training phase, we adopt a centralized training, decentralized execution (CTDE) framework. Each agent shares global state information while training, and updates its policy via multi-agent reinforcement learning algorithms such as MADDPG [26] or QMIX[27]. The objective is to maximize each agent's discounted cumulative reward:

$$\max_{\pi_i} \mathrm{E}\left[\sum_{t=0}^{\infty} \gamma^t r_i(t)\right], \quad (20)$$

where π_i denotes the policy function of agent *i*, and $\gamma \in (0,1)$ is the discount factor. To demonstrate that the system still converges to a local equilibrium after introducing dynamic Copula parameters (i.e., dynamic environmental changes), we propose the following theorem.

Theorem 2 (Existence of local Nash equilibrium in a multi-agent system under a time-varying environment). Consider a Markov game with *d* agents, where the state space S and each agent's action space A_i are either finite or compact, and the reward function $r_i(s,a)$ is continuous in both *s* and *a*. We introduce a dynamic Copula parameter $\Theta(t)$ to capture the time-varying nature of the environment, but assume that within each short time interval of length ΔT , $\Theta(t)$ can be approximated as constant. If the following conditions hold:

(1) the state and action spaces are finite or compact, and the reward function $r_i(s,a)$ is continuous;

(2) during each stage $[t, t+\Delta T)$, the exogenous parameter $\Theta(t)$ changes slowly enough that $\Theta(t) \approx \Theta(t+\tau)$ for all $\tau \in [0, \Delta T)$;

(3) Within each stage, the agents are trained for a sufficiently large number of iterations, satisfying conditions on diminishing learning rates and stochastic approximation;

(4) Each agent's policy $\pi_i(a_i | s)$ is continuously differentiable with respect to its parameters, and the reward function is differentiable with respect to the actions.

Then, at the end of each stage, there exists a set of strategies $\{\pi_i^*\}$ such that

$$\pi_i^* \approx \arg\max_{\pi_i} \mathrm{E}\left[\sum_{\tau=t}^{t+\Delta T-1} \gamma^{\tau-t} r_i\left(s(\tau), a(\tau)\right)\right], \quad (21)$$

meaning that within each stage, the system approximately attains a local Nash equilibrium. Moreover, as the number of training iterations within each stage tends to infinity, the approximation error converges to zero.

Proof of Theorem 2. Below is the detailed proof process:

(1) stagewise near-staticity and local equilibrium existence: during the *k-th* stage, which corresponds to the time interval the time interval $[k\Delta T, (k+1)\Delta T]$, we assume that for any $t \in [k\Delta T, (k+1)\Delta T]$ the environmental parameter $\Theta(t)$ satisfies

$$\| \Theta(t) - \Theta_k \| \le \epsilon_k, \quad (22)$$

where ϵ_k is sufficiently small. Owing to the structural robustness commonly observed in Markov games[28], this implies that the environment can be regarded as "approximately stationary." Concretely, for all states *s* and actions *a*,

$$\max_{(s,a)} \| P(s_{t+1} | s_t, a_t; \Theta(t)) - P(s_{t+1} | s_t, a_t; \Theta_k) \| \le \kappa\epsilon_k, \quad (23)$$

where *k* is a constant related to the structure of the game.

Therefore, over this stage, the actual Markov game can be treated as a "quasi-static" game:

$$G_k = \left(S, A, P(\cdot| s,a;\Theta_k), \{r_i\}_{i=1}^N, B, T\right), \qquad (24)$$

In the above finite state–action game setting, classical Markov game theory has established that: for a finite time horizon, one can use the Kakutani–Glicksberg–Fan fixed point theorem [29] to prove the existence of at least one (local) Nash equilibrium; for the infinite discounted time horizon case, a similar existence result can be obtained via the Shapley equation. Consequently, in each stage G_k , there exists at least one (local) Nash equilibrium strategy profile:

$$\pi_k^* = (\pi_{k,1}^*, \pi_{k,2}^*, \dots, \pi_{k,N}^*); \qquad (25)$$

(2) quasi-convergence of intra-stage strategy updates: in the *k-th* phase, consider Θ_k as a fixed parameter. Each agent can employ various learning algorithms for Markov games (e.g., multi-agent policy gradient, Q-learning, value function decomposition, etc.). Let

$$\pi_{k,n} = \left(\pi_{k,n}^1, ..., \pi_{k,n}^N\right) \qquad (26)$$

denote the policies of all agents at the *n-th* iteration. Taking policy gradient as an example, the update for a single agent *i* can be expressed as the following stochastic approximation:

$$\pi_{k,n+1}^i = \Pi\left[\pi_{k,n}^i + \alpha_{k,n} \nabla_{\pi_i} J_i\left(\pi_{k,n}^i, \pi_{k,n}^{-i}; \Theta_k\right)\right], \quad (27)$$

where $\Pi[\cdot]$ projects onto the feasible policy space (e.g., a probability simplex), $\alpha_{k,n}$ is the step size, and $\nabla_{\pi_i} J_i$ is an unbiased estimator of the gradient of the payoff function J_i (which can be derived from sampled trajectories). In the multi-agent setting, if appropriate smoothness assumptions are satisfied—along with independent sampling or a mixing Markov decision process—and if the learning rates meet certain convergence conditions (e.g., $\sum_n \alpha_{k,n} = \infty$, and $\sum_n \alpha_{k,n}^2 < \infty$), then in the fixed game environment Θ_k , the joint policy $\pi_{k,n}$ of all agents may converge to a local Nash equilibrium (or, in the worst case, to a stable fixed point). Formally,

$$\lim_{n\to\infty} \pi_{k,n} = \pi_k^*. \qquad (28)$$

If each phase is limited to M_k iterations, then there is a small discrepancy between the final policy π_{k,M_k} and π_k^* :

$$\| \pi_{k,M_k} - \pi_k^* \| \le \epsilon_k', \qquad (29)$$

where $\epsilon_k' \to 0$ as $M_k \to \infty$. In other words, if sufficiently many iterations are performed in a single phase, the policy can become arbitrarily close to a local Nash equilibrium;

(3) environmental updates and the continuous dependence of policies: Suppose that at the end of phase *k*, the environment parameters are updated according to

$$\Theta_{k+1} = \Theta_k + \delta_k, \| \delta_k \| \le \epsilon_k'', \qquad (30)$$

where ϵ_k'' is sufficiently small. From the earlier Lipschitz or continuity assumptions, it follows that for all (s,a) ,

$$\left\| P(\cdot | s,a;\Theta_{k+1}) - P(\cdot | s,a;\Theta_k) \right\| \le c_0 \| \delta_k \|, \qquad (31)$$

where c_0 is some constant. This implies that the dynamic model of game G_{k+1} is very close to that of G_k . Consequently, its local Nash equilibrium π_{k+1}^* should also be close to π_k^* . Formally,

$$\| \pi_{k+1}^* - \pi_k^* \| \le c_1 \| \delta_k \|. \qquad (32)$$

Hence, when δ_k is small, the optimal (or local Nash) strategies of the new and old phases remain close to each other. Because $\| \pi_{k+1}^* - \pi_k^* \|$ is small, using π_k^* (the converged policy from the previous phase) as the initial policy for phase (*k+1*) requires only a small number of iterations to converge to π_{k+1}^* . If only a finite number of iterations is performed, the resulting policy still approximates π_{k+1}^* . If the phase is sufficiently long or the number of iterations is large, the error can be reduced to an arbitrarily small level.

Thus, the proof presented above substantiates the conclusion stated in *Theorem 2*: within each nearly static analytical phase, multi-agent policy updates converge to a local Nash equilibrium. Furthermore, this convergence process demonstrates continuous stability concerning minor variations in environmental parameters.

Theorem 3: Global stability and convergence of the alternating iterative update algorithm. In the system described above, if the dynamic copula parameter update function $\Gamma(\cdot)$ satisfies a Lipschitz continuity condition , and if the multi-agent policies are capable of converging to a local equilibrium in a fixed environment, then by adopting an alternating iterative update strategy (i.e., updating the agents' policies first in each phase and then updating the copula parameters), the overall system parameters will converge to a fixed point, and the system will exhibit global asymptotic stability.

Proof of Theorem 3: Let *p* denote the collection of all agent policies, which belongs to the policy space *P*; let *c* denote the set of all dynamic copula parameters, which belongs to the parameter space *C* (for example, $C \subset \mathbb{R}^k$), and assume that *C* is a complete space. We define the overall parameter vector as

$$x = (p, c), \qquad (33)$$

which resides in the space

$$X = P \times C. \qquad (34)$$

An appropriate norm is chosen, for instance,

$$\| x \| = \| p \|_P + \| c \|_C . \qquad (35)$$

Under a fixed dynamic copula parameter, the agent policies can be updated to new policies after sufficient training. We denote this update process by the mapping

$$T_p : P \to P, \qquad (36)$$

and assume that T_p satisfies the Lipschitz continuity property. That is, there exists a constant $L_p \in [0,1)$ such that for any $p_1, p_2 \in P$,

$$\| T_p(p_1) - T_p(p_2) \|_P \le L_p \| p_1 - p_2 \|_P . \qquad (37)$$

Similarly, the update process for the dynamic copula parameters is represented by the mapping

$$T_c : C \to C, \qquad (38)$$

with the update given by

$$c_{t+1} = T_c(c_t) = \Gamma(c_t, Z(t), \gamma), \quad (39)$$

where the function $\Gamma(\cdot)$ satisfies a Lipschitz continuity condition. That is, there exists $L_c \in [0,1)$ such that for any $c_1, c_2 \in C$,

$$\| T_c(c_1) - T_c(c_2) \|_C \leq L_c \| c_1 - c_2 \|_C . \quad (40)$$

From the proof of the preceding theorem, it is known that T_c is a contraction mapping, and its iteration sequence converges to a unique fixed point.

We now define the overall update mapping as

$$T: X \to X, \quad \text{with} \quad T(p,c) = \left(T_p(p), T_c(c)\right). \quad (41)$$

To prove that T is a contraction mapping, let $x_1 = \left(p_1, c_1\right)$ and $x_2 = \left(p_2, c_2\right)$ be any two points in X. Then,

$$\begin{aligned} &\| T(p_1,c_1) - T(p_2,c_2) \|_X = \\ &\| T_p(p_1) - T_p(p_2) \|_P + \| T_c(c_1) - T_c(c_2) \|_C . \end{aligned} \quad (42)$$

By using the respective Lipschitz properties, we obtain

$$\begin{aligned} &\| T(p_1,c_1) - T(p_2,c_2) \|_X \leq \\ &L_p \| p_1 - p_2 \|_P + L_c \| c_1 - c_2 \|_C . \end{aligned} \quad (43)$$

Let

$$L = \max\{L_p, L_c\}. \quad (44)$$

Then,

$$\begin{aligned} &L_p \|p_1 - p_2\|_P + L_c \|c_1 - c_2\|_C \\ &\leq L\left(\| p_1 - p_2 \|_P + \| c_1 - c_2 \|_C\right) = L\|x_1 - x_2\|_X , \end{aligned} \quad (45)$$

where $L < 1$ since both L_p and L_c are less than 1. Therefore, T is a contraction mapping. According to the Banach fixed-point theorem[30], there exists a unique fixed point $x^* = (p^*, c^*)$ in the complete space X. Moreover, for any initial point $x_0 \in X$, the iterative sequence defined by $x_{n+1} = T(x_n)$ satisfies

$$\| x_n - x^* \| \leq L^n \| x_0 - x^* \| . \quad (46)$$

Through the complete mathematical proof presented above, we have demonstrated that under the assumption that both the multi-agent policy update mapping T_p and the copula parameter update mapping T_c are contraction mappings, the overall update mapping

$$T(p,c) = \left(T_p(p), T_c(c)\right) \quad (47)$$

is also a contraction mapping. By the Banach fixed-point theorem, the overall system parameters will converge globally to a unique fixed point, and the convergence rate is geometric. This result not only guarantees the local stability and robustness of the dynamic Copula parameter update process, but also provides a rigorous mathematical basis for the theoretical convergence and global stability of the system in real economic scenarios (such as the transmission of debt risk among regional governments due to economic fluctuations and policy adjustments).

E. Coupling mechanism between reinforcement learning and dynamic copula updates

Combining the previous sections, we have constructed a dynamic Copula model together with a multi-agent reinforcement learning system and devised an alternating iterative update mechanism between the two components. The specific procedure is as follows:

(1) initialization: Historical data are used to estimate the parameters of the GARCH (1,1)-*t* model for each debt risk indicator, compute the residuals, and transform them into $u_i(t)$. A Vine Copula model is then fitted to the initial sample to obtain the initial dynamic Copula parameters $\Theta_{j,j+k} \mid S(0)$;

(2) stage-wise policy training: The entire time series is partitioned into several short-term phases. Within each phase, the dynamic Copula parameters are assumed to be fixed at Θ_k. Each agent selects actions $a_i(t)$ based on the state $S(t)$ (which includes debt indicators, Copula parameters, and exogenous variables) and receives feedback according to the reward function $r_i(t)$. Using a centralized training with decentralized execution (CTDE) framework, the agents update their respective policies until convergence is reached within the phase;

(3) parameter update: Based on the most recent economic data and debt risk information, the dynamic Copula parameters are updated using the function

$$\Theta_{j,j+k|S(t+1)} = \Gamma\left(\Theta_{j,j+k|S(t)}, Z(t), \gamma\right), \quad (48)$$

ensuring via Theorem 1 that the updated parameters remain within the valid domain;

(4) alternating iteration: Steps 2 and 3 are repeated until the overall system parameters—comprising both the agent policies and the dynamic Copula parameters—stabilize, thereby achieving global convergence.

Through this alternating iterative coupling mechanism, the model is not only capable of capturing the risk contagion effects among regional governments arising from economic fluctuations and fiscal policy adjustments, but also simulates the dynamic decision-making processes of governments engaged in interactive games. This, in turn, provides a rigorous theoretical foundation and strategic recommendations for practical economic policy-making.

III. Indicator System Construction

A. Data sources

This study focuses on the 16 cities of Anhui Province, China, using relevant data from 2013 to 2022. The data primarily come from the CSMAR database purchased by the author’s institution, big data analysis platforms, the China Statistical Yearbook, the China Financial Yearbook, the Anhui Statistical Yearbook, and the Anhui Statistical Bulletin. The Lagrange multiplier method is applied to handle missing data for certain years.

B. Indicator Construction

Using the 16 cities of Anhui Province as case studies, this research develops a comprehensive indicator system comprising three primary dimensions: debt risk, economic environment, and regional contagion (Table I). This indicator framework aims to quantitatively assess the debt pressure, economic conditions, and regional contagion characteristics for each city, serving as critical input data for the multi-agent reinforcement learning model. Specifically, the multidimensional indicator system developed in this study captures key factors influencing the transmission of local government debt risks. Debt risk indicators measure the scale, growth dynamics, and default pressures associated with local government indebtedness, thereby establishing an essential foundation for understanding risk transmission mechanisms and identifying vulnerabilities in debt accumulation that could impact fiscal sustainability. Indicators of the economic environment assess the structural capacity of local governments to manage and respond effectively to debt risks. Metrics such as GDP growth, fiscal autonomy, per capita income, and fixed asset investment elucidate each region's debt-bearing capability and resilience, highlighting how economic conditions moderate contagion processes. Finally, regional contagion indicators encompass financial linkages, debt correlations, structural economic similarities, and trade interdependencies among the cities, revealing spatial channels and mechanisms through which risks propagate. These indicators facilitate the identification of regional spillover effects and provide a theoretical basis for coordinated risk mitigation strategies. Collectively, this structured and empirically actionable indicator system not only supports robust theoretical analysis but also informs practical policy formulation, thereby enhancing strategies aimed at mitigating and controlling the spread of local government debt risks.

TABLE I
DEBT RISK INDICATORS

Indicator Category	Indicator name	Symbol
Debt Risk Indicators	Debt Balance Ratio	DR
	Debt Dependency Ratio	DDR
	Interest Payment Burden Ratio	DIBR
Economic Environment Indicators	Debt Growth Rate	DGR
	GDP Growth Rate	GDPGR
	Fiscal Self-sufficiency Ratio	FSR
	Per Capita GDP	PCGDP
	Fixed Asset Investment Growth Rate	FAIGR
Regional Contagion Indicators	Fiscal Dependence Between Cities	IFD
	Debt Correlation with Neighboring Cities	NDC
	Economic Structure Similarity	ESS
	Inter-city Trade Dependence	ITD

C. Parameter Design

In the empirical analysis of the DCRL–CM model, 16 cities in Anhui Province are selected as case study subjects to construct a dynamic dependency framework based on debt risk, economic environment, and regional contagion indicators. To ensure the model's accuracy and contextual relevance, parameter ranges are carefully designed to reflect the actual economic and fiscal characteristics of these cities, thereby providing realistic initial conditions for model training. Table II presents the parameter configurations employed for the 16 cities, enabling the model to simulate the dynamic characteristics of debt risk contagion across regions with greater realism and fidelity.

TABLE II
PARAMETER DESIGN FOR THE 16 CITIES IN ANHUI PROVINCE WITHIN THE DCRL-CM MODEL

Parameter Category	Indicator Name	Range
Debt Risk Indicators	Debt-to-GDP Ratio	15% - 40%
	Debt Dependency Ratio	50% - 200%
	Debt Interest Burden Ratio	5% - 20%
Economic Environment Indicators	Debt Growth Rate	2% - 15%
	GDP Growth Rate	3% - 10%
	Fiscal Self-sufficiency Ratio	50% - 150%
	Per Capita GDP	40,000 - 120,000
	Fixed Asset Investment Growth Rate	5% - 20%
Regional Contagion Indicators	Inter-city Fiscal Dependence	5% - 15%
	Neighboring Debt Correlation	0.2 - 0.8
	Economic Structure Similarity	0.5 - 1
	Inter-city Trade Dependence	10% - 30%
	Inter-city Fiscal Dependence	5% - 15%
Reinforcement Learning and Copula Parameters	Discount Factor	0.9
	Learning Rate	0.01
	Dynamic Reward Weight	0.5
	Copula Dependency Parameter	0.3

D. Dynamic Evolution of Marginal Distributions and Copula Parameter

This study begins by employing the GARCH (1,1)-t model to estimate the marginal distributions of key city-level debt risk indicators, including debt growth rates and outstanding debt ratios. The GARCH (1,1)-t specification is well suited to capturing the volatility clustering and heavy-tailed behavior frequently observed in financial and fiscal time series. In this framework, the parameters ω, α, and β represent the constant term, the ARCH effect, and the GARCH effect, while ν denotes the degrees of freedom in the student-t distribution, reflecting the severity of tail risk. Model parameters are estimated via maximum likelihood estimation, with model adequacy assessed using standard selection criteria such as the Akaike information criterion (AIC). As shown in Table III, the estimated values of $\alpha + \beta$ approach unity for most cities, highlighting a pronounced persistence in volatility. Furthermore, the majority of cities exhibit significant heavy-tail properties, evidenced by degrees of freedom (ν) generally below 10. Economically, this implies that local governments' debt levels are particularly susceptible to extreme fluctuations in response to fiscal expansions or external shocks, leading to intensified funding needs and pronounced volatility patterns.

Following the estimation of marginal distributions, this study applies the Vine Copula framework to examine the multidimensional dependence structure of inter-city debt risks. To capture the evolving nature of these linkages under shifting macroeconomic conditions and local policy dynamics, time-varying Copula parameters are introduced. Specifically, exogenous macroeconomic variables—such as GDP growth rates and fiscal self-sufficiency levels—are embedded into the correlation coefficients and tail-dependence parameters of the Copula functions, enabling a dynamic representation of the intensity of inter-city risk linkages. Figure 1 and Table IV present average Kendall's τ values for selected city pairs over three periods: 2013–2015, 2016–2018, and 2019–2022. The results reveal a marked increase in correlation for core city pairs, such as Hefei–Wuhu and Hefei–Ma'anshan, particularly from 2016 onward, indicating a significant strengthening of debt-risk interdependence. By contrast, pairs such as Huangshan–Chizhou and Xuancheng–Tongling, which exhibited notable linkages in earlier periods, experienced a substantial decline in correlation following the economic downturn after 2019. This divergence suggests that, amid growing regional economic disparities, cities with weaker economic interdependence or dissimilar industrial structures were less likely to sustain strong debt-risk connections over time.

TABLE IV
GARCH (1,1)-T PARAMETER ESTIMATES FOR THE DEBT GROWTH RATES OF 16 CITIES

City Pair	2013–2015	2016–2018	2019–2022
Hefei – Wuhu	0.39	0.54	0.62
Hefei – Ma'anshan	0.4	0.52	0.59
Hefei – Lu'an	0.34	0.48	0.51
Wuhu – Ma'anshan	0.36	0.47	0.54
Bengbu – Huainan	0.31	0.33	0.35
Fuyang – Bozhou	0.45	0.43	0.38
Huangshan – Chizhou	0.46	0.42	0.36
Xuancheng – Tongling	0.37	0.41	0.43
Anqing – Chizhou	0.41	0.44	0.46
Huaibei – Suzhou	0.3	0.32	0.34

Building on the earlier results from the GARCH (1,1)-t model, the observed variation in volatility dynamics and tail behavior across cities—driven by economic fluctuations and fiscal policy shifts—translates into differentiated intensities of risk linkage within the Copula framework. Under conditions of macroeconomic stability and increasing fiscal interdependence among cities, regional correlation coefficients tend to rise in tandem. Conversely, external shocks that induce divergence in industrial structures or exacerbate imbalances in fiscal burdens lead to weakened linkages among certain city pairs, amplifying temporal "center–periphery" disparities. Taken together, the insights from marginal distribution modeling and dynamic Copula analysis offer a coherent and mutually reinforcing perspective on inter-city debt risk relationships. These findings establish a solid empirical foundation for the subsequent quantification of debt-risk contagion pathways and the design of coordinated response strategies within the MARL framework.

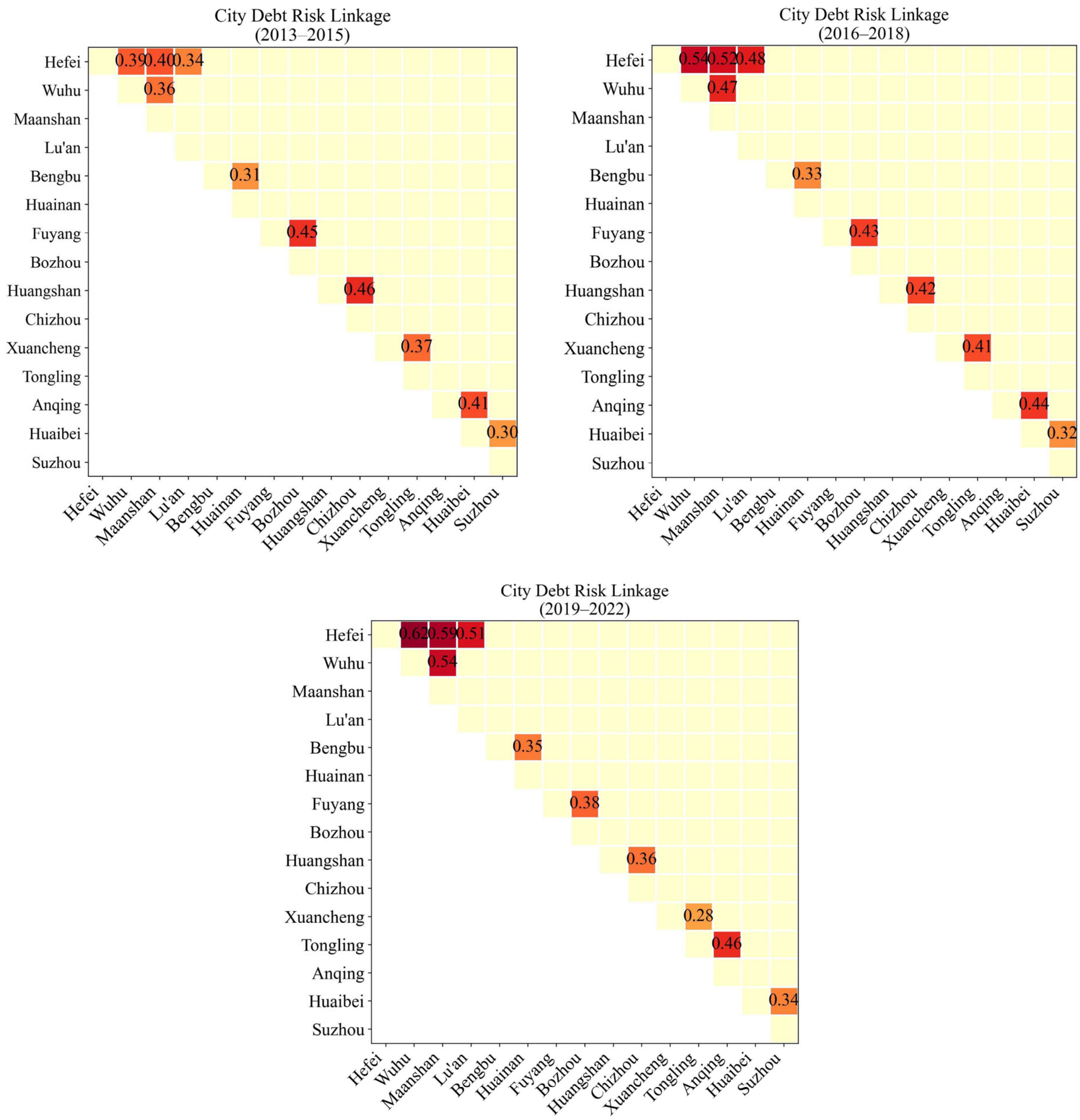

Fig. 1. City debt risk linkage (Kendall's τ) 2019-2022

TABLE III
GARCH (1,1)-T PARAMETER ESTIMATES FOR THE DEBT GROWTH RATES OF 16 CITIES

City	ω (Std. Err.)	α (Std. Err.)	β (Std. Err.)	α+β	ν (Std. Err.)	Log-Likelihood	AIC
1	0.012** (0.005)	0.098** (0.020)	0.842*** (0.030)	0.94	8.37 (1.12)	-426.35	862.71
2	0.009 (0.006)	0.123*** (0.018)	0.795*** (0.025)	0.918	7.24 (1.35)	-439.12	888.25
3	0.015** (0.007)	0.088** (0.021)	0.876*** (0.028)	0.964	9.53* (1.78)	-414.86	843.72
4	0.014** (0.005)	0.142*** (0.017)	0.771*** (0.023)	0.913	6.28** (1.25)	-453.17	924.35
5	0.016** (0.006)	0.115*** (0.019)	0.825** (0.031)	0.94	8.05 (1.47)	-429.62	869.23
6	0.010 (0.004)	0.129** (0.023)	0.802*** (0.029)	0.931	7.52* (1.66)	-432.85	877.7
7	0.018** (0.007)	0.092* (0.049)	0.844*** (0.028)	0.936	6.91** (1.30)	-438.71	891.42
8	0.013** (0.005)	0.101*** (0.015)	0.809*** (0.022)	0.91	9.20* (1.83)	-421.09	856.18
9	0.017** (0.006)	0.080** (0.025)	0.876*** (0.024)	0.956	6.57** (1.12)	-447.44	912.88
10	0.011 (0.007)	0.137*** (0.018)	0.794** (0.034)	0.931	7.03** (1.24)	-450.52	919.03
11	0.008 (0.004)	0.094** (0.031)	0.852*** (0.027)	0.946	8.66 (1.39)	-439.7	889.41
12	0.020** (0.006)	0.079* (0.041)	0.870*** (0.026)	0.949	5.98** (1.09)	-445.89	909.77
13	0.015** (0.006)	0.110** (0.022)	0.822*** (0.025)	0.932	6.24** (1.40)	-434.26	880.53
14	0.021*** (0.007)	0.087** (0.026)	0.815** (0.033)	0.902	9.32 (1.67)	-418.55	853.1
15	0.014* (0.008)	0.125** (0.020)	0.792*** (0.029)	0.917	7.15** (1.56)	-454.31	926.62
16	0.019** (0.006)	0.091** (0.027)	0.834*** (0.030)	0.925	6.02** (1.41)	-441.83	891.6

Notes: Parentheses show standard errors. *p<0.10, **p<0.05, ***p<0.01.

E. Multi-Agent Strategy Convergence and Dynamic Game Characteristics

Building on the previously established marginal distribution and dynamic Copula modeling, this study further conceptualizes the local governments of 16 cities as autonomous agents within a MARL framework, implemented using a CTDE approach. By holding Copula parameters fixed at each stage and iteratively updating decision-making strategies, the model captures the dynamic strategic interactions among cities in relation to debt issuance, risk management, and fiscal balance objectives. The analysis also assesses the convergence properties of these inter-city games and explores the emergence of strategic differentiation across regions once equilibrium is reached. The primary empirical results are presented from three perspectives: strategy convergence dynamics, reward trajectory evolution, and the emergence of core–periphery patterns in policy behavior.

The training process of MARL systems typically involves complex, nonlinear dynamics. During the centralized training phase, agents have access to global state information, including city-level debt indicators, time-varying Copula correlation coefficients, and relevant macroeconomic variables. This shared information environment facilitates coordinated strategy optimization and significantly enhances overall system performance. As shown in Figure 2, the evolution of the average reward across all cities reflects the typical trajectory of agent behavior, progressing from early-stage exploration to eventual strategic convergence.

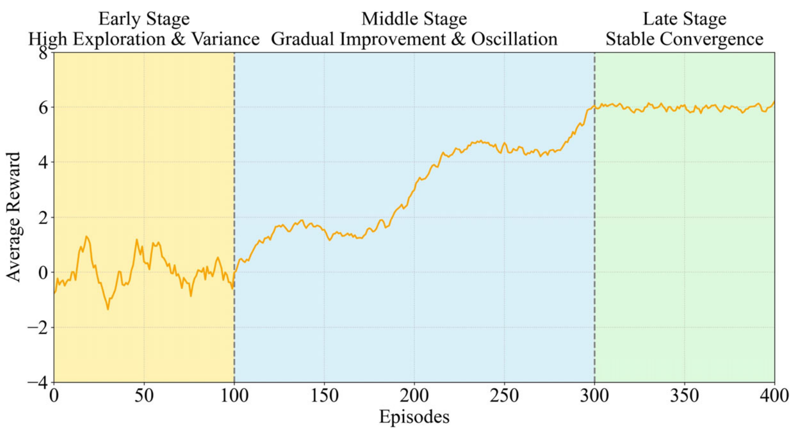

Fig 2. Dynamic evolution of average rewards in multi-Agent reinforcement learning

In the initial training phase (Episodes < 100), agent strategies are underdeveloped and often exhibit biased risk assessments. Some cities pursue aggressive debt expansion to boost short-term fiscal revenues, resulting in high volatility and low average system rewards. During the intermediate phase (Episodes 100–300), ongoing updates to value functions allow agents to refine their strategies, leading to more balanced trade-offs between revenue generation and risk control. This stage is marked by a steady increase in average rewards, albeit with fluctuations, as agents adjust to dynamic interdependencies. In the later phase (Episodes > 300), the reward trajectory stabilizes at a higher level, signaling that agents have converged toward robust local equilibria. At this stage, additional training iterations yield only marginal performance improvements. Following convergence of the training process, cities exhibit marked heterogeneity in their debt issuance behavior, risk management intensity, and fiscal balancing strategies.

Figure 3 presents average annualized debt issuance levels and corresponding risk management weights under converged strategies. Core cities (e.g., Hefei, Wuhu, Ma'anshan) typically adopt moderate borrowing (~4–5%) combined with robust risk management (>0.60), achieving superior performance (>0.70). In contrast, peripheral cities (e.g., Bozhou, Fuyang, Suzhou), facing greater fiscal pressures, opt for aggressive borrowing (~6%) with relatively weaker risk controls, resulting in comparatively lower returns. This divergence aligns with dynamic Copula dependency findings: core cities, with stronger interdependencies and higher sensitivity to external shocks, learn to adopt cautious fiscal strategies within the MARL framework, whereas peripheral cities prioritize immediate fiscal needs, accepting elevated risk.

Observed strategic convergence highlights a critical trade-off between fiscal expansion and risk mitigation among regional governments. Core cities, notably Hefei, centrally positioned within the debt-risk transmission network (Figure 4), effectively balance growth with stability, exerting positive spillovers on neighboring cities. In contrast, peripheral cities such as Fuyang and Bozhou typically adopt reactive fiscal behaviors, increasing exposure to tail risks during adverse conditions and reinforcing structural asymmetries. This pattern aligns with Copula analysis results, underscoring how external shocks (e.g., interest rate hikes) disproportionately impact peripheral cities, thereby undermining their fiscal sustainability.

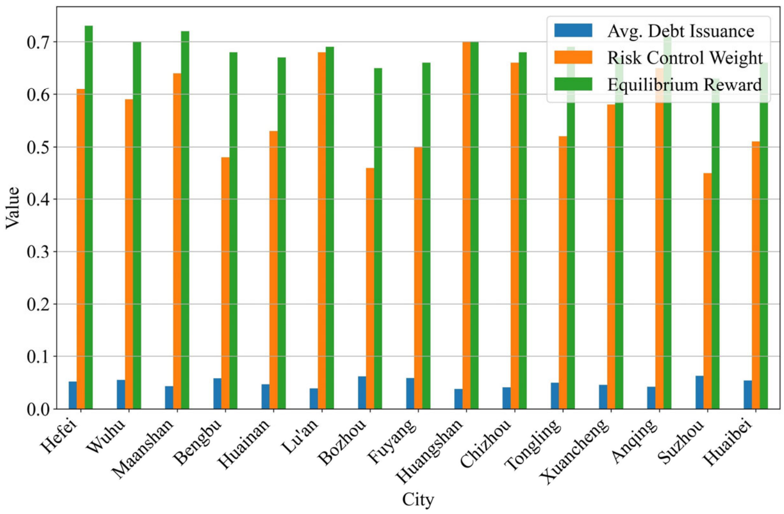

Fig 3. Key strategy metrics and equilibrium reward by city

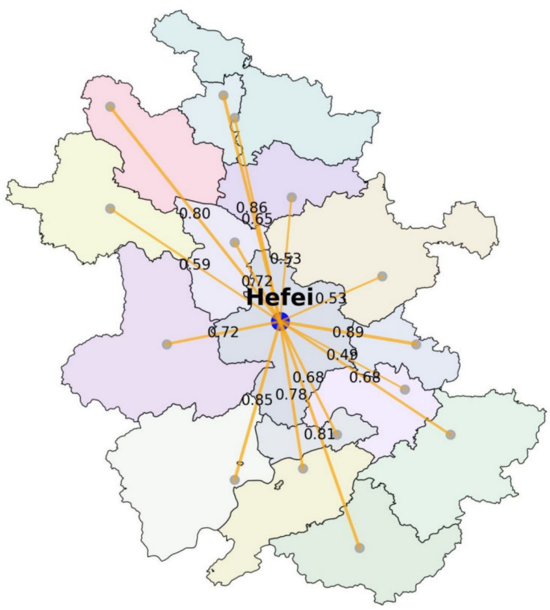

Fig 4. Regional fiscal risk network in Anhui province

The tail-risk heatmap derived from the Copula analysis in Figure 5 further validates the aforementioned findings. As clearly illustrated, Hefei exhibits notably strong tail-risk correlations with its surrounding cities, underscoring its central role in regional risk transmission. Conversely, Figure 5 also highlights relatively weaker tail-risk associations among peripheral cities, indicating that these cities tend to become isolated risk clusters when facing external shocks, thereby limiting their ability to achieve effective collaborative risk management.

F. Debt Risk Spillover Pathways and Contagion Patterns

The preceding network topology analysis underscores the structural vulnerabilities arising from aggressive debt strategies adopted by peripheral cities under fiscal constraints, revealing divergent pathways of risk accumulation under different policy regimes. Figure 6 further illustrates how policy interventions shape the dynamic evolution of inter-city risk correlations, as captured by the trends in Copula coefficients. Specifically, under stringent debt-control policies, the overall level of risk correlation increases steadily, suggesting that rigid fiscal constraints—though designed to curb systemic risk—may inadvertently intensify risk clustering due to insufficient external liquidity. In contrast, when central fiscal authorities enhance transfer payments to peripheral cities, a pronounced decline in risk correlation is observed. This pattern indicates that improved liquidity support through central transfers can effectively mitigate the accumulation and transmission of systemic risk. These findings demonstrate the stabilizing role of central government transfer policies in dampening contagion intensity across the regional debt-risk network, thereby validating the conclusions drawn from both the MARL simulations and the structural risk-network analyses.

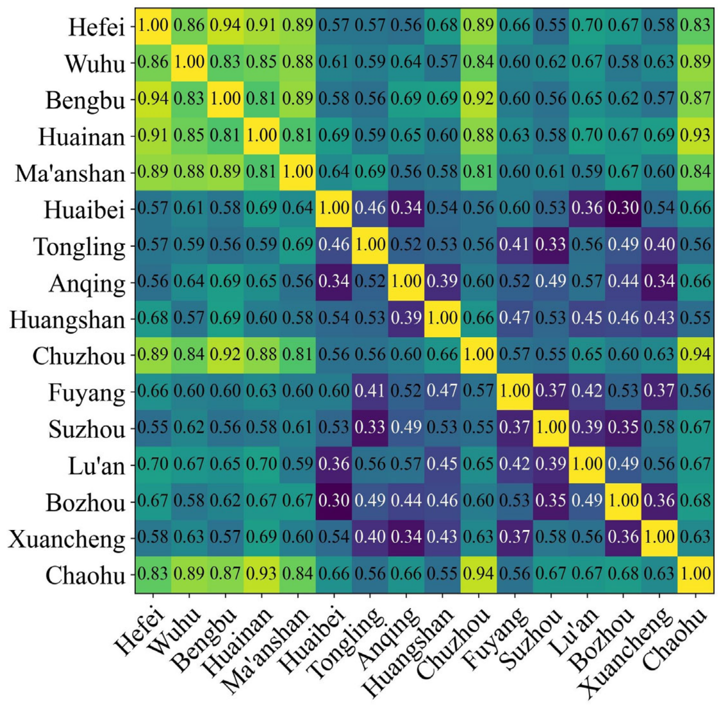

Fig 5. Copula tail-risk heatmap among Anhui province Cities

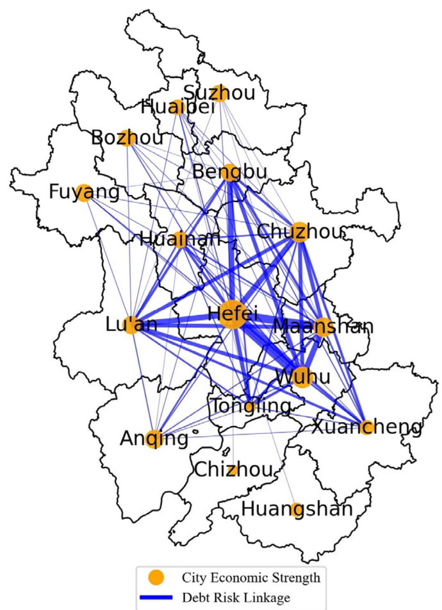

Fig 6. Network graph of debt risk linkages among cities in Anhui province

Figure 7 maps debt-risk diffusion across 16 cities under escalating shocks, corroborating insights from marginal distributions, dynamic Copulas, and MARL simulations. Peripheral cities, hampered by limited fiscal capacity, persistently issue more debt (darker nodes), securing only transient gains when conditions are benign. As external shocks intensify, default pressure rises and risk rapidly cascades along high-risk corridors—especially among Lu'an, Suzhou, and Bozhou—forming tight vulnerability clusters. Even where some cities adopt prudence, highly leveraged, densely connected nodes remain pivotal risk hubs. The pattern exposes the structural fragility of fiscally constrained peripheral clusters and the enduring challenge of systemic-risk control in a decentralized fiscal regime.

Network topology analysis highlights structural vulnerabilities stemming from aggressive borrowing by fiscally constrained peripheral cities, illustrating divergent risk trajectories under varying policy regimes. Figure 8 demonstrates that strict debt-control policies, despite aiming to reduce systemic risk, inadvertently heighten inter-city risk correlation by limiting external liquidity. Conversely, increased central government transfer payments to peripheral cities effectively reduce risk correlations, mitigating systemic risk accumulation and diffusion. Thus, central fiscal transfers play a stabilizing role, moderating risk transmission and reducing contagion intensity across the regional debt network, consistent with findings from MARL simulations and structural network analyses.

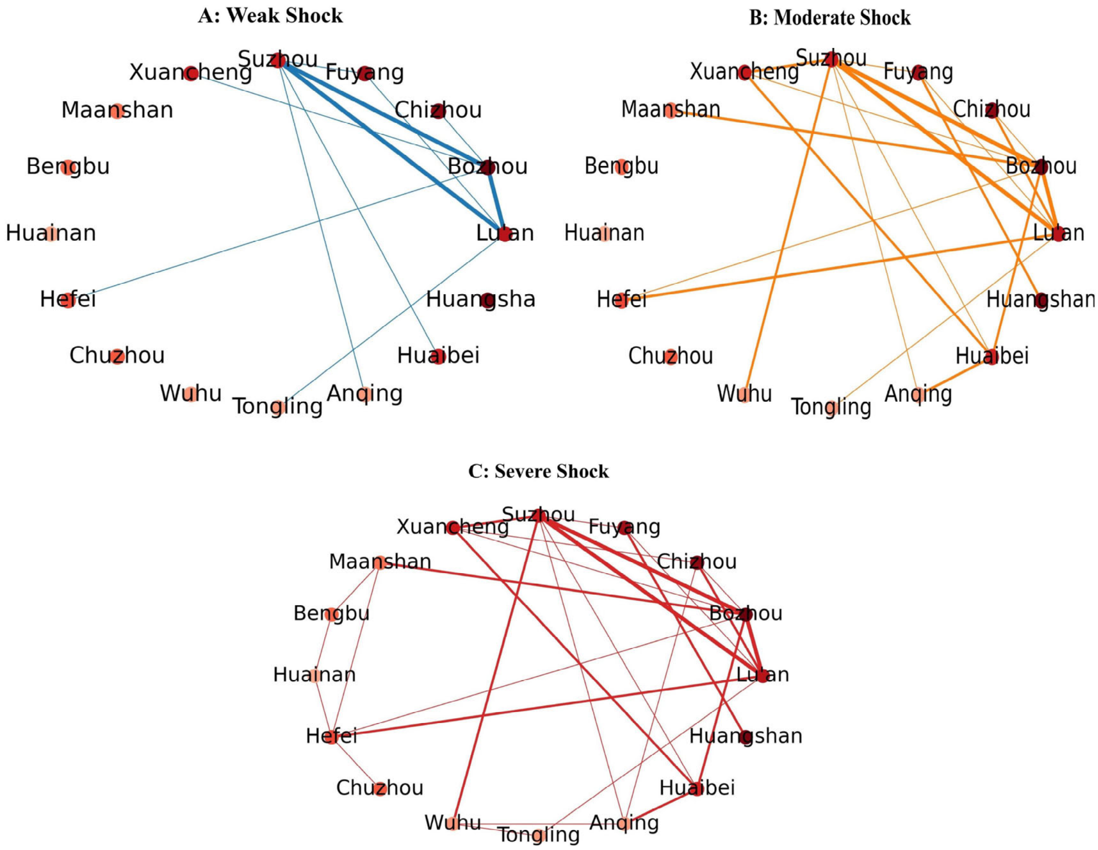

Fig 7. Network contagion and evolutionary dynamics of urban debt risks under different economic scenarios

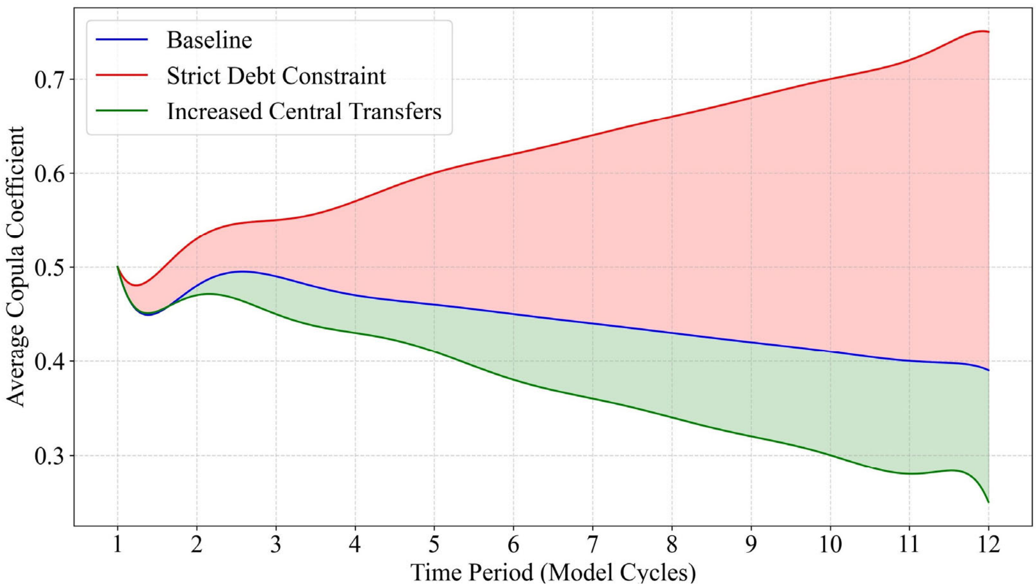

Fig 8. Trend of Copula mean coefficients under different policies

The contagion patterns described above indicate that core cities do not invariably function as stabilizers. When confronted with adverse external conditions or surging local debt levels, they may propagate tail risks outward to peripheral regions. Meanwhile, peripheral cities, constrained by limited fiscal resources and industrial diversity, are more susceptible to cascading reactions at particular junctures or through specific transmission channels. If policy interventions strategically enhance fiscal support and optimize debt-issuance regulations in peripheral regions, it could partially reshape the inherent "core-periphery" risk structure of the existing network, leading to a more decentralized and resilient inter-city linkage. Nevertheless, uneven development across cities may still concentrate vulnerabilities at critical nodes during shocks, posing ongoing challenges for regional macroprudential oversight and cross-city fiscal coordination.

G. Robustness Tests

To further validate the robustness of the model in capturing risk dependence structures, we examined and compared the performances of three distinct Copula models: Gaussian Copula, t-Copula, and Vine Copula. Figure 9 illustrates the visualization results of dependency structures under these different Copula specifications. As depicted in the Gaussian Copula, the correlations among variables appear relatively weak, with data points dispersedly distributed and notably lacking significant clustering in the tail regions. This highlights the Gaussian Copula's limitations in capturing extreme events or tail-risk contagion. In contrast, the t-Copula model exhibits pronounced heavy-tail characteristics and strong tail dependencies, evidenced by substantial data clustering in extreme regions. This trait demonstrates the t-Copula's greater suitability for modeling extreme events commonly observed in financial market risk transmission. However, considering the complexity and potential stage-specific characteristics of real-world risk contagion mechanisms, we further employed the Vine Copula model to capture more detailed dynamic variations in risk dependencies.

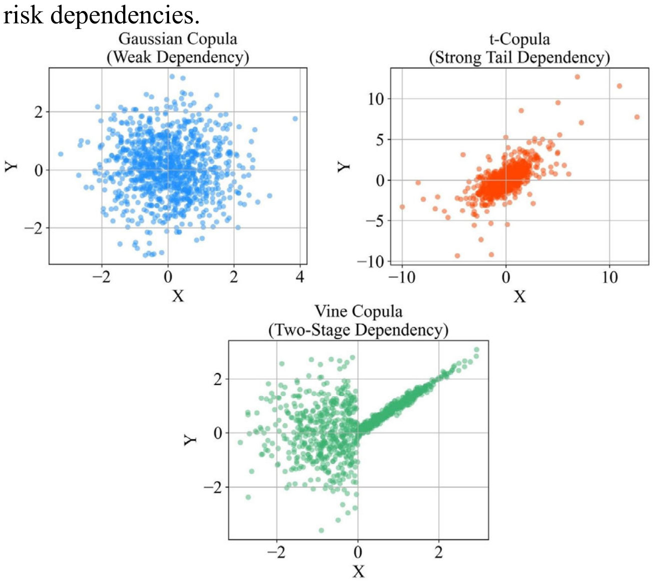

Fig 9. Comparison of different Copula models

The results in Figure 9 reveal a clear two-stage dependence structure: in the initial stage, data points show weak and dispersed dependencies, whereas after crossing a certain risk threshold, they exhibit distinctly strong and concentrated dependencies. This two-phase characteristic enables the Vine Copula model to effectively differentiate between the initial propagation and outbreak stages of regional government debt risks. Thus, the robustness analysis presented above indicates that the Vine Copula model employed in this study demonstrates robust performance and high applicability, accurately capturing the complex dependency structures and dynamic tail-risk characteristics inherent in the regional debt-risk contagion process.

Following the comparative analysis of dependency structures across various Copula models discussed earlier, this study next evaluates the robustness of marginal distributions. This step aims to confirm that the overall model consistently captures how debt risk dynamically evolves under different volatility assumptions. Figure 10 shows volatility sequences produced by three commonly used volatility models: GARCH (1,1)-t, EGARCH (1,1)-t, and GJR-GARCH (1,1)-t. Specifically, the GARCH (1,1)-t model (top of Figure 10, highlighted in blue) displays symmetric and smooth fluctuations, suggesting that the volatility response to positive and negative market shocks remains relatively balanced. In contrast, the EGARCH (1,1)-t model (middle of Figure 10, represented by the red curve) exhibits noticeable asymmetry, characterized by sustained increases in volatility following negative market shocks. The GJR-GARCH (1,1)-t model (bottom of Figure 10, shaded green) typically shows abrupt jumps in volatility in response to negative returns, clearly reflecting the leverage effect, which amplifies the impact of negative shocks. By comparing these results with Figure 10, we observe that although each volatility model reacts differently to negative shocks, the general cyclical patterns and tail-risk behavior are consistent with the contagion paths previously identified using Copula models. This indicates that substituting the GARCH (1,1)-t with either EGARCH (1,1)-t or GJR-GARCH (1,1)-t does not significantly alter the main findings. Thus, the study's key conclusions regarding the dynamic evolution and tail-risk amplification of regional government debt risk remain robust, irrespective of whether symmetric or asymmetric volatility assumptions are applied. After verifying the robustness of the Copula dependence structure and the GARCH-based marginal distributions, this study further investigates the sensitivity of the MARL model parameters to evaluate its robustness and applicability under various strategy settings. Specifically, this paper adjusted the discount factor, learning rate, and dynamic reward weights between returns and risks within the reward function, subsequently observing changes in cities' debt management strategies, risk contagion pathways, and systemic risk levels.

As indicated in Figure 11, increasing the discount factor from 0.9 to 0.99 significantly shifts agents' preferences toward long-term gains, prompting them to adopt strategies that reduce short-term debt expansion while enhancing risk controls. However, excessively high learning rates may induce substantial strategic fluctuations or impede convergence, negatively affecting the model's predictive accuracy. Additionally, varying the dynamic weights of risk and return in the reward function considerably impacts how cities balance fiscal revenues against debt risk management. Overall, despite specific model parameters influencing detailed characteristics of debt management strategies and risk contagion dynamics, the overarching risk hierarchy and contagion pathways remain stable. This outcome demonstrates that the MARL model constructed in this study maintains robust performance and adaptability across diverse reinforcement learning parameter configurations.

To thoroughly evaluate the advantages of the DCRL-CM model in controlling debt risk contagion pathways and managing risks, this study conducts a comparative analysis against the static Copula model and the traditional regression model. In managing debt risk, the DCRL-CM model demonstrates a clear advantage in controlling cross-regional contagion effects. Figure 12 presents the contagion intensity distribution of debt risks across 16 cities for the three models: DCRL-CM, static Copula, and traditional regression. Leveraging a multi-agent reinforcement learning mechanism, the DCRL-CM model dynamically adjusts its management strategies based on the risk levels in different regions. As shown in Figure 12, the overall contagion intensity for the DCRL-CM model is lower, especially in high-risk areas where it effectively mitigates contagion effects. This balanced distribution of contagion intensity across cities reflects the model's stability and adaptability in multi-regional settings. In contrast, the static Copula model, with its fixed dependency structure, lacks the ability to adapt to contagion effects in real-time under dynamic conditions. Compared to the DCRL-CM model, the contagion intensity is higher, with certain regions displaying a marked increase in contagion intensity. The traditional regression model, which relies on historical data for linear regression analysis, exhibits the highest contagion intensity among the three models, showing stronger cross-regional contagion effects along major contagion pathways.

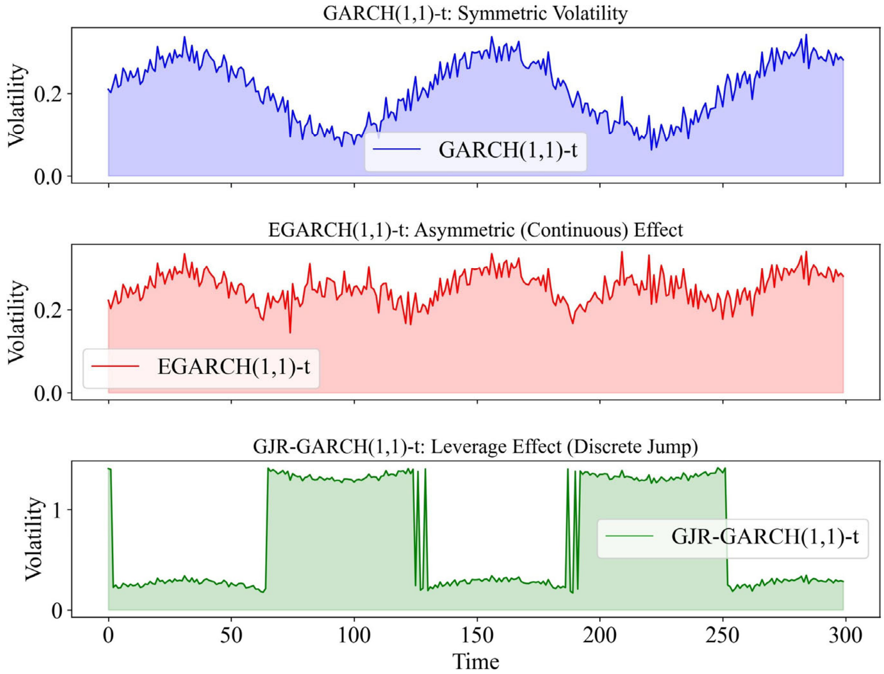

Fig 10. Robustness check: distinct marginal volatility patterns

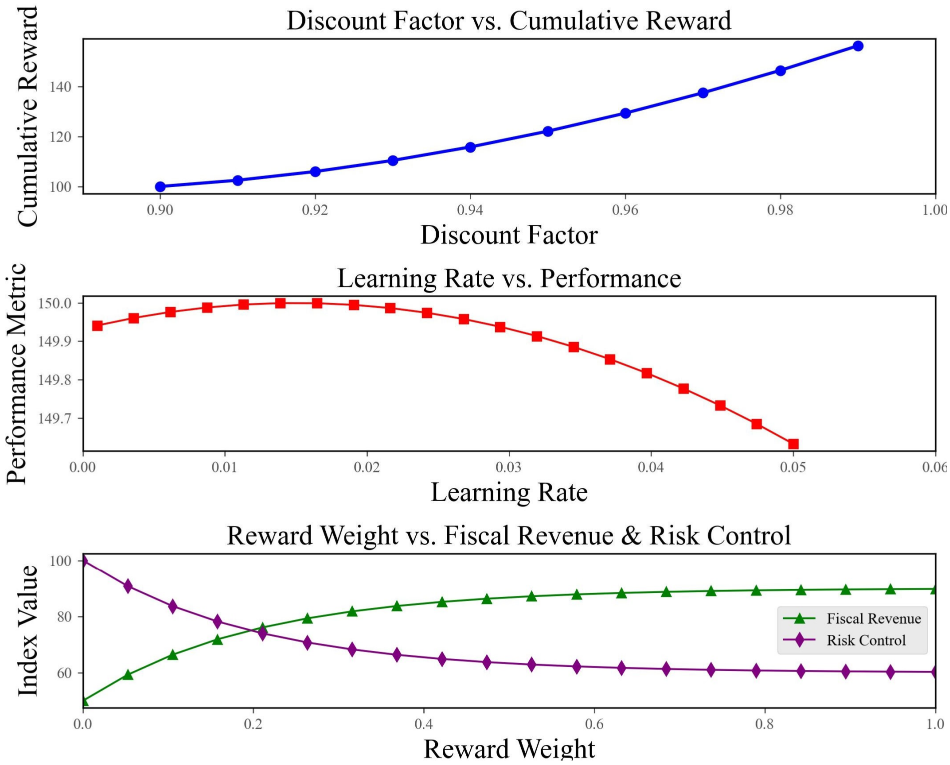

Fig 11. Parameter sensitivity analysis in reinforcement learning

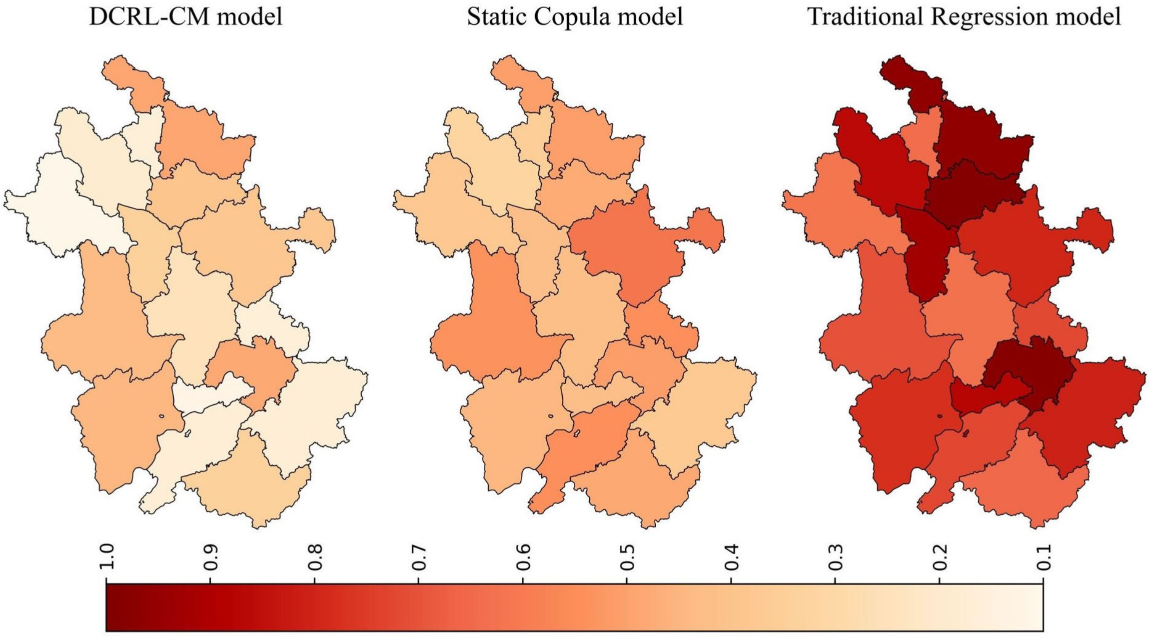

Fig 12. Contagion intensity distribution across 16 cities for DCRL-CM, static Copula, and traditional regression models

IV. DISCUSSION

The study results indicate that the DCRL－CM model, by integrating dynamic Copula with multi-agent reinforcement learning, effectively captures the evolving contagion pathways and intensities of government debt risk. This integration significantly enhances adaptability in modeling interregional debt risk dependencies and improves risk control outcomes. Specifically, the model not only demonstrates stronger risk mitigation in high-risk areas but also achieves decentralized risk management across regions through coordinated optimization among multiple agents. Compared to traditional static Copula and regression models, the DCRL-CM model offers greater flexibility in capturing cross-regional dynamic risk, particularly showing a robust real-time adjustment capacity under changing policies and market fluctuations. This innovation holds substantial application value in the fields of debt contagion risk warning and control. Additionally, the DCRL-CM model excels in predictive accuracy, with the lowest cumulative prediction error over multiple iterations, highlighting that its dynamic dependency structure and reinforcement learning mechanism capture time-varying characteristics of debt risk, slow the accumulation of error, and significantly improve the reliability of risk forecasting. This feature provides policymakers with a crucial risk monitoring tool in complex, multi-regional debt environments.

Consistent with Luo et al.[31], this study finds that dynamic Copula models effectively identify nonlinear interregional dependencies, particularly exhibiting high sensitivity to contagion risk during economic crises or policy shifts. Unlike Pham et al.'s analysis[32] focused on tail dependence, this study extends beyond tail dependence by simulating dynamic contagion pathways and using reinforcement learning to dynamically optimize contagion mechanisms, offering a more comprehensive understanding of debt risk propagation. Similar to Lee et al.'s findings [33] on the East Asian sovereign debt market, this study identifies dual contagion effects within and across regions, indicating that risk can spread both within cities and across city boundaries. However, through the application of a multi-agent coordination mechanism, the study achieves dynamic adjustments in cross-regional risk control. Furthermore, unlike Huynh et al.'s work [34] on bank system contagion, this study's DCRL-CM model uses reinforcement learning to enhance strategic flexibility in debt management, extending risk control applicability beyond local banking systems to the broader context of regional government debt environments.

While the DCRL-CM model offers certain advantages in capturing and controlling dynamic contagion of risk, there are some limitations. For instance, in adjusting dependency parameters dynamically, the model heavily relies on data, and its computational complexity can pose challenges in practical applications due to limitations in data availability and computational resources. Additionally, sample size limitations may affect the model's generalizability, suggesting that future research could expand the sample scope to verify the model's applicability across different regions. Future studies might also focus on simplifying the model's computational complexity or incorporating more external environmental variables to enhance robustness. For cities with low-risk control effectiveness, such as Huainan, further parameter optimization is needed to improve the model's risk management capabilities.

V. CONCLUSION

A. Research Conclusions

This study proposes and validates a Dynamic Copula and Reinforcement Learning-based Contagion Model (DCRL-CM), which integrates a time-varying Copula approach with multi-agent reinforcement learning (MARL) to explore the contagion paths and dynamic characteristics of regional government debt risks. The findings demonstrate the model's significant advantages in capturing the dynamic dependency structure of multi-regional debt risks and effectively controlling cross-regional risks, as reflected in the following aspects:

(1) The time-varying Copula method effectively captures the nonlinear dynamic dependency characteristics of regional debt risks under policy adjustments and external shocks. Particularly during economic downturns or periods of

heightened fiscal pressure, the model exhibits high sensitivity to tail risks. The results quantitatively reveal the amplifying effects of macroeconomic and fiscal conditions on the coupling of regional debt risks;

(2) The multi-agent reinforcement learning framework enables dynamic coordination and optimization of fiscal and risk control strategies across regions. In the context of high-risk events, city-level agents can adjust debt issuance and risk management strategies based on reinforcement learning feedback mechanisms, dynamically achieving a balance between returns and risks. This demonstrates the model's adaptability and flexibility in multi-regional dynamic environments;

(3) Compared to traditional static Copula and regression models, the DCRL-CM model exhibits significantly stronger real-time adjustment capabilities under dynamic conditions, enhancing its effectiveness in capturing and managing evolving debt risk interdependencies.

B. Policy Recommendations

Based on the study's conclusions, policy recommendations are proposed across three dimensions-fiscal management, economic structure, and risk monitoring-to mitigate the risk of debt contagion among regional governments:

(1) Regional governments should establish a fiscal coordination mechanism and emergency plans to enhance information sharing and coordinated responses. Setting up an inter-regional fiscal coordination committee with unified emergency protocols and information - sharing standards can ensure a swift and collaborative approach, with regular communication and training sessions. A permanent risk monitoring and emergency drill system should be established to simulate economic shocks, test emergency response capabilities, and continually optimize plans. Developing a fiscal information - sharing platform for real-time access to fiscal data will strengthen overall response readiness. Diversifying debt instruments and financing channels can reduce reliance on single sources and mitigate external shock impacts. Using adaptive models, such as the DCRL-CM, to dynamically adjust debt management strategies can allow flexibility in responding to economic changes;

(2) High interdependence between regional economies can exacerbate debt risk contagion. Promoting economic diversification, particularly in cities heavily dependent on specific industries, can alleviate systemic impacts of localized economic downturns, thus effectively reducing debt risk transmission. For instance, in cities with a strong reliance on manufacturing, introducing high-value technology sectors and supporting small and medium-sized enterprises (SMEs) can enhance economic resilience and diversity. Supporting local business innovation and attracting investment, especially in high - growth industries, can also foster economic diversification, lessening the adverse effects of economic volatility. By promoting industrial diversification, regional economic stability is bolstered, reducing the negative impacts of fluctuations. Additionally, strengthening collaboration and complementarity between industries across regions can reduce dependence on single sectors, increasing regional resilience and mitigating the systemic impact of debt risk;

(3) Key contagion nodes in debt risk should be closely monitored, with early warning systems in place to detect risk signals promptly and prevent further spread. Leveraging big data and artificial intelligence, an intelligent debt risk monitoring platform could be established to continuously assess the fiscal conditions of regional governments in real-time, improving data collection and analysis efficiency, and enabling faster identification of potential risks with targeted interventions. A regional data warehouse could standardize and share fiscal data, using machine learning to automatically detect risk patterns and generate alerts. Artificial intelligence can conduct multi-dimensional data correlation analysis, predicting potential contagion pathways and offering decision support. An integrated smart monitoring platform linked to local fiscal systems could enable real-time updates and automated risk assessments. Furthermore, enhancing inter-regional risk information sharing mechanisms ensures that cities can quickly access relevant debt risk information, creating a coordinated response mechanism to curb debt risk contagion in its early stages.

C. Research Limitations and Future Directions

While integrating dynamic Copula with multi-agent reinforcement learning (MARL) improves the model's adaptability to dynamic environments, it also significantly increases computational complexity. This is particularly evident when high-dimensional multi-regional data is introduced, as the model's training time and resource requirements may pose challenges for real-time applications in resource-constrained settings. To address these limitations, future research could explore more efficient MARL algorithms or adopt distributed computing frameworks to reduce computational complexity. For instance, employing graph neural networks (GNNs) could enhance the efficiency of modeling inter-regional dependency networks. GNNs are well-suited for capturing the relational structure among regions, potentially reducing the computational burden while maintaining high accuracy.

Additionally, this study primarily considers macroeconomic variables such as GDP growth rates and fiscal self-sufficiency as external shock factors. However, other dimensions—such as environmental, social, and political variables—are not fully accounted for, which may constrain the model's robustness in addressing complex external environmental changes. Future research should incorporate variables from these dimensions into the analytical framework to better capture a broader range of external shock factors. This would enhance the model's comprehensiveness and applicability in diverse and evolving contexts.

REFERENCES

[1] M. Elliott, B. Golub, and M. Jackson, "Financial Networks and Contagi on," *American Economic Review,* vol. 104, no. 10, pp. 3115-53, 2014. [Online]. Available: https://EconPapers.repec.org/RePEc:aea:aecrev:v:104:y:2014:i:10:p:3115-53.

[2] X. Yang, X. Wang, J. Cao, L. Song, and C. Huang, "Can local govern ment implicit debt raise regional financial market spillover? Evidence f rom China," *Finance Research Letters,* vol. 67, p. 105873, 2024/09/01/ 2024, doi: https://doi.org/10.1016/j.frl.2024.105873.

[3] S. Corbet, Y. Hou, Y. Hu, L. Oxley, and D. Xu, "Pandemic-related fina ncial market volatility spillovers: Evidence from the Chinese COVID-1 9 epicentre," *International Review of Economics & Finance,* vol. 71, p p. 55-81, 2021/01/01/ 2021, doi: https://doi.org/10.1016/j.iref.2020.06.022.

[4] A. J. Makin and A. Layton, "The global fiscal response to COVID-19: Risks and repercussions," *Economic Analysis and Policy,* vol. 69, pp. 3 40-349, 2021.

[5] Y. Ma and L. Lv, "Money, debt, and the effects of fiscal stimulus," *Eco nomic Analysis and Policy,* vol. 73, pp. 152-178, 2022.

[6] X. Geng and M. Qian, "Understanding the local government debt in China," *Pacific-Basin Finance Journal,* vol. 86, p. 102456, 2024.

[7] C. Jiang, Y. Li, Q. Xu, and Y. Liu, "Measuring risk spillovers from multiple developed stock markets to China: A vine-copula-GARCH-MIDAS model," *International Review of Economics & Finance,* vol. 75, pp. 386-398, 09/01 2021, doi: 10.1016/j.iref.2021.04.024.

[8] B. S. Atasoy, İ. Özkan, and L. Erden, "The determinants of systemic risk contagion," *Economic Modelling,* vol. 130, p. 106596, 2024/01/01/ 2024, doi: https://doi.org/10.1016/j.econmod.2023.106596.

[9] A. Ahdika, D. Rosadi, A. R. Effendie, and G. Gunardi, "Measuring Dynamic Dependency using Time-Varying Copulas with Extended Parameters: Evidence from Exchange Rates Data," *MethodsX,* vol. 8, 101322, 03/26 2021, doi: 10.1016/j.mex.2021.101322.

[10] P. Krupskii and H. Joe, "Flexible copula models with dynamic dependence and application to financial data," *Econometrics and Statistics,* vol. 16, pp. 148-167, 2020/10/01/ 2020, doi: https://doi.org/10.1016/j.ecosta.2020.01.005.

[11] C. Luo, C. Xie, Y. Cong, and X. Yan, "Measuring financial market risk contagion using dynamic MRS-Copula models: The case of Chinese and other international stock markets," *Economic Modelling,* vol. 51, pp. 657-671, 12/01 2015, doi: 10.1016/j.econmod.2015.09.021.

[12] Y. Huang, C. Zhou, K. Cui, and X. Lu, "A multi-agent reinforcement learning framework for optimizing financial trading strategies based on timesnet," *Expert Systems with Applications,* vol. 237, p. 121502, 2024.

[13] N. Pippas, C. Turkay, and E. A. Ludvig, "The Evolution of Reinforcement Learning in Quantitative Finance," *arXiv preprint arXiv:2408.10932,* 2024.

[14] Z. Li, V. Tam, and K. L. Yeung, "Developing a multi-agent and self-adaptive framework with deep reinforcement learning for dynamic portfolio risk management," *arXiv preprint arXiv:2402.00515,* 2024.

[15] M. Hernes, J. Korczak, D. Krol, M. Pondel, and J. Becker, "Multi-agent platform to support trading decisions in the FOREX market," *Applied Intelligence,* vol. 54, no. 22, pp. 11690-11708, 2024/11/01 2024, doi: 10.1007/s10489-024-05770-x.

[16] J. Lussange, I. Lazarevich, S. Bourgeois-Gironde, S. Palminteri, and B. Gutkin, "Modelling Stock Markets by Multi-agent Reinforcement Learning," *Computational Economics,* vol. 57, no. 1, pp. 113-147, 2021/01/01 2021, doi: 10.1007/s10614-020-10038-w.

[17] S. Gronauer and K. Diepold, "Multi-agent deep reinforcement learning: a survey," *Artificial Intelligence Review,* vol. 55, no. 2, pp. 895-943, 2022/02/01 2022, doi: 10.1007/s10462-021-09996-w.

[18] Y. Wang and H. Pham, "Modeling the Dependent Competing Risks With Multiple Degradation Processes and Random Shock Using Time-Varying Copulas," *IEEE Transactions on Reliability,* vol. 61, no. 1, pp. 13-22, 2012, doi: 10.1109/TR.2011.2170253.

[19] G. Papoudakis, F. Christianos, L. Schäfer, and S. V. Albrecht, "Comparative Evaluation of Cooperative Multi-Agent Deep Reinforcement Learning Algorithms," 2021.

[20] H. Xuan, L. Maestrini, F. Chen, and C. Grazian, "Stochastic variational inference for GARCH models," *Statistics and Computing,* vol. 34, no. 1, p. 45, 2024.

[21] C. Czado and T. Nagler, "Vine copula based modeling," *Annual Review of Statistics and Its Application,* vol. 9, no. 1, pp. 453-477, 2022.

[22] T. Li *et al.*, "Applications of multi-agent reinforcement learning in future internet: A comprehensive survey," *IEEE Communications Surveys & Tutorials,* vol. 24, no. 2, pp. 1240-1279, 2022.

[23] Y. Han, P. Gong, and X. Zhou, "Correlations and risk contagion between mixed assets and mixed-asset portfolio VaR measurements in a dynamic view: An application based on time varying copula models," *Physica A: Statistical Mechanics and its Applications,* vol. 444, pp. 940-953, 2016/02/15/ 2016, doi: https://doi.org/10.1016/j.physa.2015.10.088.

[24] F. Broucke and T. Hilberdink, "A Mean value Theorem for general Dirichlet series," *The Quarterly Journal of Mathematics,* vol. 75, no. 4, pp. 1393-1413, 2024.

[25] H. Gouk, E. Frank, B. Pfahringer, and M. J. Cree, "Regularisation of neural networks by enforcing lipschitz continuity," *Machine Learning,* vol. 110, pp. 393-416, 2021.

[26] J. Du *et al.*, "MADDPG-based joint service placement and task offloading in MEC empowered air–ground integrated networks," *IEEE Internet of Things Journal,* vol. 11, no. 6, pp. 10600-10615, 2023.

[27] M. Zhang, W. Tong, G. Zhu, X. Xu, and E. Q. Wu, "SQIX: QMIX algorithm activated by general softmax operator for cooperative multiagent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* 2024.

[28] M. Sayin, K. Zhang, D. Leslie, T. Basar, and A. Ozdaglar, "Decentralized Q-learning in zero-sum Markov games," *Advances in Neural Information Processing Systems,* vol. 34, pp. 18320-18334, 2021.

[29] J. Liu and G. Yu, "Fuzzy Kakutani-Fan-Glicksberg fixed point theorem and existence of Nash equilibria for fuzzy games," *Fuzzy Sets and Systems,* vol. 447, pp. 100-112, 2022.

[30] C. Nwaigwe and D. N. Benedict, "Generalized Banach fixed-point theorem and numerical discretization for nonlinear Volterra–Fredholm equations," *Journal of Computational and Applied Mathematics,* vol. 425, p. 115019, 2023.

[31] L. Changqing, X. Chi, Y. Cong, and X. Yan, "Measuring financial market risk contagion using dynamic MRS-Copula models: The case of Chinese and other international stock markets," *Economic Modelling,* vol. 51, pp. 657-671, 2015/12/01/ 2015, doi: https://doi.org/10.1016/j.econmod.2015.09.021.

[32] T. N. Pham, R. Powell, and D. Bannigidadmath, "Tail risk network analysis of Asian banks," *Global Finance Journal,* vol. 62, p. 101017, 2024/09/01/ 2024, doi: https://doi.org/10.1016/j.gfj.2024.101017.

[33] Y. Lee, K. Hong, and K. Yang, "Sovereign Risk Contagion in East Asia: A Mixture of Time-Varying Copulas Approach," *Emerging Markets Finance and Trade,* vol. 54, no. 7, pp. 1513-1537, 2018/05/28 2018, doi: 10.1080/1540496X.2018.1445989.

[34] T. L. D. Huynh, M. A. Nasir, S. P. Nguyen, and D. Duong, "An assessment of contagion risks in the banking system using non-parametric and Copula approaches," *Economic Analysis and Policy,* vol. 65, pp. 105-116, 2020/03/01/ 2020, doi: https://doi.org/10.1016/j.eap.2019.11.007.