

Article

Mechanism-Embedded Machine Learning-Driven Resilience Governance of Digital–Intelligent Innovation Collaboration Networks

Jingran Xing, Hua Zou * and Gupeng Zhang

School of Management, Shenyang University of Technology, Shenyang 110870, China

* Correspondence: 19708688926@163.com

Highlights

Please indicate how your work links to systems science via your contributions to systems practice, theory, and/or methodology.

- Develops a staged mechanism-embedded machine learning framework linking relationship processes, latent-state transitions, collaboration resilience, and sequential governance.
- Freezes the empirically validated relationship and state-transition mechanisms before reinforcement learning, enabling traceable governance optimization under resource constraints.

What are the main findings and/or the implications of the main findings?

- Six state-transition associations remain robust after a global BH-FDR correction, showing that collaboration conditions operate in a path-specific rather than uniformly positive manner.
- OSR-DDQN improves the mean reward by 11.56–12.83% over the best-performing standard deep Q-network while reducing budget utilization by 5.08–6.46 percentage points.

Abstract

The existing research on digital–intelligent innovation collaboration networks has not effectively connected the evolution of real collaborative relationships with governance methods, and governance models often struggle to accommodate both empirically identified evolutionary mechanisms and sequential decision-making capability. Using Chinese digital–intelligent innovation co-patent data from 2016 to 2025, this study develops a machine learning analytical framework centered on relationship evolution and collaboration resilience governance and employs OSR-DDQN to optimize dynamic intervention strategies under limited budgets. The results show that state transitions in digital–intelligent innovation collaboration exhibit pronounced path specificity, with the innovation capability, relational capital, and technological complementarity displaying differentiated conditional associations across transition paths. Further simulation experiments show that, across different budget conditions, OSR-DDQN improves the mean reward by 11.56% to 12.83% relative to the best-performing standard deep Q-network. Within the specified governance environment, once the budget reaches a minimally sufficient level, further performance gains arise primarily from optimizing the action timing, intervention targets, and stopping decisions rather than from simply increasing governance resources.



Academic Editor: Mitsuru Kodama

Received: 12 August 2026

Revised: 8 September 2026

Accepted: 9 September 2026

Published: 11 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: digital–intelligent innovation collaboration network; collaboration resilience; mechanism-embedded machine learning; deep reinforcement learning; adaptive governance

1. Introduction

The convergence of digital and intelligent technologies is reshaping the organizational boundaries of innovation activities [1,2]. Complex technological problems typically involve multiple types of knowledge, equipment, data, and application scenarios, making it increasingly difficult for a single actor to rely solely on internal resources to move from knowledge discovery to the translation of research outcomes. Firms, universities and research institutions, and public sector actors therefore form interorganizational innovation networks through joint R&D, knowledge exchange, and resource coordination [3]. Collaboration networks can broaden the scope of knowledge search and facilitate the combination of complementary resources, but they also introduce new governance problems involving partner identification, relationship coordination, knowledge integration, and structural dependence. An increase in the number of network ties does not necessarily produce a simultaneous improvement in collaboration quality and system resilience. New relationships may develop into stable collaboration through sustained interaction, but they may also terminate rapidly because of poor actor matching, rising maintenance costs, or frequent partner turnover [4,5]. Digital–intelligent innovation collaboration should therefore be understood as a dynamic process in which relationships are continuously formed, maintained, recovered, and reconfigured, rather than as a simple collection of static ties [6–8].

The existing studies explain innovation collaboration from the perspectives of network embeddedness, knowledge bases, technological proximity, geographic distance, and organizational heterogeneity, providing an important foundation for understanding the conditions under which collaborative relationships emerge and how networks perform [9–12]. However, digital–intelligent innovation collaboration exhibits pronounced temporal dependence and state heterogeneity. First, its initial formation, relationship maintenance, and recovery after interruption involve different search costs, foundations of trust, and resource conditions; combining them into a single type of relationship event can allow the average effects to obscure distinct processes. Second, the same scale of collaboration may correspond to stable deepening, rapid expansion, relationship replacement, or low-level dormancy, making it difficult for a single continuous indicator to reveal differences in the collaboration structure and functional state. Third, collaboration resilience is reflected not only in maintaining existing ties after disturbances, but also in the ability to repair interrupted relationships, reconfigure partner combinations, and re-enter higher-functioning states; it therefore cannot be equated directly with network density, connectivity scale, or short-term collaborative output [13–16].

The existing research has generally developed along two separate lines: the identification of relationship mechanisms and the optimization of dynamic governance [17]. Network models can identify the factors associated with relationship formation or state transitions, but they cannot directly determine when to intervene or which actors to target under conditions of limited budgets, diminishing action effectiveness, and the intertemporal accumulation of policy effects. Fixed rules are difficult to adjust as states change; general-purpose reinforcement learning, when lacking task-relevant state representations and resource constraints, may also produce action sequences that are difficult to interpret or infeasible to implement [18,19]. This study therefore further examines how the identified relationship and state-transition mechanisms can be used for sequential governance decisions.

1.1. Research Questions

Focusing on the formation and governance of digital–intelligent innovation collaboration resilience, this study develops four interrelated research questions along a logical sequence from relationship evolution to state transitions, resilience formation, and dynamic governance. First, how are interorganizational collaborations formed, maintained, and recovered under the influence of knowledge bases, technological relatedness, and existing relationships, and can these micro-level relationship rules explain the evolution of collaboration networks? Second, do changes in relationships exhibit qualitatively distinct latent collaboration states that can transition into one another, and how are different state transitions conditionally associated with organizational and relational conditions? Third, how do the state-transition capacity and the capacities for relationship maintenance, recovery, and reconfiguration jointly constitute collaboration resilience? Finally, taking the foregoing relationship mechanisms and state-transition structure as the empirical basis of the governance environment, can mechanism-constrained reinforcement learning produce sequential decisions that outperform standard deep Q-networks and fixed governance schemes? The analysis thus proceeds from micro-level relationship changes to actor states and collaboration resilience and then reformulates these processes as a dynamic governance problem.

1.2. Research Design and Methods

As shown in Figure 1, this study adopts a staged, sequential inferential design that links relationship mechanism identification, state-transition analysis, resilience measurement, and sequential governance. In the first stage, Chinese digital–intelligent innovation co-patents from 2016 to 2025 are used to construct an interorganizational collaboration network. Three risk processes—initial formation, relationship maintenance, and relationship recovery—are distinguished to estimate the micro-level relationship rules, and rolling validation and an out-of-time hold-out test are used to assess whether these rules can reproduce the scale, relationship updating, and major structural characteristics of the observed network. Only relationship rules supported by generative validity tests are used for subsequent recursive network generation.

In the second stage, organizational collaboration states are identified on the basis of the foregoing network evolution. Continuous models, threshold models, and a hurdle-type latent Markov model are compared to determine the latent states and their transition structures, after which the conditional associations between the organizational attributes and relationship conditions and the key transition probabilities are examined. The number of states, state profiles, and transition structure enter the next stage only after receiving support from the classification certainty, out-of-sample performance, and generative fit.

The third stage does not construct an independent resilience indicator unrelated to the preceding mechanisms. Instead, the validated state-transition capacity and the capacities for relationship maintenance, recovery, and reconfiguration are jointly incorporated into the measurement of collaboration resilience, allowing for the resilience to be traced back to the actor's current state and its underlying relationship processes.

In the fourth stage, OSR-DDQN is constructed while the preceding relationship rules, state-transition structure, and resilience measurement remain fixed, thereby translating them into a sequential governance environment under a limited budget. Reinforcement learning is allowed to learn the intervention targets, implementation sequences, and stopping points only under the constraints of the established empirical mechanisms; it does not re-estimate the preceding relationship mechanisms or state structure in response to governance outcomes. In this way, relationship generation, state transitions, resilience

measurement, and governance decisions are sequentially linked, with the empirical results from each stage providing the basis for the next.

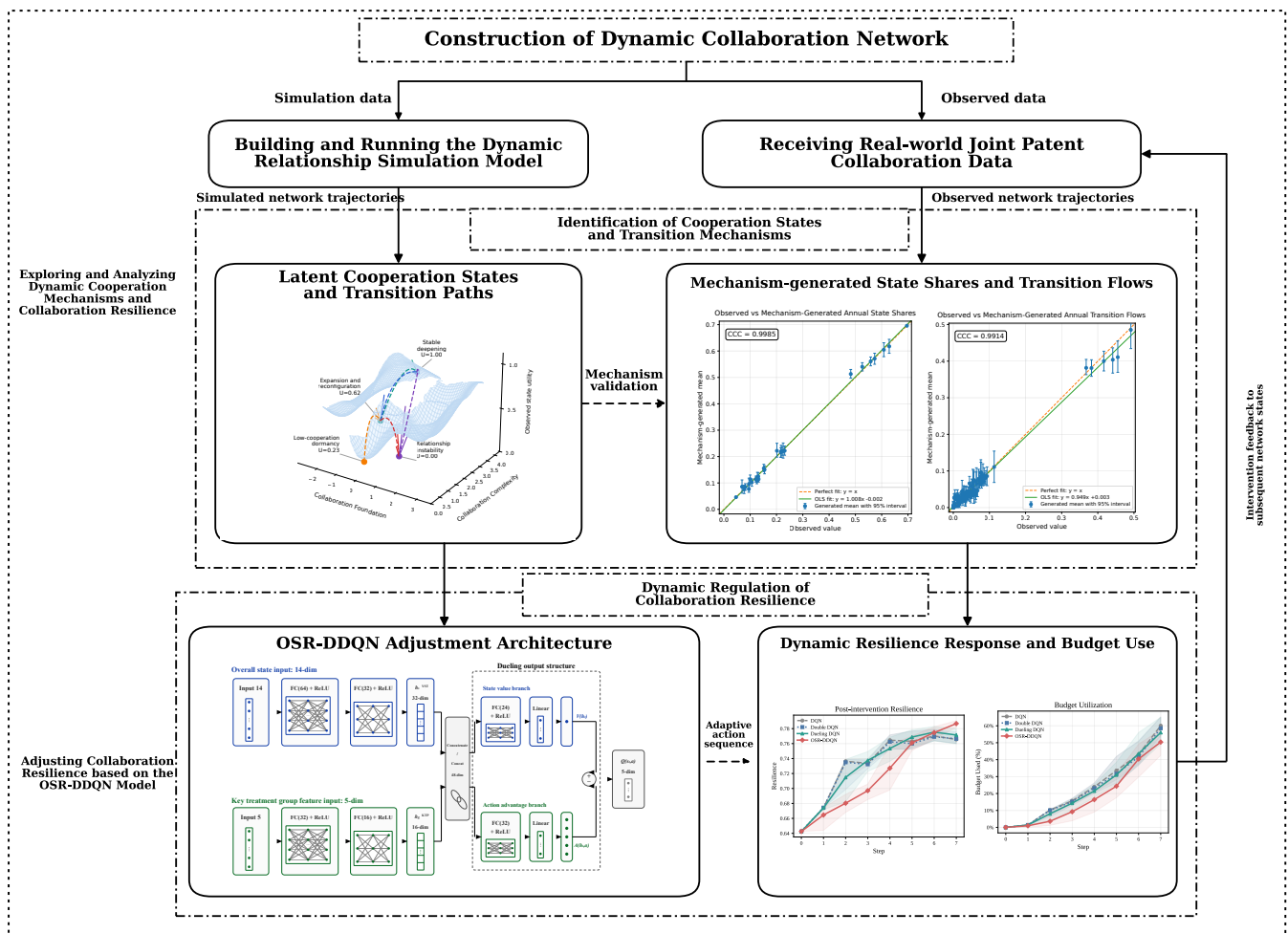


Figure 1. Research methodology framework.

1.3. Research Contributions

The contributions of this study are three-fold.

First, methodologically, this study develops a cross-level dynamic analytical framework with mechanism consistency, incorporating micro-level relationship events, actor-level state transitions, collaboration resilience, and sequential governance within a common evolutionary logic. Rather than placing network simulation, state identification, resilience measurement, and reinforcement learning side by side as separate tools, the framework allows the empirically validated relationship mechanisms to continuously constrain the subsequent network generation, state transitions, and governance decisions, thereby transmitting the mechanisms across all levels, from the micro-level relationship processes to macro-level resilience outcomes and dynamic governance. This design strengthens the interpretability and consistency across analytical levels and avoids the mechanism fragmentation and result stitching that often arise in multi-model combinations.

Second, this study interprets collaboration resilience as a state-dependent relational process capacity rather than as a proxy for network connectivity scale or static structural stability. The four latent states show that similar levels of collaboration can correspond to qualitatively different modes of relationship organization, including low-cooperation dormancy, expansion and reconfiguration, relationship instability, and stable deepening. After the global BH-FDR correction, the organizational conditions exhibit clearly differentiated

conditional associations across transition paths. These findings imply that collaboration advantages cannot be interpreted uniformly without reference to an actor's current state and transition direction; the formation of collaboration resilience must jointly consider the current state, state reachability, and the capacities for relationship maintenance, recovery, and reconfiguration.

Third, this study further distinguishes two governance constraints: insufficient resources and inappropriate resource allocation. At low-budget levels, the primary problem is that necessary actions cannot be implemented sufficiently. Within the budget range examined, however, once the budget reaches a minimally sufficient level, the final collaboration resilience tends to stabilize and performance differences arise mainly from the intervention targets, implementation sequences, action timing, and stopping conditions. This result shows that resource constraints in innovation collaboration governance are stage-dependent: increasing resource inputs and improving resource allocation are not the same problem. It also provides a basis for explaining why dynamic state feedback can outperform fixed stage-based schemes.

2. Related Literature and Theoretical Foundations

2.1. Digital–Intelligent Innovation Collaboration and Networked Knowledge Production

Digital technologies are highly recombinable and can be embedded in different products, services, and industrial settings, making innovation activities increasingly dependent on the combination of cross-domain knowledge and the joint participation of multiple types of organizations [20,21]. Firms, universities, research institutions, and public sector organizations possess different resources for technological development, basic research, and application scenarios, and have different institutional support, and a single organization is usually unable to complete complex digital–intelligent innovation independently. Interorganizational collaboration can broaden the scope of knowledge search, introduce complementary capabilities, and create more possibilities for technological recombination.

The knowledge gains from collaboration do not depend solely on the number of partners. Partners need a certain degree of knowledge complementarity, while also needing to control communication, coordination, and integration costs. Greater partner diversity facilitates access to different knowledge sources, but excessive differences may also increase the difficulties in mutual understanding [22]. Geographic and technological proximity can reduce interaction costs, whereas excessive proximity may confine collaboration to similar knowledge and fixed partners over the long term [23]. Innovation collaboration therefore continually involves a tradeoff between knowledge expansion and coordination burdens.

Co-patents record joint applicants, application dates, and technological classifications and can therefore be used to identify organizational collaborations that have produced formal technological outcomes and to observe long-term changes in collaboration networks and technological combinations. Related studies have accordingly examined how knowledge bases, network embeddedness, and relationship history affect the formation and evolution of patent collaboration [24]. However, co-patents capture only formal collaborations that reach the joint application stage and cannot cover consulting services, joint standard setting, open-source collaboration, or informal knowledge exchange. This study therefore uses them as observable evidence of formal innovation collaboration rather than equating them with all interactions between organizations.

2.2. Dynamic Evolution Mechanisms of Innovation Collaboration Relationships

The existing innovation network research mainly focuses on structural characteristics, such as centrality, structural holes, clustering, and partner diversity, and examines how these characteristics affect knowledge acquisition and innovation performance. Such

research can describe an actor's position in a network but is less able to explain how specific relationships emerge, persist, dissolve, and reform. Even when two actors have the same tie status in a given period, the relationship histories underlying those statuses may be completely different.

Establishing a new relationship requires searching for and evaluating potential partners; continuing an existing relationship depends more heavily on trust, coordination experience, and joint investment; and recovering an interrupted relationship depends on the knowledge and relational foundations left by previous collaboration. Relationship formation, maintenance, and recovery should therefore not be treated as the same process. By incorporating lagged networks and relationship histories, dynamic network models can separately examine the conditions for the development of new relationships and continuation of existing ones [25,26]. Related patent studies also show that knowledge stocks, network embeddedness, and relationship stability jointly influence subsequent collaboration choices [27].

The purpose of simulation is not to replace real data but to examine what aggregate outcomes emerge when multiple micro-level mechanisms operate together. When parameters are estimated from historical data and then subjected to out-of-time testing, simulation can connect relationship-level analysis with network-level evolution while retaining uncertainty across different evolutionary paths [28,29].

2.3. Concept, Mechanisms, and Measurement of Collaboration Resilience

Resilience does not mean that a system remains unchanged; rather, it is the capacity to maintain basic functions, cope with losses, and adjust when facing disturbances. The research on organizational resilience has gradually shifted from post-shock recovery outcomes toward the processes of anticipation, coping, and adaptation. In innovation collaboration networks, disturbances may arise from the exit of key partners, changes in technological trajectories, contractions in collaboration resources, or rising interorganizational coordination costs. Actors therefore need to not only maintain valuable existing relationships, but also repair interrupted relationships and identify new partners when the original collaboration structure fails.

The existing studies of network resilience often use connectivity, network efficiency, and structural losses following node failures to evaluate a network's capacity to withstand disruption [30]. These indicators can reflect structural changes caused by external shocks, but they pay less attention to how collaborative relationships themselves are formed, dissolved, and re-established. In innovation networks, having more connections does not necessarily imply greater resilience. Excessive partner concentration, deeper technological dependence, or persistent tie dissolution can expose highly collaborative actors to substantial risk; actors with fewer connections may also recover collaborations by relying on existing relational capital. Collaboration resilience therefore needs to consider the current functional state of collaboration, the probability that actors enter more favorable states, and their capacities for relationship maintenance and recovery [31,32].

The existing studies rarely place the changes in collaboration states and specific relationship processes within the same measure. The collaboration scale alone cannot distinguish stable deepening from short-term activity generated by frequent partner replacement; structural stability alone may also overlook actors' capacities to recover and reorganize relationships. This study therefore conceptualizes resilience from both the collaboration states and relationship processes and focuses on whether actors can maintain effective collaboration, repair interrupted relationships, and re-enter higher-functioning states.

2.4. Reinforcement Learning for Innovation Network Governance

Innovation collaboration governance usually requires sequential implementation and is jointly constrained by budgets and action timing. For sequential governance problems with delayed effects and resource constraints, reinforcement learning can use intertemporal feedback to evaluate the outcomes of actions taken at different stages. This study further considers the intertemporal effects of relationship repair, partner activation, and structural consolidation, as well as the possible decay in effectiveness caused by repeated actions.

Reinforcement learning compares the outcomes of actions across stages through environmental feedback and is suitable for problems involving sequential decisions and delayed effects. DQN uses a neural network to estimate long-term action values, Double DQN reduces value overestimation, and Dueling DQN further separates the value of the state itself from the incremental contribution of specific actions [33–35]. These methods provide tools for sequential decision making, but a more complex algorithmic structure does not necessarily imply a more effective governance strategy.

Reinforcement learning results are readily affected by random seeds, hyperparameters, reward design, and evaluation samples. If the parameters are repeatedly adjusted based on test results or only the best-performing training results are reported, a model's advantages may be substantially overstated. Reinforcement learning studies therefore need to strictly separate training, validation, and testing, and assess result stability through multiple independent training runs [36,37].

Deep reinforcement learning has already been applied to the dynamic governance of social systems, such as online communities, indicating that state feedback can support intertemporal intervention decisions [38]. In innovation collaboration networks, however, relationship formation, state transitions, and resilience changes are still generally studied separately from governance optimization.

2.5. Literature Review

As shown in Table 1, the existing studies have advanced the analysis of innovation collaboration relationship evolution, empirical network simulation, network resilience assessment, and sequential governance, yet the studies most closely related to the present work still stop at different analytical boundaries. Jin et al. [12] use patent data to examine how network embeddedness and knowledge stocks affect collaboration dynamics and demonstrate the importance of relationship history and network structure for subsequent collaboration choices, but their analysis ultimately remains focused on the changes in collaborative relationships themselves. Neuländtner et al. [29] further use empirical multi-agent simulation to study intercity R&D collaboration in China, linking micro-level rules to macro-level network evolution, but they do not proceed to identify the organization-level latent collaboration states or translate state transitions into collaboration resilience and a governance environment.

In resilience research, Tóth et al. [15] and Hu and Ma [16] discuss system resilience from the perspectives of technological network structure and changes in urban innovation networks, respectively, showing that network relationships and their dynamic adjustment are closely related to resilience. However, these studies rarely trace resilience further back to specific processes, such as relationship formation, maintenance, recovery, and reconfiguration. In governance research, Vinyals et al. [17] show that empirical simulation and reinforcement learning can be combined for innovation diffusion policy optimization, while Du et al. [38] apply deep reinforcement learning to resilience regulation in social systems. Yet, the decision environments in these studies are not jointly constructed from empirically generated innovation collaboration relationship mechanisms and latent-state transitions. Accordingly, the key unresolved issue in the existing literature is not whether the above

methods can separately be applied to dynamic networks, state identification, resilience assessment, or policy optimization, but how these components can be grounded in the same set of verifiable empirical mechanisms and connected into a continuous explanatory chain from relationship events to state transitions and from resilience measurement to sequential governance.

Table 1. Studies most closely related to this study and their analytical boundaries.

Representative Study	Main Advance	Closest Connection to This Study	Unresolved Issue
Jin et al. [12]	Dynamic mechanisms of patent collaboration	Relationship history, knowledge bases, and network embeddedness	Does not further connect latent states, collaboration resilience, and governance decisions
Neuländtner et al. [29]	Empirical mechanism-driven simulation of collaboration networks	Recursive generation of macro-level networks from micro-level rules	Does not proceed to organization-level state identification, resilience measurement, or sequential governance
Tóth et al. [15]; Hu and Ma [16]	Network structure and innovation network resilience	Explains resilience differences through network change	Does not systematically connect resilience to relationship formation, maintenance, recovery, and latent-state transitions
Vinyals et al. [17]	Multi-agent simulation and reinforcement learning policy optimization	Intertemporal decision making in a simulation environment	Governance target is innovation diffusion; the environment is not jointly generated by collaboration relationship mechanisms and state transitions
Du et al. [38]	Deep reinforcement learning and resilience regulation in social systems	State feedback and dynamic intervention	The research object is not an innovation collaboration network, and the governance environment lacks constraints from empirical relationship mechanisms
This study	Relationship mechanisms, latent states, collaboration resilience, and sequential governance	Continuous analysis from real relationship processes to dynamic governance	Empirically identified and validated relationship rules remain embedded throughout network generation, state transitions, resilience measurement, and governance decisions

The key to the integration developed here is that the output of each preceding stage is frozen and passed to the next. Relationship mechanisms are used to recursively generate networks; state transitions and relational process reserves are used to construct collaboration resilience; and reinforcement learning learns governance strategies under frozen mechanism, cost, and budget constraints. This produces a traceable computational chain from relationship evolution and state transitions to resilience measurement and sequential governance.

3. Exploration and Analysis of Digital–Intelligent Innovation Collaboration and Its Dynamic Mechanisms

3.1. Construction of a Multi-Agent Simulation Model for Digital–Intelligent Innovation Collaboration

Digital–intelligent innovation collaboration is not a one-off partner-matching event but a dynamic relational process in which knowledge resources, relational capital, and network structure continuously interact. Accordingly, this study takes organization pairs

as the micro-level units of interaction; uses co-patents as the observable outputs of formal collaborative innovation; sequentially models relationship formation, maintenance, and recovery; and further identifies network evolution, latent collaboration states, and collaboration resilience [39].

(1) Relationship states: Collaborative relationships are history-dependent, and their initial formation, maintenance, and recovery, respectively, involve different conditions related to partner search, relationship continuation, and reuse of relational capital. Accordingly, the three candidate sets are defined as follows:

$$\begin{aligned} P_t &= \{(i, j) : i, j \in V_t, i < j\}, \mathcal{H}_{t-1} = \bigcup_{\ell < t} E_\ell^C, \\ S_t^M &= \{e \in P_t : e \in E_{t-1}^C\}, \\ S_t^R &= \{e \in P_t : e \in \mathcal{H}_{t-1}, e \notin E_{t-1}^C\}, S_t^F = \{e \in P_t : e \notin \mathcal{H}_{t-1}\}. \end{aligned} \quad (1)$$

In Equation (1), V_t , P_t , and E_t^C denote, respectively, the set of organizations in period t , the set of all candidate organization pairs, and the set of collaboration edges actually observed in period t . $\mathcal{H}_{t-1}$ denotes the set of historical collaboration edges formed through period $t - 1$, and e denotes any candidate organization pair. The superscripts C, M, R, and F denote collaboration edges, relationship maintenance, relationship recovery, and initial relationship formation, respectively.

(2) Collaboration payoff and relationship selection: In the model, the collaboration payoff jointly incorporates information on the actors' knowledge bases, technological and relational foundations, network embeddedness, geographic proximity, and organizational type. Because the payoff cannot be observed directly, this study identifies its relative magnitude using the probability that a relationship occurs. The candidate organization pairs for initial formation are numerous while the events are sparse, so a probabilistic model capable of accommodating nonlinearity and interactions is used; maintenance and recovery have explicit relationship histories and are modeled separately using Logit models [40]. The estimation and selection of the three processes are written in a unified form as follows:

$$\begin{aligned} \hat{p}_{ij,t}^{(r)} &= \hat{\Phi}_r(x_{ij,t}^{(r)}), \\ \Pi_{ij,t}^{(F)} &= \tau^{-1} \ln \hat{p}_{ij,t}^{(F)} + 1/2 \ln(a_{i,t} a_{j,t}) + \ln \Omega_{ij,t}^{(F)}, \\ \Pi_{ij,t}^{(r)} &= \ln \hat{p}_{ij,t}^{(r)} + \ln \Omega_{ij,t}^{(r)}, r \in \{M, R\}, \\ y_{ij,t}^{(r)} &= \chi_{ij,t}^{(r)} 1[\Pi_{ij,t}^{(r)} + u_{ij,t}^{(r)} \geq \kappa_{t,g}^{(r)}], u_{ij,t}^{(r)} \sim \text{Gumbel}(0, 1). \end{aligned} \quad (2)$$

In Equation (2), $r \in \{F, M, R\}$ indexes the three relationship processes of initial formation, maintenance, and recovery; $\hat{p}_{ij,t}^{(r)}$ is the predicted probability that organization pair (i, j) undergoes process r in period t . $\Phi_r(\cdot)$ denotes the corresponding probability model, $x_{ij,t}^{(r)}$ is the ex ante feature vector, $\Pi_{ij,t}^{(r)}$ is the log selection weight of a candidate relationship, and τ is the temperature parameter for the initial formation process. $a_{i,t}$ and $a_{j,t}$ are the attractiveness weights of the two endpoint actors, $\Omega_{ij,t}^{(r)}$ is the network structure adjustment term, and $y_{ij,t}^{(r)}$ is the relationship–event indicator. $\chi_{ij,t}^{(r)}$ is the candidate eligibility and capacity constraint indicator, $\kappa_{t,g}^{(r)}$ is the endogenous selection threshold for type group g_r , and $u_{ij,t}^{(r)}$ is a random disturbance following the standard Gumbel distribution. The predicted probability represents the relative payoff of relationship behavior within the corresponding risk set; the remaining terms capture, in sequence, actor attractiveness, network structure adjustment, candidate eligibility and the degree–capacity constraints, and the endogenous threshold jointly determined by the total relationship volume and type quotas. The detailed components are reported in Table 2. The candidate models for the

three relationship processes, the nested rolling out-of-time validation, the unified selection rule, and the final parameter-freezing procedure are reported in Section SB.

Table 2. Information components of collaboration payoff.

Information Category	Main Variables	Role
Actor foundation	Lagged patents, IPC breadth, activity level, and actor age	Characterizes knowledge inputs and absorptive capacity
Technological and relational foundation	IPC similarity, historical collaboration intensity, duration, and relationship memory	Characterizes technological relatedness and existing relational capital
Network and spatial structure	Degree, common partners, bridging, geographic proximity, and actor type	Characterizes embedded opportunities, coordination costs, and organizational differences
Opportunity and capacity	Year, candidate set size, total relationship volume, type quotas, and degree-capacity	Constrains the feasible range of relationship realization

3.2. Simulation Data and State Indicators

The observed data were drawn from Chinese patents related to digital–intelligent innovation from 2016 to 2025. The patent data were obtained from the incoPat database using a combined search based on IPC technology classifications and keywords; the complete search expression and processing rules are reported in Section SA. The raw records were first harmonized in terms of field names, deduplicated by application number, restricted to the study years, and subjected to field-level quality checks, yielding a frozen master table of 63,558 patents. The applicants were then separated, the organizational actors were identified and their names standardized, and the individual applicants were removed, resulting in 8612 co-patents containing at least two distinct organizations. The organizational applicants on the same patent were paired to construct undirected collaborative relationships and aggregated by year, producing 6972 annual cross-group collaboration edges, as reported in Table 3. This study adopts an open actor pool: an organization enters the observable actor records once it first exhibits relevant patent activity; absence of a relevant patent in a given year is treated only as temporary inactivity, while the actor identity and prior relationship history are retained. The annual changes in the actor pool therefore reflect patent activity and the data-visibility boundary, and relationship formation, maintenance, recovery, and interruption are interpreted as collaboration evolution processes only after the actor pool has been defined.

Co-patents record the joint applicants, application time, and technological classifications within the same traceable technological output. They are therefore suitable for identifying formal innovation collaboration that has generated patentable outputs and for constructing organization-level collaboration networks that are comparable over time [41]. Each simulation trajectory is initialized using the observed 2017 network and prior relationship history and recursively generates the networks for 2018–2025. The latent-state and actor–year mechanism analyses use 2019–2024 as the estimation period, with 2025 reserved as an out-of-time hold out; the relationship formation, maintenance, and recovery models follow the independent rolling model selection, development-end validation, and fully held-out 2025 protocol described in Section SB. The observed sample and 30 independent trajectories use the same annual granularity, relationship definitions, and indicator definitions. Each trajectory is fitted separately rather than averaged ex ante, thereby preserving uncertainty across simulation paths.

Table 3. Data foundation and analytical samples.

Data Category	Sample Size	Purpose
Patents related to digital–intelligent innovation	63,558 patents	Search and topic screening
Organizational co-patents	8612 patents	Identify formal interorganizational collaboration
Annual cross-group collaboration edges	6972 edges	Construct the main network
Relationship–year panel	90,340 rows	Construct risk sets for existing and historical relationships
Observed mechanism sample	5416 observations	Actor–year records for 2019–2025
Aligned simulation mechanism sample	162,480 observations	5416 valid mechanism observations × 30 trajectories

The same number of connections may arise from the continued deepening of stable partnerships or from short-term expansion following frequent partner replacement. To distinguish these processes, which have different governance implications, this study constructs state observations along two dimensions: the functional level of collaboration and the mode of relationship renewal.

$$y_{it} = Z \left[\ln(1+s_{it}), \ln(1+d_{it}), \ln(1+n_{it}^F), \ln(1+n_{it}^M), \frac{n_{it}^F + n_{it}^R + n_{it}^D}{\max\{1, d_{i,t-1} + n_{it}^F + n_{it}^R\}} \right]^T \quad (3)$$

In Equation (3), $y_{i,t}$ is the five-dimensional state observation vector for actor i in period t , and $Z(\cdot)$ is the common scaling operator determined from the training sample. $s_{i,t}$ and $d_{i,t}$ denote the collaboration intensity and partner breadth, respectively; $n_{i,t}^F$, $n_{i,t}^M$, $n_{i,t}^R$, and $n_{i,t}^D$ denote the numbers of initially formed, maintained, recovered, and interrupted relationships, respectively. The recovery and interruption counts do not occupy separate state dimensions; together with the number of initially formed relationships and lagged partner size, they are used to calculate the fifth indicator, the relationship turnover rate. $(\cdot)^T$ denotes vector transposition. The first four elements of the state vector are collaboration intensity, partner breadth, new relationship formation, and relationship maintenance, and the fifth is the relationship turnover rate calculated according to Equation (3).

3.3. Analysis of Collaboration Transition Associations Based on a Latent State-Transition Model

The same scale of collaboration may correspond to different states, and actors may shift from one mode of relationship organization to another within a relatively short period [42,43]. This study therefore first compares the continuous and threshold models and then adopts a latent-state framework to identify discrete collaboration states. A Bernoulli HMM based on binary hurdle indicators is used to select the number of states. Once the number of states is fixed, the Gaussian HMM in Equation (4) is used to estimate the state profiles and transition mechanisms from the five-dimensional continuous observations.

$$y_{it} | S_{it} = k \sim N(\mu_k, \text{diag}(\sigma_k^2))$$

$$a_{it,kj} = \Pr(S_{it} = j | S_{i,t-1} = k, z_{it}) = \frac{\exp(\beta_{kj}^\top z_{it})}{\sum_{h=1}^K \exp(\beta_{kh}^\top z_{it})}, \quad \beta_{kK} = 0 \quad (4)$$

In Equation (4), $S_{i,t}$ denotes the latent collaboration state of actor i in period t ; $k, j, h \in \{1, \dots, K\}$ are the state indices; and K is the total number of candidate states.

μ_k and σ_k^2 are, respectively, the vectors of observed means and variances for state k ; and $\text{diag}(\sigma_k^2)$ is the diagonal covariance matrix. $a_{i,t,kj}$ is the conditional probability of transitioning from state k to state j , and $z_{i,t}$ is the vector of transition covariates. β_{kj} is the corresponding coefficient vector, while $\beta_{kK} = 0$ indicates that state K is used as the reference outcome.

To use the latent-state transition results to further identify the actors' transition propensities, this study defines the relative transition advantage as the ratio of the conditional probability of a target state to that of a reference state under the same origin state. Let b_j denote the reference outcome corresponding to origin state j . The linear predictor for actor i transitioning from state j to state k in period t is then defined as

$$\eta_{jk,it} = \log \frac{\Pr(S_{i,t+1} = k \mid S_{it} = j, x_{it})}{\Pr(S_{i,t+1} = b_j \mid S_{it} = j, x_{it})} = \alpha_{jk} + \beta_{jk}^T x_{it} \quad (5)$$

In Equation (5), α_{jk} is the transition intercept, β_{jk} is the covariate coefficient, and x_{it} contains the organizational attributes and relational conditions measured before the transition occurs.

By the definition of log odds, the direction of transition propensity can be represented by the sign of the linear predictor:

$$\text{sgn}(\eta_{jk,it}) = \text{sgn}[P_{jk,it} - P_{jb_j,it}] \quad (6)$$

In Equation (6), $\eta_{jk,i,t}$ denotes the log odds of transitioning from origin state j to target state k relative to transitioning to reference state b_j . Here, j , k , and b_j denote the origin state, target state, and reference outcome, respectively, while i and t denote the actor and period. $P_{jk,i,t}$ and $P_{jb_j,i,t}$ are the conditional probabilities of the target and reference transitions, respectively, and $\text{sgn}(\cdot)$ is the sign function.

To provide an intuitive representation of the distributional characteristics of latent states under different collaboration conditions and their typical transition patterns, Figure 2 presents a conceptual landscape of latent collaboration-state transitions. The vertical axis represents the latent collaboration states, while the two horizontal axes represent the innovation and relational foundations and technological and structural support, respectively. Different regions of the surface correspond to the four latent states of stable deepening, expansion and reconfiguration, relationship instability, and low-cooperation dormancy.

As shown in Figure 2, four typical paths are illustrated. Path $a \rightarrow b$ represents a gradual adjustment within the stable-deepening state; path $c \rightarrow d$ represents partner configuration and structural optimization within expansion and reconfiguration; path $e \rightarrow f$ represents movement toward the instability region after relational conditions deteriorate; and path $g \rightarrow h$ represents prolonged residence in a low-collaboration, low-connectivity dormant region. The state transitions are position-dependent. An improvement of the same magnitude may activate collaboration for an actor in low-cooperation dormancy but consolidate relationships for an actor already in expansion. Discussing the variable associations without conditioning them on the current state can therefore produce excessively averaged conclusions. The transition directions are also asymmetric: the relational accumulation and structural support required to enter a high-functioning state need not be the mirror image of the conditions associated with deterioration. The transition-sensitive regions and boundaries in the figure are intended to convey this conditional dependence and directional asymmetry.

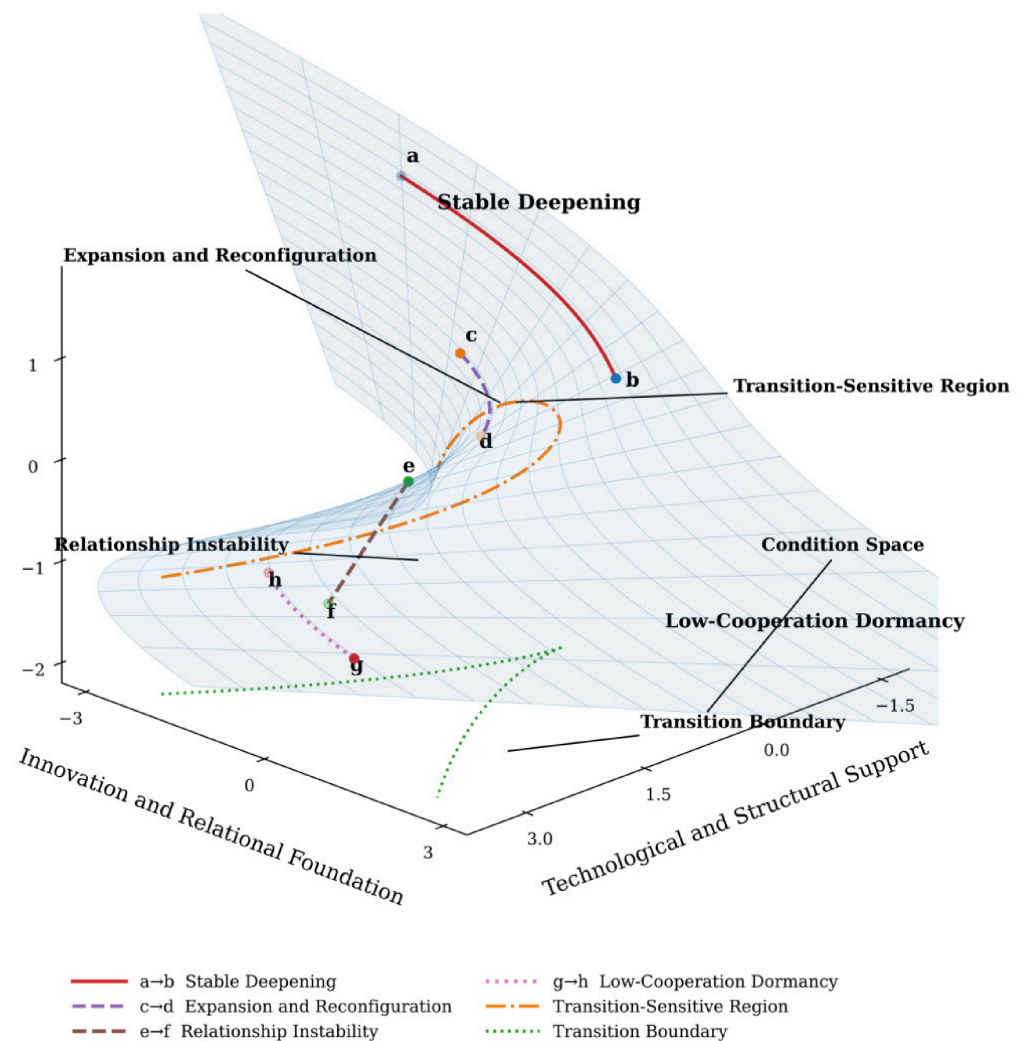


Figure 2. Conceptual landscape of latent collaboration-state transitions.

3.4. Calculation of Digital–Intelligent Innovation Collaboration Resilience

The collaboration resilience cannot simply be equated with a high level of collaboration. Accordingly, this study defines collaboration resilience as an actor's composite capacity to maintain effective relationships under disturbances, repair interrupted relationships, and re-enter higher-functioning states through necessary partner reconfiguration.

To avoid misclassifying frequent exits and unstructured partner replacement as adaptation, the state utility penalizes turnover that exceeds the level of initial relationship formation and applies equal-weight shrinkage between the actor-specific transition matrix and the overall transition matrix:

$$r_k = \frac{1}{4} \sum_{m=1}^4 \mu_{km} - 0.35 \max(\mu_{k5} - \mu_{k3}, 0), \quad u_k = \frac{r_k - \min_h r_h}{\max_h r_h - \min_h r_h} \quad (7)$$

$$A_i^* = 0.5A_i^{NH} + 0.5A^H, \quad T_i = \pi_i A_i^* (0.7\mathbf{u} + 0.3\mathbf{h}), \quad \mathbf{h} = (0, 0, 1, 1)^\top$$

In Equation (7), r_k is the raw collaboration utility of state k , μ_{km} is the mean of that state on observed indicator m , and u_k is the normalized state utility. $\mathbf{h}$ is the favorable-state indicator vector for the expansion-and-reconfiguration and stable-deepening states; A_i^{NH} and A^H denote the actor-specific and overall state-transition matrices, respectively; and A_i^* is the transition matrix obtained by the equal-weight shrinkage of the two. π_i

is the posterior probability vector for the current state of actor i , and T_i is that actor's state-transition capacity.

$$\begin{aligned} q_i^M &= \frac{n_i^M + 1}{n_i^M + n_i^D + 2}, & q_i^R &= \frac{n_i^R + 1}{n_i^R + n_i^D + 2} \\ q_i^Q &= \frac{n_i^F + n_i^R + 1}{n_i^F + n_i^R + n_i^D + 2}, & q_i^L &= \frac{n_i^M + n_i^R + 1}{n_i^F + n_i^M + n_i^R + n_i^D + 2} \\ \tilde{B}_i &= \rho_i \frac{q_i^M + q_i^R + q_i^Q + q_i^L}{4} + (1 - \rho_i)0.5, & \rho_i &= \frac{E_i}{E_i + 8}, \quad E_i = \sum_r n_i^r \end{aligned} \quad (8)$$

In Equation (8), q_i^M , q_i^R , q_i^Q , and q_i^L denote, respectively, the resistance, recovery, reconfiguration, and continuity ratios of actor i . E_i is the number of valid relationship events, $\rho_i = E_i / (E_i + 8)$ is the evidence reliability weight, and B_i is the relational process reserve after shrinkage by the evidence weight.

Finally, the state-transition capacity and relational process reserve are combined using a parallel reliability form:

$$R_i = 1 - (1 - T_i)(1 - B_i) = T_i + (1 - T_i)B_i \quad (9)$$

In Equation (9), R_i denotes the composite collaboration resilience of actor i , T_i denotes the state-transition capacity, and B_i denotes the relational process reserve. The composite collaboration resilience ranges from 0 to 1. The two components enter a bounded complementary aggregation with a diminishing incremental contribution. The mathematical form permits partial substitution: when $B_i = 0$, Equation (9) reduces to $R_i = T_i$; when $T_i = 0$, it reduces to $R_i = B_i$. Therefore, a high value for one component can still yield a high composite R_i when the other component is low, and T_i and B_i should be reported alongside R_i .

The diminishing incremental contribution follows directly from $\frac{\partial R_i}{\partial T_i} = 1 - B_i$ and $\frac{\partial R_i}{\partial B_i} = 1 - T_i$: as either component increases, the marginal contribution of the other decreases. The construct rationale, strict out-of-time criterion validity, and parameter-sensitivity tests for the collaboration resilience are reported in Section 4.4 and Section SD. The parameter sensitivity is examined using one-factor-at-a-time perturbations, so the conclusions are limited to single-parameter changes.

4. Experimental Results for Digital–Intelligent Innovation Transition Associations and Collaboration Resilience

4.1. Experimental Data, Parameter Settings, and Evaluation Methods

This study consistently follows the temporal validation principle of using the past to predict the future [44]. The relationship models first undergo rolling model-family comparisons over 2022–2024; the final models are then fitted using data from 2018 to 2023, with 2024 used for development-end validation and calibration of formation probabilities, and 2025 completely held out. The latent-state and actor–year mechanism models are estimated using the 2019–2024 data, with 2025 held out. Criterion validity of collaboration resilience is evaluated under a separate complete re-freezing through 2023, and validated using relationship shocks in 2024 and recovery outcomes in 2025.

The experiment first estimates the initial relationship formation, relationship maintenance, and relationship recovery, and then recursively generates the annual collaboration networks after the parameters are frozen. The network model is evaluated by comparing the observed and simulated annual statistics and distributions of node-level relationship characteristics, thereby avoiding the use of agreement in total edge counts as a substitute for structural validity. The analysis then compares the continuous, threshold, and

latent-state models to determine the number of states and identify the state profiles and the baseline transition structure. Holding the state definitions fixed, it further estimates the conditional associations between organizational attributes and relational conditions and transition probabilities, and uses the observed annual state changes to evaluate the overall reproduction capability of the state-transition model.

4.2. *Validity Tests of Dynamic Collaboration Network Simulation*

4.2.1. Simulation Test of Aggregate Network Characteristics

The validity of a dynamic network model cannot be judged solely by its proximity to edge counts or mean degree because the same aggregate scale may be generated by different relationship processes and local structures [45]. Following this logic, the 2025 validation simultaneously examines the network scale, relationship events, local clustering, connectivity, and cross-domain collaboration.

To ensure that the 2025 evaluation remains fully *ex ante*, the forecasting stage uses only historical information available through 2024. First, the 2025 active actor pool is generated from the continuation of existing actors and the historical expansion range of the actor pool. The predicted number of active actors is 13,974, compared with an observed 14,243, corresponding to an error of 1.89%. Once the actor pool is locked, network expansion is further constrained using the historical conversion rates through 2024, at which point actors enter the collaboration network and establish their first collaborative relationships; the specific relationships are still generated by the frozen formation, maintenance, and recovery rules. The observed 2025 actor list, actor count, collaboration edges, and patent outcomes are used only for evaluation after the simulation is completed. This two-stage treatment explicitly separates changes in the observable actor pool from collaboration relationship generation: the first stage tests the actor-visibility boundary determined by annual patent activity, whereas the second tests the generative capacity of the formation, maintenance, and recovery rules conditional on the given actor pool. The changes in actor counts are therefore not directly interpreted as changes in collaboration mechanisms.

Figure 3 compares annual statistics of the observed network with those of 30 recursively simulated networks from 2019 to 2025. Table 4 summarizes the mean errors across five rolling out-of-time test folds from 2021 to 2025, and separately reports the fully recursive 2025 hold-out results. The five-fold mean absolute relative errors for node count, edge count, mean degree, and the number of initially formed relationships are 0.78%, 0.92%, 1.70%, and 1.01%, respectively. The corresponding errors for relationship maintenance and recovery are 4.55% and 4.68%, higher than for initial formation but still below 5%.

The 2025 predictive simulations yield relative deviations of +0.56%, −1.48%, −2.01%, and −1.32% for network nodes, network edges, mean degree, and initially formed relationships, respectively, and −5.93% and +3.72% for relationship maintenance and recovery. Across 14 network indicators, the absolute relative deviation does not exceed 5.93%, the absolute standardized difference does not exceed 1.09, and every two-sided Monte Carlo *p*-value is at least 0.323. The errors are not concentrated in any single structural dimension, such as local clustering, cross-type collaboration, or cross-province collaboration. Thus, even after completely isolating the 2025 actor and relationship information, the frozen rules can simultaneously reproduce the network scale, relationship renewal, and principal structural characteristics.

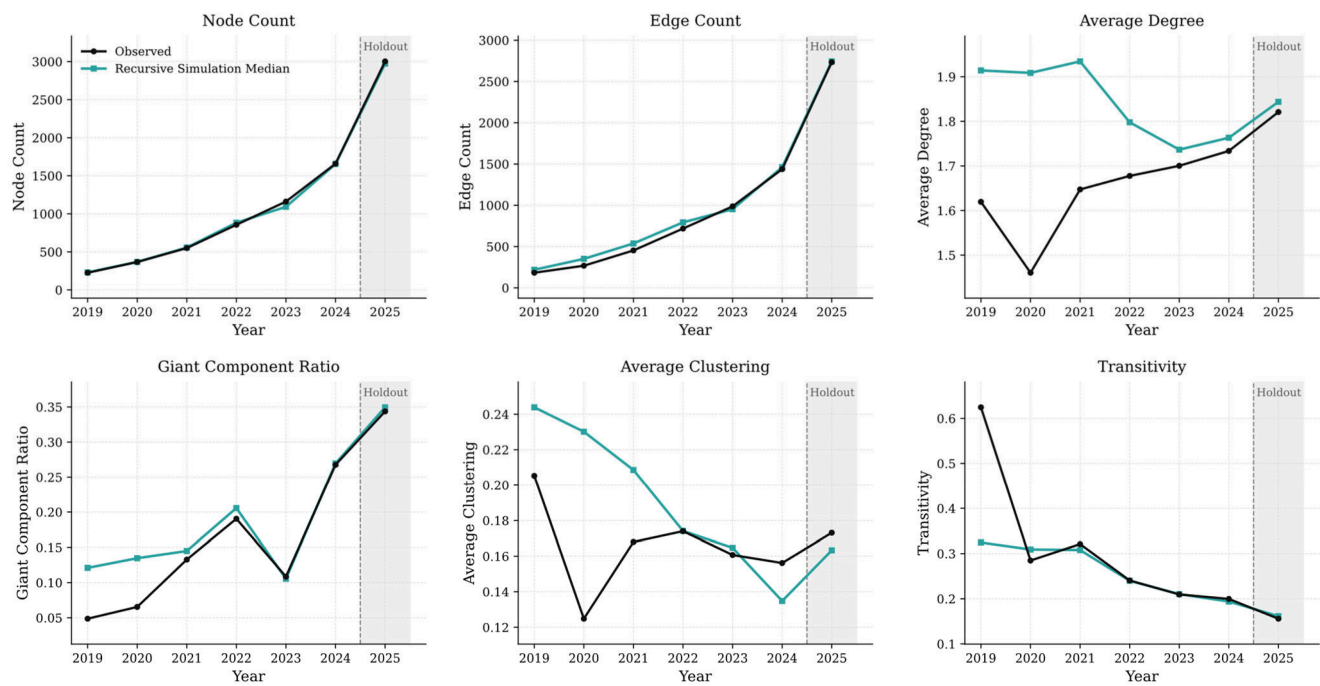


Figure 3. Comparison of annual statistics between the observed and recursively simulated networks.

Table 4. Rolling validation and 2025 hold-out results for the dynamic collaboration network.

Indicator (Evaluation Dimension)	Five-Fold Mean Absolute Relative Error (%)	Five-Fold Mean Signed Error (%)	Observed Value in 2025	Mean of 2025 Recursive Simulations (SD)	2025 Relative Deviation (%)	Standardized Difference z	Two-Sided MC <i>p</i> -Value
Number of network nodes	0.78	−0.71	2967	2983.567 (42.202)	+0.56	−0.39	0.968
Number of network edges	0.92	+0.92	2712	2671.900 (44.461)	−1.48	+0.90	0.452
Mean degree	1.70	+1.70	1.82811	1.791 (0.040)	−2.01	+0.91	0.452
Number of initially formed relationships	1.01	+0.94	2455	2422.700 (44.811)	−1.32	+0.72	0.516
Number of maintained relationships	4.55	−0.26	180	169.333 (13.197)	−5.93	+0.81	0.516
Number of recovered relationships	4.68	+3.39	77	79.867 (10.064)	+3.72	−0.28	0.774
Mean edge weight	3.24	−0.75	1.21202	1.260 (0.044)	+3.98	−1.09	0.323
Share of giant connected component	3.73	+2.81	0.346478	0.352 (0.014)	+1.64	−0.41	0.839
Mean clustering coefficient	3.62	+0.42	0.17454	0.179 (0.006)	+2.33	−0.67	0.645
Transitivity	2.28	−1.26	0.156234	0.158 (0.010)	+1.17	−0.18	0.968
Share of cross-type collaboration	2.08	−2.08	0.500	0.503 (0.017)	+0.53	−0.16	0.968
Share of cross-province collaboration	2.78	+0.03	0.310472	0.303 (0.015)	−2.37	+0.48	0.645
Maximum degree	0.72	−0.18	35	34.967 (0.320)	−0.10	+0.10	1.000
99th percentile of degree	3.37	−2.04	11.34	11.220 (0.410)	−1.06	+0.29	0.452

Note: The two-sided Monte Carlo *p*-value measures the tail position of an observed statistic in the simulation distribution. An MC *p*-value < 0.05 indicates a significant model fit deviation for the corresponding network feature, whereas an MC *p*-value ≥ 0.05 indicates that the current test does not detect a significant lack of fit.

4.2.2. Simulation Tests of Relationship Mechanisms

The preceding dynamic network simulations evaluated the generative validity using multidimensional statistics on the network scale, relationship events, and local clustering. The agreement on aggregate network features alone, however, cannot rule out the possibility that different micro-level relationship rules could generate similar network patterns. To distinguish the network generation fit from the evidence on relationship mechanisms, this study adds competing specifications, out-of-time prediction, conditional permutation, and matched-risk-set tests at the organization pair level, directly examining whether the key mechanism information retains stable incremental explanatory and predictive value after controlling for other observed conditions.

For statistical inference, the parametric models report two-sided p -values. The bootstrap and conditional permutation tests for competing specifications report empirical p -values in the prespecified direction, together with the 95% confidence intervals. The significance thresholds are uniformly set at $p < 0.001$, $p < 0.01$, and $p < 0.05$. The LogLoss, ROC-AUC, and PR-AUC reported jointly within the same permutation experiment are adjusted for multiple comparisons using the Holm method. The out-of-time validation and parameter freezing for the relationship models are reported in Section SB; the core conditional permutation and matched-risk-set results are shown in Table 5.

Table 5. Core results of direct validation of relationship mechanisms.

Relationship Process	Mechanism Tested	Specific Factor Tested	Core Result	p -Value	Assessment
Relationship formation	Spatial proximity	Same-province and same-city spatial information	$\Delta\text{ROC-AUC} = -0.2073$	$p < 1.0 \times 10^{-3}$	Significant
Relationship formation	Spatial proximity	Dyadic spatial information	$\Delta\text{PR-AUC} = -0.0247$	$p = 0.0015$	Significant
Relationship formation	Spatial proximity	Same province relative to cross-province	OR = 6.7385	$p = 3.11 \times 10^{-48}$	Significant
Relationship formation	Organizational-type combination	Organizational-type-combination information	$\Delta\text{ROC-AUC} = +0.0020$	$p = 0.062$	Not significant
Maintenance of existing relationships	Relationship history features	Prior collaboration history of both parties	$\Delta\text{LogLoss} = +0.0331$	$p = 0.009$	Significant
Recovery of interrupted relationships	Relationship history features	Dormancy duration, historical collaboration intensity, and relationship memory	$\Delta\text{LogLoss} = +0.0046$	$p = 0.013$	Significant
Recovery of interrupted relationships	Relationship history features	Pair-specific historical collaboration information	$\Delta\text{ROC-AUC} = -0.0254$	$p = 0.011$	Significant
Recovery of interrupted relationships	Historical collaboration intensity	Cumulative historical collaboration intensity	$\beta = 0.3768$; OR = 1.4576	$p = 0.0066$	Significant

Note: The ‘Mechanism tested’ column identifies only the mechanism being evaluated. The confidence intervals for differences in the AUC are obtained using a cluster bootstrap. A positive $\Delta\text{LogLoss}$ indicates an increase in the predictive loss after the mechanism is removed. The conditional permutation jointly reports the LogLoss, ROC-AUC, and PR-AUC and applies a Holm correction for multiple comparisons.

The initial formation results show that spatial proximity carries stable relationship-level information. Removing or permuting the spatial proximity information significantly

reduces the model's ability to identify first-time collaborations, and the matched-risk-set results likewise show a higher propensity for first-time collaboration among organization pairs located in the same province. By contrast, the independent contribution of organizational-type combinations is not statistically significant. Spatial proximity therefore exhibits a stable relationship-level association with initial relationship formation, consistent with the theoretical mechanism that proximity reduces the interaction costs, whereas the organizational type primarily functions as a control and source of heterogeneity.

Maintenance of existing relationships and recovery of interrupted relationships show pronounced dependence on the relationship history. Removing the prior collaboration history significantly worsens maintenance predictions. In the recovery process, the rolling tests, conditional permutation, and matched-risk-set results for the relationship history features are all statistically significant, indicating that prior collaboration provides stable incremental explanatory and predictive information about whether interrupted relationships are re-established.

Taken together, the model reproduces not only aggregate characteristics, such as network scale, clustering, and relationship events, but also stable and statistically significant changes associated with distinct relationship processes when specific relational information is controlled or perturbed. Spatial proximity primarily supports initial relationship formation, whereas relationship history primarily supports maintenance and recovery. The model therefore retains information on the micro-level relationship mechanisms underlying network evolution rather than merely fitting the aggregate network form.

4.3. Collaboration-State Transition Structure

4.3.1. Comparison of Continuous, Threshold, and Latent-State Models

The preceding comparisons confirm that, with parameters frozen, the dynamic network model can reproduce the overall evolution of the observed network reasonably well. Nevertheless, agreement on aggregate network characteristics may still mask substantial actor-level heterogeneity, and the network structure and resilience need to be understood through dynamic state transitions [16,30]. To assess whether changes in digital-intelligent innovation collaboration can be adequately explained by a single continuous process, this study compares dynamic linear, PSTR/LSTAR, SETAR, and lagged-threshold models.

Figure 4 and Table 6 show that PSTR/LSTAR yields the lowest AIC, whereas the dynamic linear model yields the lowest BIC. Under repeated actor resampling, the 2025 hold-out RMSE and MAE of the four model classes are also highly similar. Smooth-threshold and discrete-threshold specifications increase the functional-form flexibility, but do not produce stable and substantively meaningful out-of-sample advantages. Thus, simply adding threshold or nonlinear forms to a continuous model is insufficient to distinguish different modes of relationship organization, such as expansion, maintenance, turnover, and dormancy. The analysis therefore proceeds to latent-state models.

The collaboration intensity, partner breadth, initial formation, relationship maintenance, and relationship turnover all contain many zero values. The number of latent states can be assessed using information criteria, marginal likelihood, and related model-comparison evidence, while the latent-state structure in zero-inflated processes should also be distinguished from zero-valued features at the observation level [46–48]. State number selection and subsequent mechanism estimation are therefore separated into two steps. First, the five process indicators are converted into binary hurdle indicators, and a Bernoulli HMM is used to compare $K = 1$ through $K = 8$ and identify a state number with a clear classification structure. Once K is determined, the analysis returns to the five-dimensional continuous standardized observations and uses a Gaussian HMM to estimate the state profiles and non-homogeneous transition mechanisms. The first step serves state number

selection; the second retains information on the state intensity and supports mechanism estimation. The two steps therefore perform different statistical tasks.

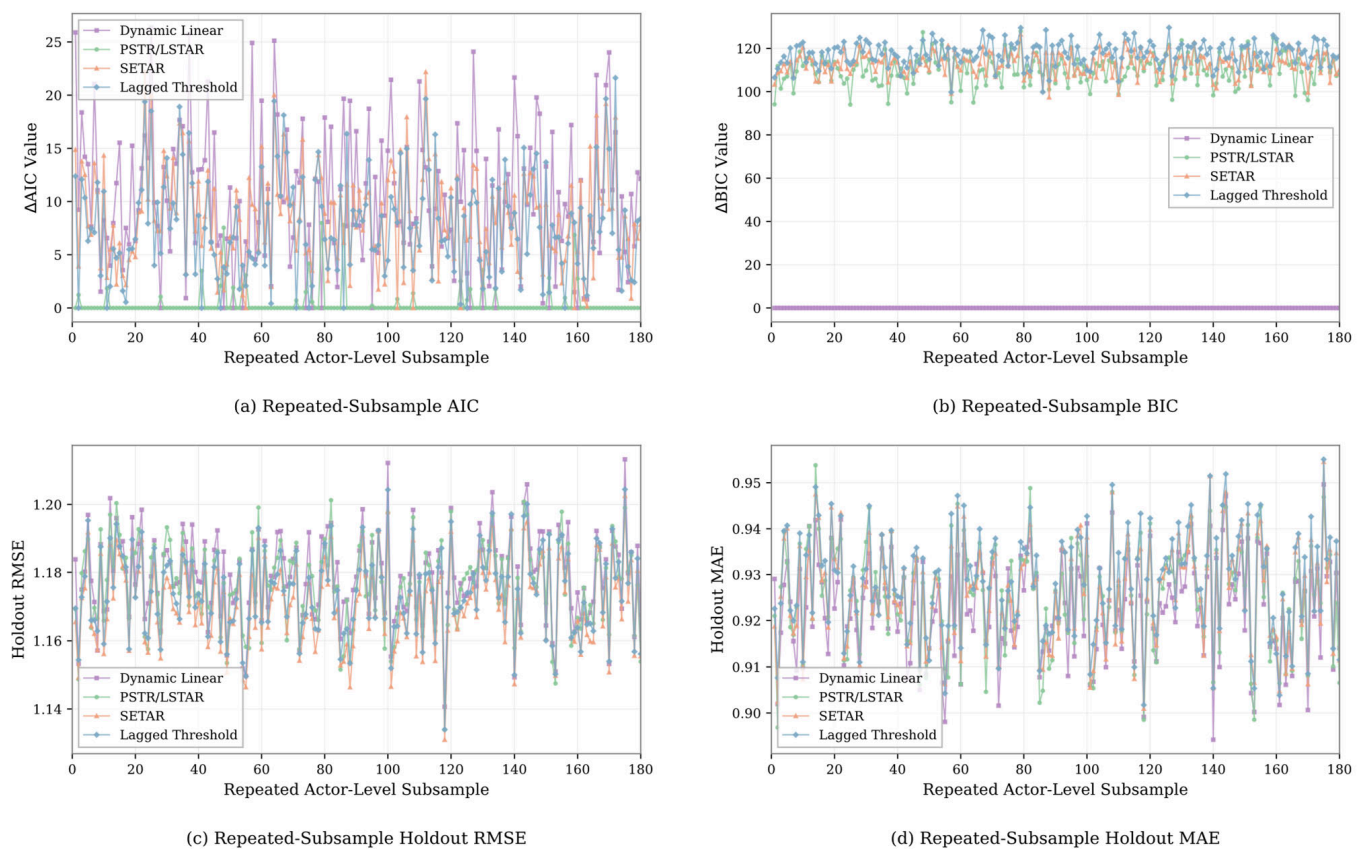


Figure 4. Repeated resampling comparison of continuous and threshold models.

Table 6. Information criterion comparison of continuous and threshold models.

Model	AIC	BIC
Dynamic Linear	12,373.89	12,499.91
PSTR/LSTAR	12,348.63	12,606.96
SETAR	12,365.30	12,617.33
Hysteresis Threshold	12,360.53	12,618.86

4.3.2. State Number Selection and Profiles of the Four States

In Table 7, BIC favors six states, whereas ICL supports four states. The four-state model simultaneously exhibits high classification certainty, reasonable state shares, and good state separation. From five states onward, the classification entropy increases markedly, and the additional states mainly represent further subdivisions of existing categories. Considering both the statistical performance and the requirements of subsequent transition interpretation, this study adopts the four-state model. The robustness to alternative state numbers using continuous indicators is reported in Section SC.1.

The four states are named according to their posterior-weighted profiles on the five observed indicators of collaboration intensity, partner breadth, new relationship formation, relationship maintenance, and relationship turnover, together with their sample shares and expected durations. All five indicators are standardized using the training sample, so positive and negative values indicate deviations above and below the overall mean, respectively.

Table 7. Bernoulli HMM state number selection based on binary hurdle indicators.

Number of States K	Maximum Log-Likelihood Log L	Number of Parameters p_K	BIC	Classification Entropy H	ICL
1	−10,097.74	5	20,236.98	0.00	20,236.98
2	−6984.17	13	14,076.25	7.53	14,091.31
3	−6035.36	23	12,261.64	272.71	12,807.05
4	−5003.46	35	10,297.45	16.45	10,330.34
5	−4831.31	49	10,069.36	228.57	10,526.49
6	−4658.62	65	9856.79	272.80	10,402.40
7	−4653.07	83	9995.10	353.92	10,702.95
8	−4641.58	103	10,138.14	447.07	11,032.29

Relationship instability accounts for 7.05% of observations and has an expected duration of 1.00 year. Collaboration intensity, partner breadth, new relationship formation, and relationship maintenance are all below their overall means, with standardized values of −0.642, −0.652, −0.587, and −0.327, respectively, whereas relationship turnover reaches 1.341. Low-cooperation dormancy accounts for 57.06% and has an expected duration of 4.15 years. Its first four indicators are likewise low, but relationship turnover falls to −0.788. Although both states exhibit low levels of collaboration, relationship instability represents short-term relational disruption under high turnover, whereas low-cooperation dormancy reflects a persistent lock-in characterized by low activity and low turnover.

Expansion and reconfiguration accounts for 24.60% of observations and has an expected duration of 1.64 years. Collaboration intensity, partner breadth, and new relationship formation are 0.879, 0.968, and 1.115, respectively, while relationship maintenance remains at −0.327 and relationship turnover is 1.341. This state is therefore dominated by expansion through new relationships and adjustment of partner configurations, with existing relationships not yet fully consolidated. Stable deepening accounts for 11.28% and has an expected duration of 2.80 years. Collaboration intensity and partner breadth rise to 1.729 and 1.593, relationship maintenance reaches 2.569, and relationship turnover falls to 0.222, indicating high collaboration intensity, broad partner reach, and persistent relationships. These empirical differences provide the basis for naming the four states relationship instability, low-cooperation dormancy, expansion and reconfiguration, and stable deepening.

4.3.3. Key State-Transition Associations and Visualization

After the four states are identified, the organizational attributes and relational conditions are incorporated into the state-transition probability equations to examine their conditional associations with the actors' relative propensity to move from the current state to the next state. In latent-state models, the covariates in transition probabilities can identify systematic differences in state-transition propensities across organizational and relational conditions [49], while the time-varying conditions can characterize the dynamic associations in transition probabilities across contexts [50].

The established state profiles are held fixed. Each organization's state posterior probabilities are used as soft weights, and the model includes innovation capability, relational capital, technological complementarity, local clustering, partner diversity, cross-type and cross-province collaboration, relationship turnover rate, and actor type. The bootstrap resamples entire organizations while retaining their complete annual sequences. All 60 non-intercept transition effects are jointly subjected to a Benjamini–Hochberg FDR correction. Table 8 reports only the six core associations that survive the global BH-FDR correction; the complete coefficients, intervals, and multiplicity-adjusted results are provided in Section SC.2.

Table 8. Key state-transition coefficients from the non-homogeneous latent-state transition model.

Transition Category	State Transition	Key Variable	Coefficient (Bootstrap <i>p</i> -Value; Global BH-FDR <i>q</i> -Value)
Collaboration activation	Low-cooperation dormancy → expansion and reconfiguration (relative to remaining dormant)	Innovation capability	0.357 (<i>p</i> = 0.002; <i>q</i> = 0.024)
Collaboration activation	Low-cooperation dormancy → expansion and reconfiguration (relative to remaining dormant)	Lagged relational capital	0.698 (<i>p</i> = 0.002; <i>q</i> = 0.024)
Collaboration activation	Low-cooperation dormancy → expansion and reconfiguration (relative to remaining dormant)	University/research institution status	0.743 (<i>p</i> = 0.002; <i>q</i> = 0.024)
Collaboration deepening	Expansion and reconfiguration → stable deepening (relative to remaining in expansion)	Lagged technological complementarity	0.634 (<i>p</i> = 0.002; <i>q</i> = 0.024)
Suppression of instability	Expansion and reconfiguration → relationship instability (relative to stable deepening)	Lagged relational capital	−2.056 (<i>p</i> = 0.002; <i>q</i> = 0.024)
Maintenance of stability	Stable deepening → expansion and reconfiguration (relative to remaining stable)	Lagged relational capital	−2.019 (<i>p</i> = 0.004; <i>q</i> = 0.040)

Note: The table reports only the effects with *q* < 0.05 after the global BH-FDR correction across all 60 non-intercept effects. The parentheses contain, in order, the two-sided actor-block bootstrap *p*-value and the global BH-FDR *q*-value. The bootstrap uses 500 resamples. The complete results are reported in Section SC.2.

To translate the coefficients surviving the global BH-FDR correction into transition conditions, this study constructs core relative transition propensities according to the actual reference outcome for each pairwise Logit equation. State numbers 1 through 4 denote relationship instability, low-cooperation dormancy, expansion and reconfiguration, and stable deepening, respectively; IA, RC, UR, and TC denote innovation capability, lagged relational capital, the university/research institution dummy, and lagged technological complementarity. The terms to the right of the colon in each expression denote the reference transition outcome:

$$\eta_{23:22} = \alpha_{23:22} + 0.357IA + 0.698RC + 0.743UR \quad (10a)$$

$$\eta_{34:33} = \alpha_{34:33} + 0.634TC \quad (10b)$$

$$\eta_{31:34} = \alpha_{31:34} - 2.056RC \quad (10c)$$

$$\eta_{43:44} = \alpha_{43:44} - 2.019RC \quad (10d)$$

Using Equation (10a) as an example, $\eta_{23:22}$ compares the transition from low-cooperation dormancy to expansion and reconfiguration with remaining dormant; $\eta_{31:34}$ in Equation (10c) compares the transition from expansion and reconfiguration to relationship instability with the transition to stable deepening. For any pairwise state comparison, the linear predictor is converted through the Logit function into the conditional probability of the target transition:

$$\Pr(S_{i,t+1} = k \mid S_{i,t} = j, S_{i,t+1} \in \{k, b_j\}, x_{i,t}) = \frac{\exp(\eta_{jk,it})}{1 + \exp(\eta_{jk,it})} \quad (11)$$

Among all 60 non-intercept effects, six remain at *q* < 0.05 after the global BH-FDR correction. Subsequent analysis of transition associations is based only on these results; effects that are nominally significant or supported only within a single transition comparison are

not treated as mechanism evidence in the main text. The complete results are reported in Section SC.2.

The six results are concentrated on four transition paths. Innovation capability, lagged relational capital, and university/research institution status are positively associated with a transition from low-cooperation dormancy to expansion and reconfiguration; lagged technological complementarity is positively associated with a transition from expansion and reconfiguration to stable deepening; and lagged relational capital is negatively associated with both a transition from expansion and reconfiguration to relationship instability and a transition from stable deepening to expansion and reconfiguration. These coefficients represent relative transition propensities that are conditional on the included covariates and do not constitute estimates of causal effects.

Figure 5 places the four collaboration states and their typical transitions within a common state landscape. The horizontal axis, the collaboration foundation, summarizes the innovation capability, relational capital, and technological complementarity; the vertical axis, the collaboration complexity, summarizes the partner diversity, cross-domain collaboration, and relationship turnover; and the surface height U represents the composite collaboration utility. The colored dashed lines illustrate typical transition paths and positional relationships among the four states. Substantive judgments about transition associations remain based on the transition coefficients supported after the global BH-FDR correction.

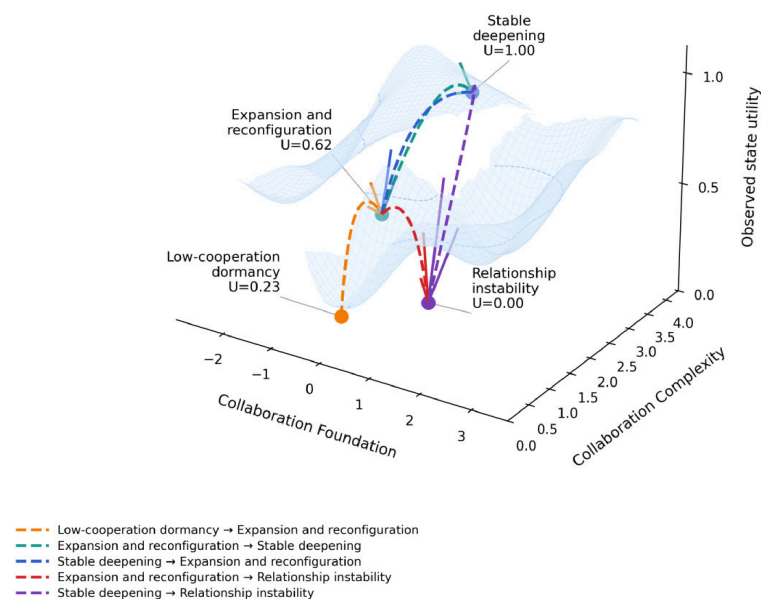


Figure 5. Landscape of key transitions among digital-intelligent innovation collaboration states.

4.3.4. Fit of the Collaboration-State Transition Structure to Observed Data

To examine the empirical support for the state landscape in the fully held-out period, Figure 6 maps the 2025 organizational observations onto the same raw state utility surface used in Figure 5. The horizontal axis, vertical axis, and utility surface retain the definitions from Figure 5, and the orange points indicate the actual locations of held-out organizations in the state space. The fully held-out 2025 sample contains 1389 actor-year observations, of which 1310 organization points fall within the plotting range and empirical support domain of the original surface. The hold-out sample exhibits differentiated clustering in low-utility, transitional, and higher-utility regions, consistent with the empirical distributional structure represented by the four collaboration states.

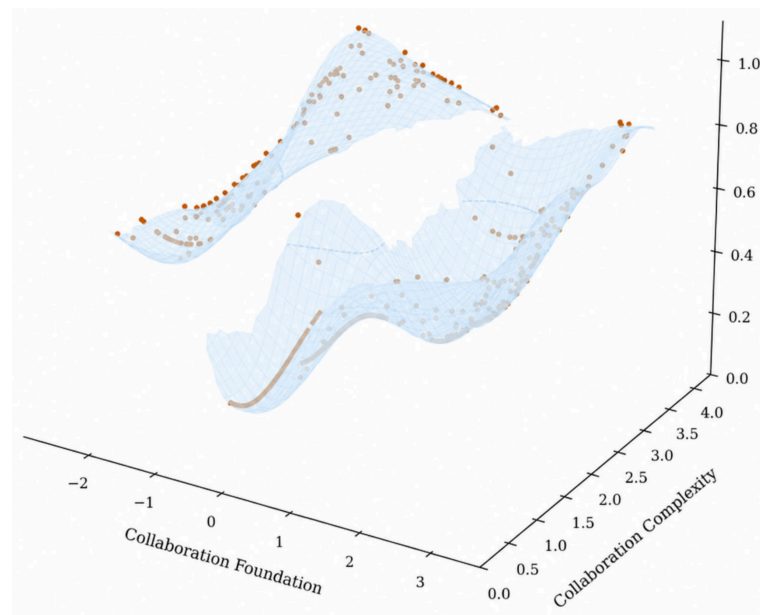


Figure 6. Distribution of organizational collaboration states in 2025.

Table 8 and Equations (10a)–(10d) report the local transition associations that receive statistical support after the global BH-FDR correction; the overall transition structure is evaluated through a generative fit test. After recursively generating 500 sets of state trajectories from the estimated transition probabilities, the model-generated results are compared with the observed state changes. As shown in Figure 7, the model-generated annual state shares and annual state-transition flows are concentrated close to the line of perfect agreement. Lin’s concordance correlation coefficients for the annual state shares and transition flows are 0.9985 and 0.9914, respectively [51,52]. The fitted lines largely overlap with the line of perfect agreement, and most observed values fall within the generative intervals. These results indicate that the estimated transition structure can reproduce the overall state composition and transition flows in the sample, providing empirical support for a subsequent resilience measurement and governance simulation under the frozen transition structure.

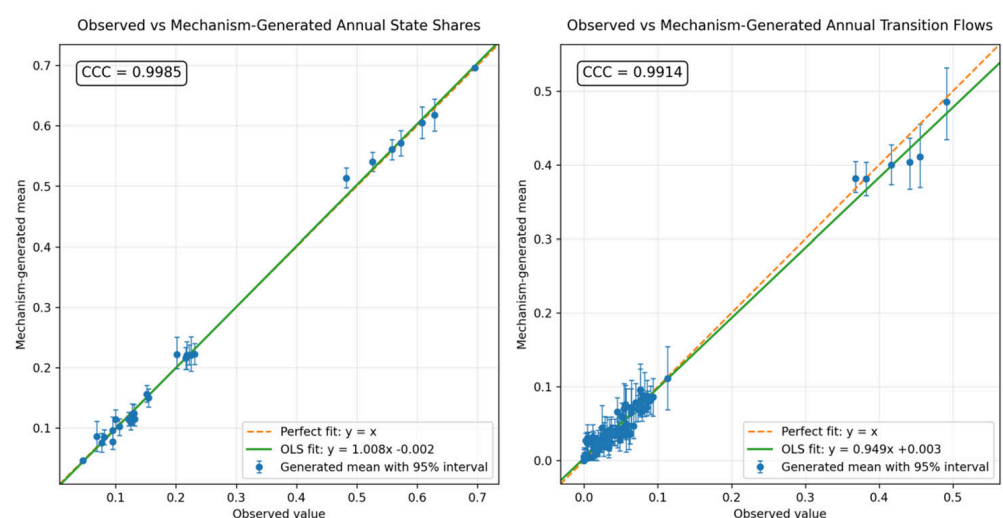


Figure 7. Comparison of observed state changes and mechanism-generated results.

4.4. Predictive Validity of the Collaboration Resilience Measure

To evaluate the ability of collaboration resilience to discriminate actual relationship recovery, the criterion validity test freezes all data-driven parameters using information available only through 2023 and validates the measure against relationship shocks in 2024 and recovery outcomes in 2025. Among 277 shocked actors, the strict R_{2023} yields ROC-AUC values of 0.6447 for recovery with the same key partner and 0.6545 for recovery of at least one interrupted relationship, and a Spearman correlation of 0.1800 with the continuous recovery ratio ($p = 0.0026$). The paired bootstrap comparisons with the original specification do not indicate a systematic decline in the validity metrics. These results support the idea that collaboration resilience retains out-of-time discrimination for recovery under strict temporal freezing. The complete experimental design, confidence intervals, and stratified results are reported in Section SD.

5. Construction of an Improved Double DQN Model for Digital–Intelligent Innovation Collaboration Resilience Governance

5.1. Markov Decision Modeling of the Governance Problem

The preceding analysis identifies the formation, maintenance, and recovery of digital–intelligent innovation collaboration relationships and estimates the collaboration states and their transition structures. These results, however, do not directly determine the action timing, governance targets, or action combinations under resource constraints. This study therefore formulates collaboration resilience governance over 2019–2025 as a finite-horizon sequential decision problem and develops an Organization–System Resilience-based Double Deep Q-Network (OSR-DDQN), implemented in Python 3.11 using the PyTorch 2.2.2 framework, with Double DQN as the learning backbone. The model uses a dual-branch state representation to retain information on both the overall system and priority governance groups, and combines Dueling value decomposition, dynamic costs, same-source counterfactual rewards, and a full-trajectory return target to address action comparison, resource constraints, the natural evolution baseline, and interannual effects.

(1) Markov decision process and state space: To translate the frozen state-transition structure and relationship process rules into an executable sequential decision problem, this study represents collaboration resilience governance over 2019–2025 as a finite-horizon Markov decision process [53]:

$$\mathcal{M} = \langle S, A, P, \mathcal{R}, \gamma \rangle \quad (12)$$

Here, S denotes the state space, A is the action space, P denotes the state-transition rules frozen from the calibration results in Sections 3 and 4, R is the reward function, and γ is the discount factor. The year 2018 is used as the initial state, and 2019–2025 constitute seven annual decision steps. In period t , the agent observes state s_t and selects action a_t . Without re-estimating the frozen parameters, the environment generates the next-period state s_{t+1} and returns reward r_t according to the baseline transition probabilities, relationship process rules, action effectiveness, and budget constraints.

The state space is constructed by translating the preceding state identification, transition probabilities, and resilience measurement into governance information rather than by arbitrarily expanding the input variables. Using only the average system resilience may conceal concentrated risks among actors in relationship instability and low-cooperation dormancy, whereas using only priority group indicators may neglect the overall network structure and resource constraints. To retain both types of information, the state representation is divided into an overall system branch and a priority governance group branch. The overall system branch contains 11 state, risk, relational process, and structural features derived from the preceding measurement and identification results, together with

three budget and time-constraint variables added for sequential decision making, for a total of 14 dimensions. The priority governance group branch further aggregates five governance-need features from the state distribution, transition probabilities, relational process reserve, and network structure indicators. Together, the two types of information form a 19-dimensional state vector; the specific variables, sources, and governance meanings are shown in Table 9.

$$s_t = [s_t^{sys}, s_t^{grp}], s_t^{sys} \in \mathbb{R}^{14}, s_t^{grp} \in \mathbb{R}^5 \quad (13)$$

Table 9. Composition and sources of OSR-DDQN state variables.

ID	State Variable	Source or Construction	Meaning
S1	System collaboration resilience	Collaboration resilience measure in Equation (9), Section 3.4	Composite resilience level of the collaboration system in the current period
S2	Share in relationship instability	Posterior shares of the four states in Section 4.3	Proportion of actors in the relationship instability state
S3	Share in low-cooperation dormancy	Posterior shares of the four states in Section 4.3	Proportion of actors in the low-cooperation dormancy state
S4	Share in expansion and reconfiguration	Posterior shares of the four states in Section 4.3	Proportion of actors in the expansion-and-reconfiguration state
S5	Share in stable deepening	Posterior shares of the four states in Section 4.3	Proportion of actors in the stable-deepening state
S6	System deterioration risk	Calculated from state-transition probabilities in Section 4.3	Composite risk of not entering expansion and reconfiguration or stable deepening in the next period
S7	System recovery potential	Calculated from state-transition probabilities in Section 4.3	Conditional probability that low-functioning actors transition to expansion and reconfiguration or stable deepening
S8	Relational process reserve	Relational process reserve in Equation (8), Section 3.4	Composite capacity for relationship resistance, recovery, reconfiguration, and continuity
S9	Evidence reliability	Evidence reliability treatment in Section 3.4	Extent to which relationship–event evidence supports process capacity measurement
S10	Network embeddedness	Constructed from network structure indicators in Section 3.1	Existing connectivity and relational embeddedness of the system
S11	Partner diversity	Constructed from partner structure variables in Section 4.3	Diversity of cross-type and cross-regional partner combinations
S12	Remaining budget ratio	Decision environment definition in Section 5.1	Current remaining budget as a proportion of the initial budget
S13	Budget ratio	Decision environment definition in Section 5.1	Budget ceiling in the current scenario relative to the maximum feasible action cost in the baseline
S14	Decision-time progress	Decision environment definition in Section 5.1	Relative position of the current year within the 2019–2025 decision horizon
G1	Share of low-functioning actors	S2 + S3	Combined proportion of actors in relationship instability and low-cooperation dormancy

Table 9. Cont.

ID	State Variable	Source or Construction	Meaning
G2	Composite risk of priority governance groups	$0.55 \times S2 + 0.35 \times S3 + 0.10 \times S6$	Jointly reflects instability, dormancy, and further deterioration risk
G3	Recovery gap of priority governance groups	$1 - S7$	Insufficient capacity of low-functioning actors to transition toward favorable states
G4	Relational process reserve gap of priority governance groups	$1 - S8$	Insufficient capacity for relationship resistance, recovery, reconfiguration, and continuity
G5	Network structure leverage of priority governance groups	$G1 \times [0.5 \times S10 + 0.5 \times (1 - S11)]$	Joint representation of the scale of low-functioning actors and the room for structural adjustment
	Total	14-dimensional overall system state + 5-dimensional priority governance group features	Forms the complete OSR-DDQN state input

(2) Actions and their effects on state transitions: The action space is constructed as $A = \{0, 1, 2, 3, 4\}$ based on the characteristics of the four states, governance objectives, and the frozen transition structure. Here, $a = 0$ denotes no intervention, under which the system evolves naturally according to the frozen transition structure while retaining the relational capital accumulated by previous actions; $a = 1$ denotes recovery for instability-state actors, targeting actors in relationship instability or with insufficient recovery capacity and strengthening relationship maintenance and recovery; $a = 2$ denotes collaboration activation for dormant actors, targeting actors in low-cooperation dormancy or with insufficient recovery potential and increasing their opportunity to enter expansion and reconfiguration; $a = 3$ denotes relationship consolidation for expansion-state actors, targeting actors in expansion and reconfiguration but with insufficient stability, facilitating the consolidation of new relationships and increasing the opportunity to enter stable deepening; and $a = 4$ denotes coordinated multi-mechanism intervention, jointly applying the foregoing governance modules when multiple types of risk are concentrated simultaneously.

The real-world instruments in Table 10 illustrate the possible implementation forms corresponding to each type of governance action. What the model actually selects are governance mechanisms, such as recovery, activation, and consolidation, rather than any specific policy instrument. Coordinated multi-mechanism intervention denotes the combined application of multiple governance modules to different target groups; its coverage therefore reflects the union of the target groups and does not imply that a single measure simultaneously acts on all covered actors.

Let the baseline probability that actor i transitions from state j to state k in period t be $P(j \rightarrow k | i, t)$, let Δ^a denote the transition-direction matrix for action a , and let $m_{a,t}$ denote the realized action multiplier jointly determined by target matching and repetition decay. The intervention modifies the log weights of the target states and renormalizes them to obtain the post-action transition probabilities:

$$P^{(a)}(j \rightarrow k | i, t) = \frac{P(j \rightarrow k | i, t) \exp(m_{a,t} \Delta_{jk}^a)}{\sum_{\ell=1}^4 P(j \rightarrow \ell | i, t) \exp(m_{a,t} \Delta_{j\ell}^a)} \quad (14)$$

Table 10. Operational definitions of governance actions and corresponding real-world implementation.

Action	Operational Definition and Target Group	Potential Corresponding Governance Instruments	Typical Implementing Actors	Resource and Implementation Characteristics
A0: No intervention	No additional governance effect; the system evolves according to the established state-transition structure	Routine monitoring; deferral of additional support	No additional dedicated implementing actor	No additional action cost
A1: Recovery for instability-state actors	Targets actors in relationship instability or with insufficient recovery capacity and promotes maintenance of existing relationships and recovery of interrupted relationships	Project coordination, relationship repair, joint validation, and public technology platform support	Science and technology authorities, project-leading organizations, and public technology platforms	Concentrated targets and relatively limited coverage
A2: Collaboration activation for dormant actors	Targets actors in low-cooperation dormancy or with insufficient recovery potential and promotes the re-establishment of collaborative relationships	Industry–university–research matching, partner search, cross-regional technology matching, and joint project organization	Science and technology authorities, industrial parks, technology-transfer organizations, and public service platforms	A relatively large pool of potential targets and higher coverage costs
A3: Relationship consolidation for expansion-state actors	Targets actors in expansion and reconfiguration but with insufficient stability and promotes consolidation of new relationships and stable deepening	Technology matching, joint validation, project coordination, and knowledge-interface management	Collaboration-leading organizations, project-management bodies, and technology platforms	Concentrated targets, with emphasis on improving relationship quality
A4: Coordinated multi-mechanism intervention	Combines recovery, activation, and consolidation modules across different target groups when multiple risks coexist	Combined deployment of instruments corresponding to A1–A3	Coordination among multiple governance actors	Broad coverage, with the highest coordination complexity and relative cost

Action effectiveness varies with the degree of match between current governance needs and the action target, and decays with repeated use of the same action; the costs do not decline proportionally with effectiveness, preventing the model from obtaining an unrealistic low-cost advantage by repeatedly selecting actions that have already become ineffective. The realized action multiplier is defined as

$$m_{a,t} = \frac{0.45 + 1.05v_{a,t}}{1 + \delta_a n_{a,t-1}}. \quad (15)$$

Here, $v_{a,t} \in [0, 1]$ denotes the match score between the action and current risks and capacity gaps, $n_{a,t-1}$ is the cumulative number of times that action has been used before period t , and δ_a is the repetition decay coefficient fixed before training.

(3) Dynamic costs and budget constraints: If each action type was assigned a fixed cost that did not vary with the state conditions, the model could not reflect the annual differences in the scale of governance targets and implementation conditions. To improve the interpretability of resource constraints, this study expresses the action cost as the prod-

uct of coverage scale, implementation intensity, and coordination complexity, producing a dynamic dimensionless cost:

$$C_{a,t} = q_{a,t} \iota_{a,t} K_a \iota_{a,t} = \min \left\{ 1, \frac{\iota_a (0.45 + 1.05 \iota_{a,t})}{1.50} \right\}, K_a = \frac{M_a}{4} \quad (16)$$

The coverage ratio in action costs is determined period by period from the current shares of the four states and the baseline transition structure. For combined actions, coverage is calculated over the union of target groups to avoid charging the same actor repeatedly. The full-horizon budget is defined relative to the maximum feasible action cost in each year of the baseline environment:

$$B_\rho = \rho \sum_{t=1}^T \max_{a \in \{1,2,3,4\}} C_{a,t}^{(0)}, \rho \in \{0.30, 0.50, 0.70\}, b_{t+1} = b_t - C_{a,t} \quad (17)$$

The budget ratio in Equation (17) represents the full-horizon budget relative to the maximum feasible action cost in the baseline environment. To prevent a single year's intervention from inducing a state jump outside the historically supported range, the model further imposes a constraint on the total variation distance relative to the same-source no-intervention trajectory:

$$TV \left(\tilde{p}_{t+1}^{(a)}, p_{t+1}^{(0)} \right) = \frac{1}{2} \sum_{k=1}^4 \left| \tilde{p}_{k,t+1}^{(a)} - p_{k,t+1}^{(0)} \right| \leq \kappa \quad (18)$$

The threshold is determined from empirical quantiles of natural year-to-year state changes and restricted to the interval 0.05–0.30; the maximum realized value in the formal scenarios is 0.2992. This constraint limits only the incremental intervention effect in the current period and does not repeatedly offset the relational capital accumulated by earlier actions, thereby balancing continuity of policy effects with the empirical feasibility of state changes. The diagnostics for coverage, intensity, and cost across governance actions are reported in Table 11. Governance actions are thus jointly constrained by the budget, target coverage, coordination complexity, state matching, repeated implementation decay, and the magnitude of annual state change; broad-coverage actions cannot be implemented without cost or repeated indefinitely.

Table 11. Dynamic cost diagnostics.

Action	Governance Action	Mean Target Coverage	Mean Intensity	Complexity	Mean Cost	Cost Range
1	Recovery for instability-state actors	0.111	0.459	0.50	0.026	0.015–0.033
2	Collaboration activation for dormant actors	0.820	0.628	0.50	0.258	0.219–0.306
3	Relationship consolidation for expansion-state actors	0.154	0.466	0.50	0.036	0.027–0.048
4	Coordinated multi-mechanism intervention	0.865	0.423	1.00	0.366	0.336–0.392

Table 11 shows that the differences in the dynamic costs are driven primarily by the target group size and the number of governance modules. The mean cost of collaboration activation for dormant actors and coordinated multi-mechanism intervention is 0.258 and 0.366, respectively, substantially higher than the 0.026 and 0.036 associated with recovery for instability-state actors and relationship consolidation for expansion-state actors. Broad-coverage actions therefore have a cost–benefit basis only when governance needs are

sufficiently concentrated, and their use is constrained by the model's dynamic cost and budget mechanisms.

5.2. Adaptive Improvements and Learning Mechanisms of OSR-DDQN

DQN can approximate state-action value functions in high-dimensional states; Double DQN mitigates Q-value overestimation by separating action selection from value evaluation; and Dueling DQN further decomposes the state value and action advantage [33–35,54]. The improvements in this study address three specific issues: the state input must retain information on both the overall system and priority governance groups; the reward must distinguish policy effects from natural evolution while incorporating resource costs; and the learning target must accommodate intertemporal accumulation of governance effects. The corresponding state representation, counterfactual reward, and intertemporal learning target are specified below.

(1) Same-source counterfactual reward: The reward function must distinguish policy effects from natural system evolution. Directly using pre–post differences would incorrectly count natural recovery as a policy benefit and could also misattribute system-wide deterioration to action failure. Therefore, each intervention trajectory is paired with a same-source no-intervention trajectory sharing the same initial network, baseline parameters, and stochastic process, and the difference between the two is used to define the policy-relative advantage:

$$\Delta R_t = R_t^I - R_t^B \quad (19)$$

R_t^I and R_t^B denote the collaboration resilience in period t for the intervention trajectory and the same-source no-intervention trajectory, respectively. On this basis, the reward function jointly incorporates the change and level of relative advantage, terminal resilience gap reduction, dynamic cost, structural imbalance, and an invalid action penalty:

$$r_t = 0.30(\Delta R_{t+1} - \Delta R_t) + 0.08\Delta R_{t+1} + 1.251_{\{t=T\}} \frac{\Delta R_T}{1-R_T^B} - 0.08 \frac{C_{a,t,t}}{B_\rho} - 0.14\Omega_t - 0.021_{\{a_t \notin A_t\}} \quad (20)$$

The structural penalty includes additional costs for abnormal increases in state concentration, losses in partner diversity, and consecutive repetition of the same action beyond a threshold. It is designed to restrict two types of strategies: temporarily raising the average resilience by concentrating resources and repeatedly adding interventions while the effects of prior actions remain active. The relative advantage defined here represents the difference between the action and no-action trajectories under the same simulated stochastic process and is not equivalent to a causal treatment effect identified from real-world intervention data.

To summarize the key design constants in the governance environment that are not directly estimated from the sample, Table 12 reports the baseline settings and functions of repetition decay, dynamic costs, and the reward function.

The foregoing constants mainly determine the relative scales of governance effects, resource costs, and multi-objective rewards. Their function is to translate theoretical mechanisms into a computable reinforcement learning environment, not to interpret the specific numerical values as empirical estimates of real-world policy effects or fiscal costs. Defining reward preferences, action costs, and the feasible action set through design parameters and examining the result stability through repeated experiments and sensitivity analyses is common in multi-objective reinforcement learning and simulation-based policy optimization [55]. Local sensitivity tests for the 13 fixed design constants are reported in Section SH. Within the prespecified perturbation ranges, the main performance and

collaboration resilience conclusions remain stable, indicating that the main text conclusions do not depend on the exact value of any single constant.

Table 12. Key design constants in the reinforcement learning governance environment and their rationale.

Parameter Category	Parameter	Baseline Value	Rationale
Repeated action decay	$\delta 1$	0.65	Represents diminishing marginal effectiveness of repeated governance
	$\delta 2$	0.80	Same as above
	$\delta 3$	0.75	Same as above
	$\delta 4$	0.95	Same as above
Dynamic cost	Base term	0.45	Adjusts the baseline resource burden of an action
	Response term	1.05	Adjusts cost responsiveness to implementation intensity
	Normalization term	1.50	Places the costs of different governance actions on a common scale
Reward function	Path-gain weight	0.30	Retains resilience improvement along the governance trajectory
	Advantage-level weight	0.08	Retains persistent advantage relative to the no-intervention baseline
	Terminal-gap weight	1.25	Emphasizes the objective of terminal resilience improvement
	Cost weight	0.08	Incorporates governance resource consumption
	Structural-constraint weight	0.14	Constrains short-term improvements generated through structural imbalance
	Invalid action penalty	0.02	Discourages ineffective or unnecessary repeated actions

(2) Network architecture improvements based on Double DQN: OSR-DDQN uses Double DQN as its value-learning backbone. The online network selects the next action, while the target network evaluates its value, thereby reducing Q-value overestimation induced by the maximization operator in DQN. On this basis, the architectural improvements focus on dual-branch state encoding and Dueling value decomposition.

Although Double DQN can mitigate value overestimation, a single encoding network may compress the hierarchical distinction between overall system information and information on priority governance groups. When the system-average state is favorable, concentrated risks among low-functioning actors may be diluted; conversely, excessive focus on risk groups may overlook the overall network structure and budget conditions. OSR-DDQN therefore employs separate encoders for the overall system state and priority governance group features. The system branch takes 14 state variables as inputs and uses two fully connected layers with 64 and 32 neurons to extract information on the overall resilience, state transitions, relational processes, network structure, and decision constraints. The group branch takes five governance-need features as inputs and uses two fully connected layers with 32 and 16 neurons to emphasize the scale, risk, and capacity gaps of low-functioning actors. Both encoders use ReLU activations, and their outputs are concatenated into a 48-dimensional joint state representation:

$$h_t = \left[f_{\text{sys}}(s_t^{\text{sys}}), f_{\text{tar}}(s_t^{\text{tar}}) \right] \in \mathbb{R}^{48} \quad (21)$$

Figure 8 illustrates the OSR-DDQN workflow from state input and action selection to environmental feedback and parameter updating. After dual-branch encoding, the overall system state and priority governance group features enter the Dueling value-and-advantage decomposition. A budget-based action mask first removes infeasible actions, after which an ϵ -greedy policy selects among the feasible actions. The outcome is generated by the frozen transition environment, compared with the same-source no-intervention

baseline, written to the experience replay buffer, and then used to compute the Double DQN target and update model parameters.

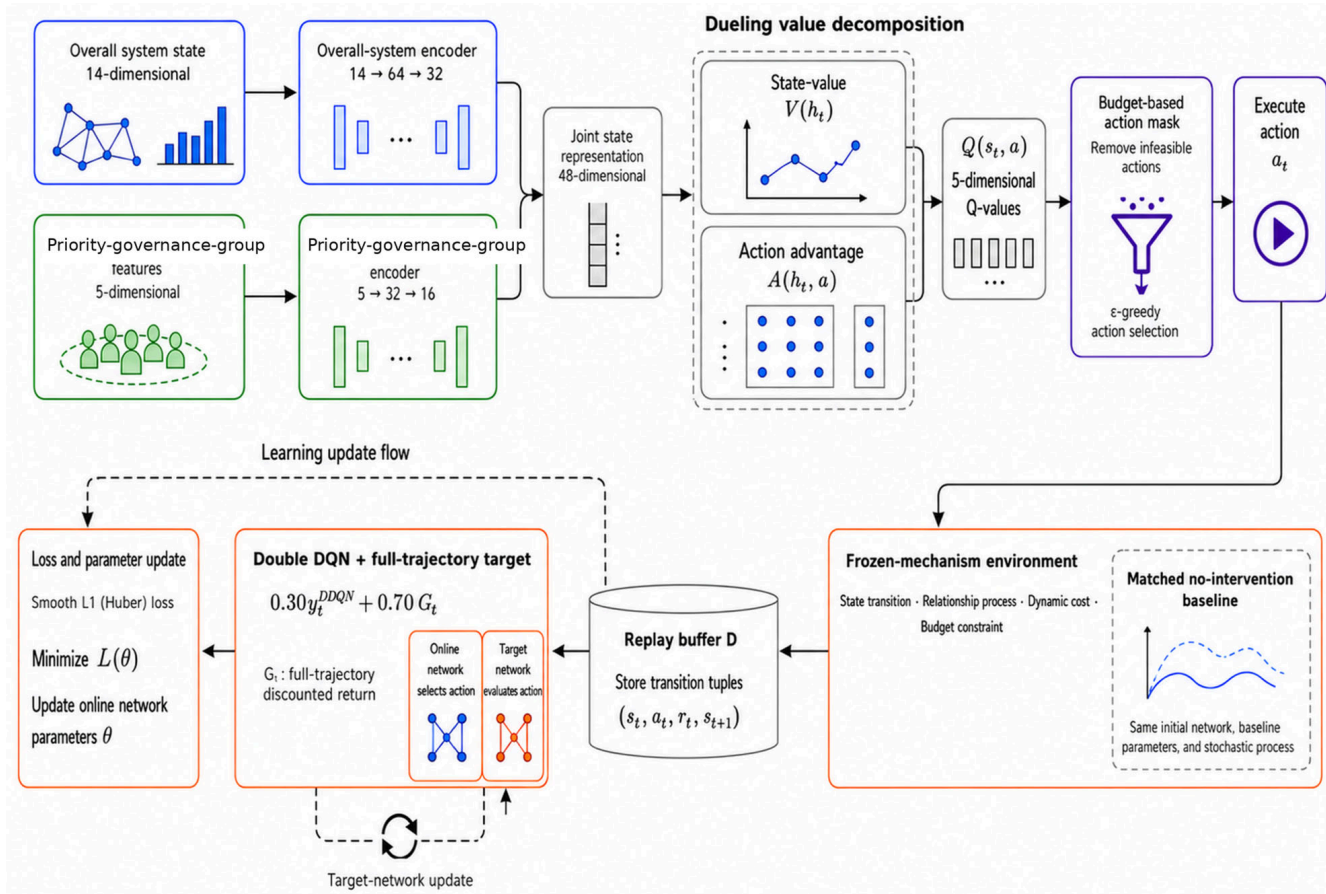


Figure 8. Decision and learning framework of OSR-DDQN.

To improve discrimination when the values of different governance actions are close, the joint state representation is further passed through a Dueling architecture that separately estimates the baseline value $V(s)$ of the current state and the advantage $A(s, a)$ of each action relative to that state. The final Q-value is expressed as

$$Q(s_t, a; \theta) = V(h_t; \theta_V) + A(h_t, a; \theta_A) - \frac{1}{|\mathcal{A}|} \sum_{a' \in \mathcal{A}} A(h_t, a'; \theta_A) \quad (22)$$

The Dueling architecture separately estimates the state value and action advantage, reducing the extent to which the overall level of the system state masks differences among actions when several action values are similar, and thereby compares the incremental value of each action relative to the current state.

Figure 9 further details the hierarchy of the dual-branch network. The system branch encodes $14 \rightarrow 64 \rightarrow 32$ dimensions, while the priority governance group branch encodes $5 \rightarrow 32 \rightarrow 16$ dimensions; the two are concatenated into a 48-dimensional joint state representation. The state-value stream then uses 24 hidden units to output a scalar state value, while the action-advantage stream uses 32 hidden units to output five action advantages, which are combined according to Equation (22) to obtain the Q-value for each action. Figure 8 emphasizes the end-to-end decision-and-learning loop, whereas Figure 9 corresponds to the internal network hierarchy and dimensional configuration.

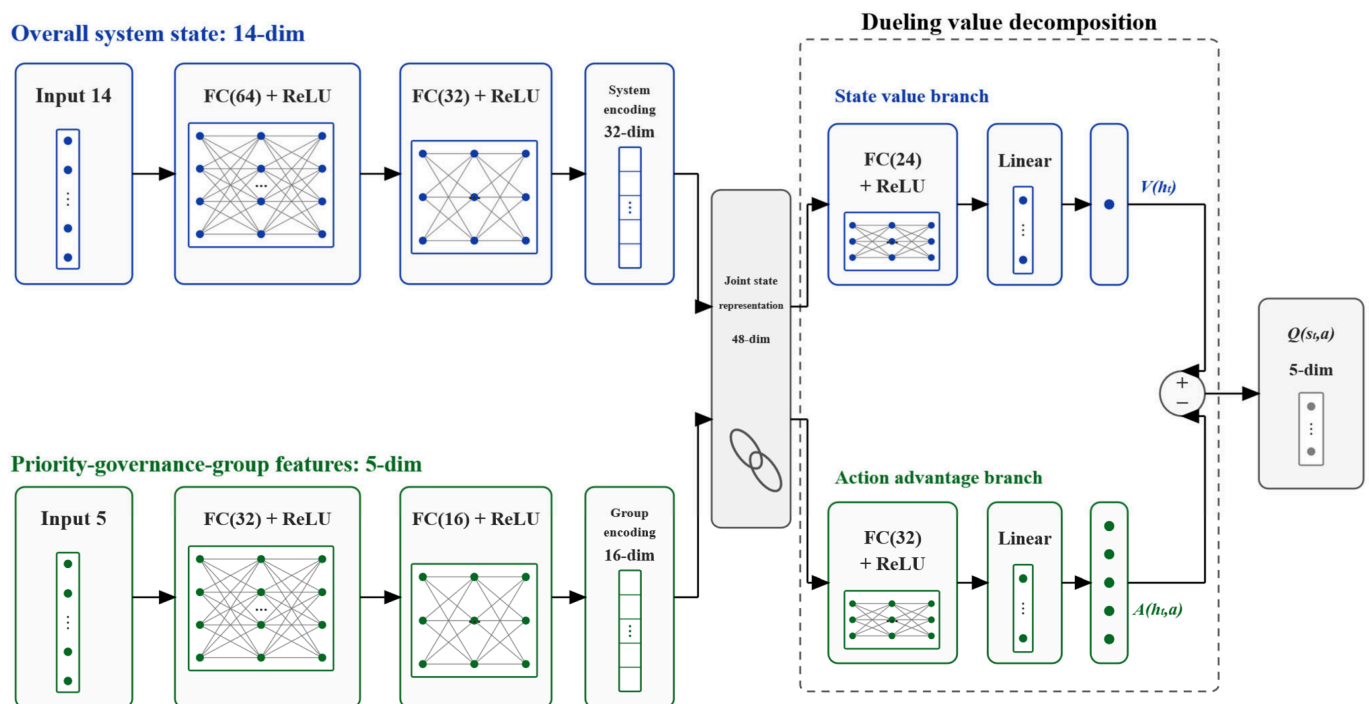


Figure 9. Dual-branch network architecture of OSR-DDQN.

(3) Full-trajectory return target and loss function: The one-step temporal-difference update in Double DQN can recursively propagate future value, but the present task contains only seven annual decision steps, and the benefits of relationship repair and network reconfiguration may emerge in later years. If learning relies solely on a one-step target, the late benefits generated by early actions must propagate backward through multiple TD updates. To strengthen this intertemporal credit signal, OSR-DDQN combines the discounted full-trajectory return with the Double DQN temporal-difference target:

$$G_t = \sum_{\tau=t}^T \gamma^{\tau-t} r_{\tau} \quad (23)$$

$$\hat{y}_t = 0.70G_t + 0.30y_t^{\text{DDQN}} \quad (24)$$

$$L(\theta) = \mathbb{E}[\ell_{\text{Huber}}(Q(s_t, a_t; \theta) - \hat{y}_t)] \quad (25)$$

The full-trajectory return and one-step temporal-difference target receive weights of 0.70 and 0.30, respectively, and optimization uses the Smooth L1 (Huber) loss. The remaining training parameters, reward scaling, and target-network update settings are reported in Table 13.

Table 13. Formal training parameters selected on the validation set.

Parameter	Value	Parameter	Value
Candidate configuration/training duration	H06/20	Episodes per epoch	180
Episodes per independent training run	3600	Discount factor γ	0.95
Learning rate	0.001	Optimizer	AdamW
Batch size	64	Replay buffer capacity	5000
Initial/minimum ϵ	0.999/0.010	ϵ decay	0.9965
Gradient updates per episode	2	Target-network update interval	120
Gradient clipping	5.0	Weight decay	1×10^{-5}
Full-trajectory/TD-target weights	0.70/0.30	Reward scaling	50

5.3. Experimental Design and Model Evaluation Protocol

To prevent data leakage and unequal tuning opportunities, the 30 baseline simulation trajectories are divided into 18 training trajectories, six internal validation trajectories, and six fully held-out test trajectories, with non-overlapping identifiers [36,37,54]. The four primary models are each developed under eight candidate hyperparameter configurations and 10 independent tuning seeds, for a total of 320 training runs. The configuration and training duration are selected solely according to the interquartile mean (IQM) of validation-set reward; the test set is not involved in selection.

After the hyperparameters are frozen, each of the four primary models is independently trained using 30 new formal random seeds, yielding 120 formal training runs. Each seed is evaluated over six fully held-out trajectories, six policy effectiveness levels, and three budget anchors, for 108 test scenarios. Statistical inference treats the independently trained seed as the replication unit: the test scenarios are first aggregated within each seed, and comparisons are then made across the 30 seeds. The complete data split and training parameters are reported in Tables 13 and 14.

Table 14. Strict training and evaluation design.

Stage	Setting	Number of Independent Training Runs/Scenarios	Purpose
Baseline trajectory split	18 training/6 validation/6 test	30 trajectories	Environment scenarios isolated
Hyperparameter candidates	4 primary models $\times$ 8 configurations $\times$ 10 tuning seeds	320 training runs	Select configuration by validation-set IQM
Formal primary models	4 models $\times$ 30 formal seeds	120 training runs	Compare DQN architectures
Tests for each formal seed	6 test trajectories $\times$ 6 policy effectiveness levels $\times$ 3 budget anchors	108 scenarios	Fully held-out evaluation
Random-seed isolation	10 tuning seeds/30 formal seeds	Non-overlapping	Prevents tuning to formal seeds
Total independent training	320 development + 120 formal	440 runs	Primary-model experiments reported in the main text

Note: The formal seeds range from 20261001 to 20261030. The hyperparameters, training duration, and aggregation rules for the formal results are determined entirely outside the test set; the test set is not used for any model selection.

6. OSR-DDQN Governance Performance and Strategy Comparison

This section evaluates OSR-DDQN along three dimensions: model performance, resource use, and dynamic strategy, using a unified comparison with DQN, Double DQN, and Dueling DQN. All performance results are obtained under the same frozen natural evolution mechanisms and prespecified governance-effect functions. They therefore represent the relative performance of different decision algorithms within the simulated environment rather than causal effects of real-world policy interventions.

The evaluation first compares the four reinforcement learning models in terms of cumulative reward, final collaboration resilience, resilience improvement, and budget utilization, and then examines the dynamic differences using resilience trajectories and budget-consumption processes. The budget response patterns are subsequently assessed, together with comparisons against no intervention, one-step cost-aware greedy governance, and a mechanism-guided fixed scheme.

6.1. Evaluation Metrics and Statistical Inference Protocol

The evaluation metrics jointly cover decision returns, resilience improvement, and resource consumption. The cumulative reward reflects the overall quality of an action sequence under benefits, costs, and structural constraints; the final collaboration resilience measures the terminal state; the absolute gain and relative improvement characterize the improvement over natural evolution; and the gap reduction ratio further accounts for the remaining distance between the baseline state and the theoretical upper bound of the scale. This multidimensional evaluation avoids treating the same resilience outcome achieved at a substantially higher cost as equally effective.

Let the final resilience of the same-source no-intervention trajectory be R_T^B and that of the governance trajectory be R_T^I . Then,

$$\Delta R^{\text{abs}}(T) = R^I(T) - R^B(T) \quad (26)$$

$$U_T^{\text{raw}} = \frac{R_T^I - R_T^B}{R_T^B} \quad (27)$$

$$G_T^{\text{close}} = \frac{R_T^I - R_T^B}{1 - R_T^B} \quad (28)$$

The absolute gain captures the direct resilience difference between the governance trajectory and natural evolution. The raw relative improvement uses the baseline resilience as the denominator and therefore measures the relative magnitude of policy-induced gain. The gap reduction ratio uses the remaining distance to the upper bound of the index scale, one, as the denominator and facilitates comparison of improvement across different baseline levels. The formal statistics use the three budget anchors $\rho = 0.30, 0.50$, and 0.70 . Within each formal seed, the 36 test scenarios under the same budget are first averaged; the means, IQM, bootstrap intervals, and paired tests are then calculated across the 30 seeds.

6.2. Outcome and Process Comparisons with Standard Deep Q-Networks

Table 15 summarizes the fully held-out performance of the four primary models at the three formal budget anchors $\rho = 0.30, 0.50$, and 0.70 . It jointly reports the reward distributions, final collaboration resilience, relative improvement, and budget utilization to compare outcome quality and resource cost across models.

Table 15. Fully held-out performance of the primary models at three formal budget anchors.

Model	Mean Reward	Median	IQM	95% Interval for the Mean	Final Resilience	Relative Improvement	Gap Reduction	Budget Utilization
Low budget ($\rho = 0.30$)								
DQN	0.428	0.440	0.436	[0.417, 0.438]	0.764	18.48%	33.49%	93.02%
Double DQN	0.434	0.443	0.439	[0.424, 0.445]	0.765	18.72%	33.93%	92.22%
Dueling DQN	0.466	0.481	0.474	[0.454, 0.477]	0.773	19.87%	36.01%	88.39%
OSR-DDQN	0.520	0.521	0.520	[0.517, 0.523]	0.787	22.04%	39.95%	81.93%
Medium budget ($\rho = 0.50$)								
DQN	0.468	0.467	0.466	[0.464, 0.472]	0.767	18.98%	34.40%	61.85%
Double DQN	0.471	0.472	0.470	[0.466, 0.476]	0.767	19.06%	34.56%	59.87%
Dueling DQN	0.490	0.497	0.493	[0.480, 0.501]	0.772	19.78%	35.86%	56.75%
OSR-DDQN	0.548	0.546	0.547	[0.544, 0.551]	0.787	22.10%	40.07%	50.63%
High budget ($\rho = 0.70$)								
DQN	0.479	0.477	0.479	[0.475, 0.483]	0.766	18.89%	34.24%	44.80%
Double DQN	0.482	0.483	0.482	[0.478, 0.486]	0.767	18.98%	34.41%	43.21%
Dueling DQN	0.495	0.491	0.496	[0.484, 0.506]	0.770	19.48%	35.31%	41.77%
OSR-DDQN	0.559	0.556	0.558	[0.555, 0.563]	0.787	22.08%	40.04%	36.69%

Table 15 shows that OSR-DDQN achieves mean rewards of 0.520, 0.548, and 0.559 under low, medium, and high budgets, respectively, while the final collaboration resilience remains 0.787 across all three budgets. Its budget utilization is 5.08–6.46 percentage points lower than that of Dueling DQN. As the budget increases, the mean reward continues to rise while the terminal resilience stabilizes. Section SI further shows that 0.787 is close to the searched reachable frontier of 0.791740 in the original environment, but is not a fixed upper bound; a higher terminal resilience remains reachable when the governance capacity is expanded.

Figures 10 and 11 show that resilience gains of the standard DQN-family models are concentrated mainly in the early stages and then flatten, whereas OSR-DDQN continues to improve in the middle and later stages and surpasses the other models around step 6. This dynamic pattern is consistent with the state-transition structure identified in Section 4 and with the design motivation of OSR-DDQN to use intertemporal value information to accommodate delayed benefits.

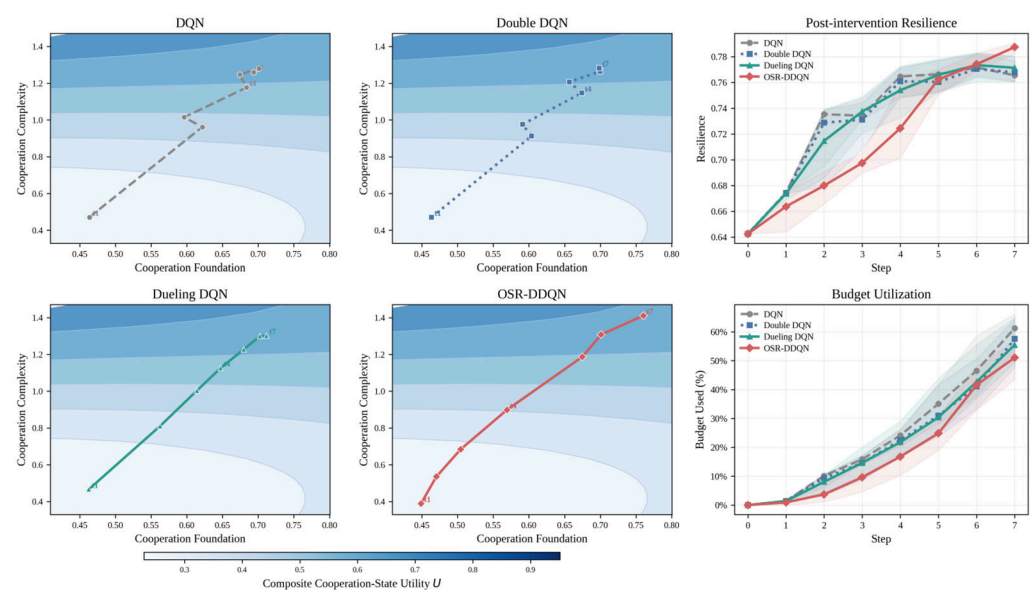


Figure 10. Collaboration-state landscape paths and dynamic performance comparison of the primary models.

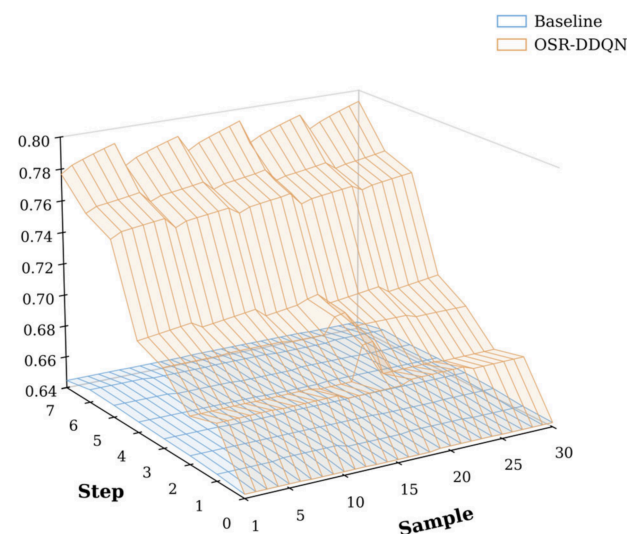


Figure 11. Dynamic trajectories of collaboration resilience in representative test scenarios.

The seed-level paired comparisons are consistent with the overall ranking. Relative to Dueling DQN, the best-performing standard model at each budget, OSR-DDQN achieves mean reward advantages of 0.054, 0.057, and 0.064 under low, medium, and high budgets, corresponding to relative mean reward gains of 11.56%, 11.66%, and 12.83%. The ablation results in Section SE further show that removing the full-trajectory return target produces the largest reductions in reward and terminal resilience, supporting its interpretation as the principal source of OSR-DDQN's incremental performance. Section SF uses local parameter perturbations and complete retraining across environments to examine environmental robustness; those results provide supplementary evidence and are not substitutes for the formal comparison in Table 15.

To further assess whether the ranking depends on the exact reward–weight combination, Table 16 summarizes the budget-averaged results for OSR-DDQN and Dueling DQN under the six local perturbations in Section SG. Because the perturbations alter the reward scale, the absolute rewards are compared only within the same setting.

Table 16. Main results for OSR-DDQN relative to Dueling DQN under local reward–weight Perturbations.

Local Perturbation	OSR Reward	Reward Difference and 95% Interval	Holm p	OSR Final Resilience	Resilience Difference and 95% Interval	Holm p
Relative importance of resilience −20%	0.514	+0.047 [0.032, 0.064]	0.0117	0.786	+0.012 [0.009, 0.016]	0.0117
Relative importance of resilience +20%	0.562	+0.067 [0.054, 0.079]	0.0117	0.787	+0.017 [0.014, 0.020]	0.0117
Relative importance of cost −20%	0.555	+0.063 [0.045, 0.080]	0.0117	0.787	+0.016 [0.012, 0.021]	0.0117
Relative importance of cost +20%	0.532	+0.062 [0.046, 0.076]	0.0117	0.788	+0.016 [0.012, 0.020]	0.0117
Relative importance of structural constraints −20%	0.552	+0.066 [0.053, 0.079]	0.0117	0.787	+0.017 [0.013, 0.020]	0.0117
Relative importance of structural constraints +20%	0.533	+0.062 [0.047, 0.077]	0.0117	0.786	+0.016 [0.012, 0.020]	0.0117

Note: Each setting uses 10 new paired seeds. Results are first equally averaged within each seed across three budgets of 0.30, 0.50, and 0.70. Differences are calculated as OSR-DDQN minus Dueling DQN. Brackets contain seed-level paired-bootstrap 95% intervals, and p -values are Holm-adjusted results from the two-sided exact sign-flip tests across the six settings.

Table 16 shows that, across the six local perturbations, OSR-DDQN's budget-averaged reward advantage ranges from 0.047 to 0.067 and its final collaboration resilience advantage ranges from 0.012 to 0.017. The six Holm-adjusted p -values are 0.0117 for both outcomes. When the results are further disaggregated by the three budget levels, all 18 comparisons for the mean reward, final collaboration resilience, and resilience efficiency per unit of budget remain positive and survive the BH-FDR correction; the complete results are reported in Section SG. The main-text model ranking is therefore stable within the examined $\pm 20\%$ local weight range, but this does not imply that the same advantage would hold under arbitrary reward functions or extreme objective specifications.

Figure 10 displays the state and performance trajectories but does not report the frequency with which individual governance actions are selected. Accordingly, no particular policy instrument is interpreted here as a high-frequency model choice. Combining the figure with the transition results supported by the global BH-FDR correction in Section 4 allows the four states to be linked to specific governance tasks. When the share of low-

cooperation dormancy is high, priority may be given to helping actors with innovation foundations and relational capital re-establish collaboration, with university/research institution status serving as important information for identifying activation targets. For actors in expansion and reconfiguration, technological matching and relationship consolidation may receive priority, with the governance objective of improving conditions associated with stable deepening. When the risk of relationship instability rises, attention may focus on existing collaborations that retain relational capital, together with governance instruments such as project coordination, joint validation, and knowledge-interface management.

6.3. Budget Sensitivity and Comparison with Conventional Strategies

Table 15 evaluates the method rankings at the three budget anchors $\rho = 0.30, 0.50$, and 0.70 . Figure 12 further extends the budget ratio from 0 to 1 to display the response shape; the formal statistical inference remains based on the three budget anchors, and curves outside those anchors are not interpreted as independent experimental points. $\rho = 0$ denotes the same-source no-intervention baseline.

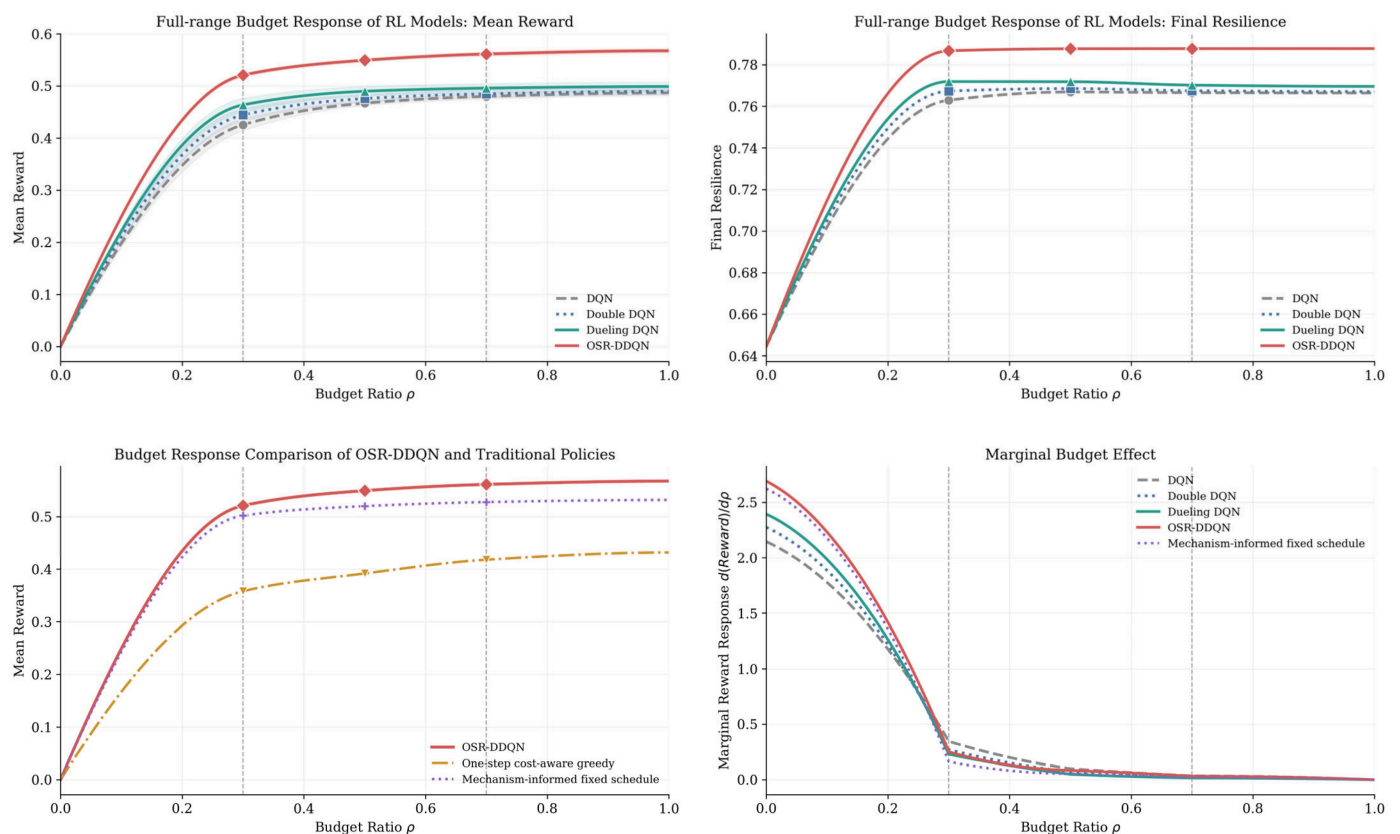


Figure 12. Sensitivity across the full budget range and comparison with conventional strategies.

Three conventional strategy baselines are compared. Under the no-intervention strategy, the system evolves naturally according to the frozen mechanisms, and its counterfactual reward is defined as 0. The one-step cost-aware greedy strategy selects, in each period, the feasible action with the highest current net reward; it accounts for the immediate cost but does not explicitly evaluate the effects on subsequent years. The mechanism-guided fixed scheme prespecifies staged actions based on the four-state structure and transition associations supported by the global BH-FDR correction, but does not adjust them in response to the current state. Comparing these three baselines with OSR-DDQN distinguishes the effects of introducing governance, short-term optimization, fixed stage-based rules, and adaptive intertemporal allocation.

Figure 12 shows that a larger budget does not increase the terminal resilience proportionally. When the budget ratio increases from 0.30 to 0.70, OSR-DDQN's mean reward rises from 0.520 to 0.559, whereas the final collaboration resilience remains approximately 0.787. This pattern suggests that action feasibility is the primary constraint at low budgets. Once basic governance actions become executable, an additional budget primarily affects action matching, implementation cost, and the remaining room for improvement, while its marginal effect on the terminal resilience is limited.

OSR-DDQN's mean reward increases by 0.028 when the budget rises from 0.30 to 0.50, but by only 0.011 when it rises from 0.50 to 0.70, indicating diminishing marginal gains. Together with the decline in budget utilization as the budget increases, this pattern suggests that, within the frozen governance environment used here, additional resources do not necessarily translate into higher terminal resilience. Curves outside the formal budget anchors are likewise shown only to illustrate trends and are not treated as independent statistical findings.

At budget ratios of 30%, 50%, and 70%, the mechanism-guided fixed scheme achieves mean rewards of approximately 0.502, 0.520, and 0.528, respectively, all higher than those of the one-step cost-aware greedy strategy. OSR-DDQN reaches 0.520, 0.548, and 0.559, and its reward advantage over the fixed scheme expands from 0.018 to 0.031. This result indicates that mechanism-guided staging improves the baseline performance of a fixed scheme, but as the budget increases and the set of feasible actions expands, a strategy capable of reallocating actions according to the current state and remaining budget yields larger relative gains.

As a real-world policy design concept requiring further validation, governance agencies may consider a rolling budget allocation rather than requiring a one-time allocation to be fully spent. At each year or project stage, the target actors could first be reclassified according to their relationship instability, low-cooperation dormancy, expansion and reconfiguration, and stable deepening, after which action decisions could be based on the recovery gap, relational process reserve, and remaining budget. For persistently dormant actors with innovation foundations or relational capital, collaboration matching and relationship activation support could receive priority. For actors already in expansion and reconfiguration, governance could focus on improving technological matching and consolidating new relationships. For actors who have regained stability and face low marginal returns to additional input, temporary non-intervention and budget retention should remain feasible options.

7. Discussion and Implications

7.1. Main Findings

After the global BH-FDR correction, the robust state-transition associations are concentrated in six results. The innovation capability, relational capital, and university/research institution status are associated with a higher relative propensity for actors in low-cooperation dormancy to enter expansion and reconfiguration; technological complementarity is associated with a higher relative propensity for actors in expansion and reconfiguration to enter stable deepening; and relational capital is also associated with a lower relative propensity for transition from expansion and reconfiguration to relationship instability and from stable deepening to expansion and reconfiguration. These findings indicate pronounced path specificity between collaboration conditions and collaboration activation, collaboration deepening, and state adjustment. Relational capital, in particular, exhibits directionally stable conditional associations across multiple transition paths.

The budget experiments further show that resource input and terminal resilience are not linearly related. As the budget rises from 0.30 to 0.70, OSR-DDQN's final collaboration

resilience remains 0.787 while the mean reward continues to increase, and its budget utilization is 5.08–6.46 percentage points lower than that of Dueling DQN. Low budgets primarily constrain the feasibility of necessary actions; after the budget reaches a minimally sufficient level, performance differences arise more from the choice of action targets, implementation timing, and decisions about additional input. The widening reward advantage of OSR-DDQN over the mechanism-guided fixed scheme as budgets increase is also consistent with the interpretation that dynamic state feedback becomes more valuable as the feasible choice set expands.

The ablation results further show that the incremental contributions of OSR-DDQN's components are uneven. Removing the full-trajectory return target produces the largest reductions in the mean reward and terminal resilience, whereas removing the priority intervention group branch or the Dueling component produces smaller changes. OSR-DDQN's incremental performance advantage is therefore driven primarily by the full-trajectory return mechanism, with the other components providing supplementary gains.

7.2. Theoretical Contributions

First, this study advances the interpretation of collaboration resilience from a static network outcome to a joint account of state transitions and relationship processes. Existing studies of network resilience typically emphasize connectivity, structural loss, or aggregate recovery outcomes, whereas the four states identified here show that similar collaboration scales can imply fundamentally different relational prospects and that stable deepening can transition back into expansion and reconfiguration. Collaboration resilience therefore cannot be judged solely from current tie counts or structural stability; it also depends on whether actors can maintain effective relationships, repair interrupted collaboration, and enter more favorable subsequent states. This interpretation integrates the current state, state reachability, and relational process reserve within a common dynamic framework and gives collaboration resilience an explicit foundation in relationship mechanisms.

Second, this study reveals the path specificity of associations with collaboration conditions. After the global BH-FDR correction, different organizational conditions receive support only on particular transition paths, indicating that their empirical associations depend on the actor's origin and target states and cannot be summarized as uniformly positive or negative effects. This provides a basis for a state-dependent interpretation of collaboration resilience.

Third, this study distinguishes two types of governance constraints: resource insufficiency and allocation efficiency. At low budgets, the main constraint is the feasibility of necessary actions. Once the budget reaches a minimally sufficient level, the terminal resilience stabilizes and the role of additional resources shifts toward matching action targets, optimizing implementation timing, and improving resource-use efficiency. The reachable-frontier test further shows that this plateau reflects diminishing marginal returns under the specified governance conditions rather than a fixed upper bound on resilience.

7.3. Managerial and Policy Implications

The specific action proposals in this section are derived primarily from the conditional associations in the observational data and relative strategy performance in the calibrated simulations. They should therefore be understood as governance hypotheses and candidate designs requiring validation rather than as policy prescriptions already demonstrated to be effective in real-world interventions. Their practical effectiveness still requires further examination using observational policy interventions, experiments or quasi-experiments, expert evaluation, and pilot programs.

7.3.1. Stage-Specific Governance for Collaboration Activation and Stable Deepening

For actors in low-cooperation dormancy, governance should focus on collaboration activation. After the global BH-FDR correction, the innovation capability, relational capital, and university/research institution status are stably associated with a higher relative propensity for dormant actors to enter expansion and reconfiguration. Governance may therefore prioritize low-collaboration actors with innovation foundations and existing relational foundations and consider instruments, such as industry–university–research matching, public technology platforms, and coordination with universities and research institutions, with the objective of reducing the matching costs of re-establishing collaboration.

After actors enter expansion and reconfiguration, governance priorities shift toward collaboration deepening. Technological complementarity is stably associated with a higher relative propensity for expansion-state actors to enter stable deepening. It can therefore serve as important information for partner screening and target identification during expansion [56], and measures, such as complementary technology matching, joint R&D, interface coordination, and joint validation, may be considered, with relationship consolidation rather than simple growth in partner numbers as the main objective of this stage.

7.3.2. Using Relational Capital to Identify Collaboration Activation and Stabilization Needs

Relational capital is a condition with comparatively stable directional associations in the state-transition results and can therefore serve as important information for identifying governance targets. For low-collaboration actors, historical partners that still possess technological fit and a foundation of trust may be reactivated first, reducing renewed search and coordination costs. For actors in expansion and stable states, prior relational capital may inform whether further expansion or relationship repair should receive priority. Because prolonged interruption, personnel mobility, and changes in technological trajectories can weaken existing relational foundations [57], collaboration records should be used together with joint trials, personnel exchange, and periodic review rather than treating historical collaboration frequency as equivalent to currently usable relational capital.

7.3.3. Dynamic Budgeting and Stage-Specific Exit

The simulation results suggest dynamic budgeting as a governance design, with the emphasis shifting from simply increasing the input intensity to comparing different input sequences and stopping conditions. At low budgets, resource guarantee schemes for relationship repair, partner activation, and necessary relationship consolidation could be tested first and compared with equal resource allocation. At medium and high budgets, further tests could reduce additional input to actors who have entered stable deepening or for whom earlier actions remain effective, while retaining flexible reserves for newly emerging instability, dormancy, or unstable expansion.

Stage-specific non-intervention can also be incorporated into this governance design. Annual or project-milestone review points could be established to determine the next-period action based on whether relationships have recovered, whether the current state has improved, whether prior relational capital remains effective, and the marginal return to additional input. Actors who have regained stability could temporarily suspend subsidies and matching support, whereas actors who remain outside stable deepening after expansion could receive priority for relationship consolidation and technological matching. Larger budgets need not automatically expand intervention coverage; unused resources could be retained for governance needs that emerge later.

7.4. Limitations and Future Research

This study has several limitations. First, the state-transition analysis is based on annual patent and organizational observations. Although all covariates enter the model using prior-period information and actor-block bootstrap, multiple-comparison correction, and out-of-time validation to improve robustness, these procedures cannot substitute for causal identification. State-transition coefficients represent conditional associations after controlling for observed conditions rather than causal effects of organizational attributes or relational conditions on transition outcomes. Accordingly, interpretations of technological complementarity, relational capital, and organizational identity should be understood as empirical associations consistent with theoretical mechanisms; they do not establish that these factors cause state changes through specific coordination, knowledge integration, or relationship maintenance processes. Future research could combine policy shocks, quasi-natural experiments, process surveys, or governance pilots to identify specific transmission mechanisms and causal effects more directly.

Second, co-patents capture only collaborations that produce formal technological outputs and cannot adequately cover open-source collaboration, joint standard setting, consulting exchanges, or informal knowledge flows. The relationship evolution and collaboration resilience identified here therefore primarily characterize formal technological collaboration networks. Future research could integrate multiple data sources, such as coauthored publications, project contracts, standards participation, corporate investment, and open-source contributions, to compare the relationship evolution, state transitions, and resilience across different collaboration vehicles.

Third, the analysis focuses mainly on organization pairs and organization–year states and does not fully characterize multilevel interactions among organizations, individuals, technological modules, and specific collaborative projects. Personnel mobility, replacement of key technological modules, and project life cycles may all alter the processes of relationship maintenance, recovery, and reconfiguration. Future research could develop multilayer dynamic networks integrating organizations, individuals, and technologies to identify how interactions across levels affect collaboration resilience.

Finally, the governance experiments are conducted in an annual simulation environment calibrated from historical data. The dynamic costs represent the relative resource burdens of different governance actions and cannot be directly converted into actual fiscal expenditures. The model also does not explicitly represent intra-year administrative coordination, policy-response lags, or the implementation sequence among specific governance actors. The present analysis is therefore primarily intended to compare the relative performance of different governance mechanism combinations under common constraints. Practical application will require further calibration using real costs, implementation cycles, and governance capacity.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/systems14091136/s1>.

Author Contributions: Conceptualization, J.X. and H.Z.; methodology, J.X.; software, J.X.; validation, J.X., H.Z. and G.Z.; formal analysis, J.X.; investigation, J.X.; resources, J.X.; data curation, J.X.; writing—original draft preparation, J.X.; writing—review and editing, J.X., H.Z. and G.Z.; visualization, J.X.; supervision, H.Z.; project administration, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to thank all individuals who provided guidance, suggestions, and technical support during the preparation of this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Shi, Y.; Cui, T.; Kurnia, S. Value co-creation for digital innovation: An interorganizational boundary-spanning perspective. *Inf. Manag.* **2023**, *60*, 103817. [\[CrossRef\]](#)
- Bettiol, M.; Capestro, M.; di Maria, E.; Grandinetti, R. Leveraging on intra- and inter-organizational collaboration in Industry 4.0 adoption for knowledge creation and innovation. *Eur. J. Innov. Manag.* **2023**, *26*, 328–352. [\[CrossRef\]](#)
- Cohen, M.; Fernandes, G.; Godinho, P. Measuring the impacts of university-industry R&D collaborations: A systematic literature review. *J. Technol. Transf.* **2025**, *50*, 345–374. [\[CrossRef\]](#)
- Nalmpanti, A.D.; Wong, C.Y.; Oghazi, P. Collaborating for innovation: The inhibiting role of constraints. *J. Innov. Knowl.* **2024**, *9*, 100504. [\[CrossRef\]](#)
- Wang, Y.; Zhang, J.; Yan, Y.; Guan, J. The bidirectional causality of tie stability and innovation performance. *Res. Policy* **2024**, *53*, 105102. [\[CrossRef\]](#)
- Powell, W.W.; Koput, K.W.; Smith-doerr, L. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Adm. Sci. Q.* **1996**, *41*, 116–145. [\[CrossRef\]](#)
- Ahuja, G. Collaboration networks, structural holes, and innovation: A longitudinal study. *Adm. Sci. Q.* **2000**, *45*, 425–455. [\[CrossRef\]](#) [\[PubMed\]](#)
- Schilling, M.A.; Phelps, C.C. Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Manag. Sci.* **2007**, *53*, 1113–1126. [\[CrossRef\]](#)
- Boschma, R. Proximity and innovation: A critical assessment. *Reg. Stud.* **2005**, *39*, 61–74. [\[CrossRef\]](#)
- Phelps, C.; Heidl, R.; Wadhwa, A. Knowledge, networks, and knowledge networks: A review and research agenda. *J. Manag.* **2012**, *38*, 1115–1166. [\[CrossRef\]](#)
- Audretsch, D.B.; Belitski, M. Geography of knowledge collaboration and innovation in Schumpeterian firms. *Reg. Stud.* **2024**, *58*, 821–840. [\[CrossRef\]](#)
- Jin, Q.; Chen, H.; Wang, X.; Xiong, F. How do network embeddedness and knowledge stock influence collaboration dynamics? Evidence from patents. *J. Informetr.* **2024**, *18*, 101553. [\[CrossRef\]](#)
- Linnenluecke, M.K. Resilience in business and management research: A review of influential publications and a research agenda. *Int. J. Manag. Rev.* **2017**, *19*, 4–30. [\[CrossRef\]](#)
- Duchek, S. Organizational resilience: A capability-based conceptualization. *Bus. Res.* **2020**, *13*, 215–246. [\[CrossRef\]](#)
- Tóth, G.; Elekes, Z.; Whittle, A.; Lee, C.; Kogler, D.F. Technology network structure conditions the economic resilience of regions. *Econ. Geogr.* **2022**, *98*, 355–378. [\[CrossRef\]](#)
- Hu, X.; Ma, H. Understanding the resilience of intercity innovation networks. *Humanit. Soc. Sci. Commun.* **2025**, *12*, 1611. [\[CrossRef\]](#)
- Vinyals, M.; Sabbadin, R.; Couture, S.; Sadou, L.; Thomopoulos, R.; Chapuis, K.; Lesquoy, B.; Taillandier, P. Toward AI-designed innovation diffusion policies using agent-based simulations and reinforcement learning: The case of digital tool adoption in agriculture. *Front. Appl. Math. Stat.* **2023**, *9*, 1000785. [\[CrossRef\]](#)
- Wachi, A.; Shen, X.; Sui, Y. A survey of constraint formulations in safe reinforcement learning. In Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, Jeju, Republic of Korea, 3–9 August 2024; pp. 8262–8271. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhao, T.; Li, G.; Zhao, T.; Chen, Y.; Xie, N.; Niu, G.; Sugiyama, M. Learning explainable task-relevant state representation for model-free deep reinforcement learning. *Neural Netw.* **2024**, *180*, 106741. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yoo, Y.; Boland, R.J., Jr.; Lyytinen, K.; Majchrzak, A. Organizing for innovation in the digitized world. *Organ. Sci.* **2012**, *23*, 1398–1408. [\[CrossRef\]](#)
- Nambisan, S.; Lyytinen, K.; Majchrzak, A.; Song, M. Digital innovation management: Reinventing innovation management research in a digital world. *MIS Q.* **2017**, *41*, 223–238. [\[CrossRef\]](#)
- Belitski, M.; Delgado-márquez, B.L.; Pedauga, L.E. Your innovation or mine? The effects of partner diversity on product and process innovation. *J. Prod. Innov. Manag.* **2024**, *41*, 112–137. [\[CrossRef\]](#)
- Abbasiharofteh, M.; Broekel, T.; Mewes, L. The role of geographic distance and technological complexity in U.S. interreg. co-patenting over almost two centuries. *Environ. Plan. A Econ. Space* **2024**, *56*, 2003–2022. [\[CrossRef\]](#)

24. Barbosa, S.; Sáiz, P.; Zofío, J.L. The emergence and historical evolution of innovation networks: On the factors promoting and hampering patent collaboration in technological lagging economies. *Res. Policy* **2024**, *53*, 104990. [\[CrossRef\]](#)
25. Snijders, T.A.B.; van de Bunt, G.G.; Steglich, C.E.G. Introduction to stochastic actor-based models for network dynamics. *Soc. Netw.* **2010**, *32*, 44–60. [\[CrossRef\]](#)
26. Krivitsky, P.N.; Handcock, M.S. A separable model for dynamic networks. *J. R. Stat. Soc. Ser. B* **2014**, *76*, 29–46. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Jiang, W.; Hong, Y.; Cui, M.; Li, M. Collaborative innovation network embeddedness and a firm's technological impact: Does prior networking experience matter? *J. Technol. Transf.* **2025**, *50*, 2802–2836. [\[CrossRef\]](#)
28. Belik, I.; Bhattacharya, P.; Knudsen, E.S. A case for simulated data and simulation-based models in organizational network research. *Res. Policy* **2024**, *53*, 105058. [\[CrossRef\]](#)
29. Neuländtner, M.; Scherngell, T.; Liu, X.; Ji, X.; Ding, X.; Paier, M. Dynamics of regional knowledge creation and inter-city R&D collaboration in China: Evidence from an empirical agent-based simulation model. *Ann. Reg. Sci.* **2025**, *74*, 77. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Qi, X.; Mei, G. Network resilience: Definitions, approaches, and applications. *J. King Saud Univ.—Comput. Inf. Sci.* **2024**, *36*, 101882. [\[CrossRef\]](#)
31. Wang, L.; Yang, X.; Zhu, F.; Klakegg, O.J. How governance of interorganizational projects develops resilience: Mediating role of resource reconfiguration. *J. Manag. Eng.* **2023**, *39*, 4022076. [\[CrossRef\]](#)
32. Du, Y.; Wang, Q.; Song, Y.; Xin, Y. How cross-regional collaborative innovation networks affect regional economic resilience: Evidence from 283 cities in China. *Technol. Forecast. Soc. Change* **2025**, *215*, 124057. [\[CrossRef\]](#)
33. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533. [\[CrossRef\]](#) [\[PubMed\]](#)
34. van Hasselt, H.; Guez, A.; Silver, D. Deep reinforcement learning with double Q-learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*; AAAI Press: Palo Alto, CA, USA, 2016; Volume 30, pp. 2094–2100.
35. Wang, Z.; Schaul, T.; Hessel, M.; Hasselt, H.; Lanctot, M.; Freitas, N. Dueling network architectures for deep reinforcement learning. In *Proceedings of the 33rd International Conference on Machine Learning*, New York, NY, USA, 20–22 June 2016; pp. 1995–2003.
36. Agarwal, R.; Schwarzer, M.; Castro, P.S.; Courville, A.C.; Bellemare, M. Deep reinforcement learning at the edge of the statistical precipice. In *Proceedings of the Advances in Neural Information Processing Systems*, Virtual, 6–14 December 2021; Volume 34, pp. 29304–29320.
37. Patterson, A.; Neumann, S.; White, M.; White, A. Empirical design in reinforcement learning. *J. Mach. Learn. Res.* **2024**, *25*, 1–63.
38. Du, Y.; Hu, B.; Ma, X.; Long, L. User engagement and its resilience enhancement mechanisms in online communities: An integrated approach with deep reinforcement learning. *Inf. Process. Manag.* **2025**, *62*, 104265. [\[CrossRef\]](#)
39. Devarakonda, R.; Pavićević, S.; Reuer, J.J. Microfoundations of R&D alliance formation: The interplay of scientist mobility and the cooperative context of collaboration. *Strategy Sci.* **2022**, *7*, 56–70. [\[CrossRef\]](#)
40. Ji, X.; Peng, S.; Yang, S. Imbalanced binary classification under distribution uncertainty. *Inf. Sci.* **2023**, *621*, 156–171. [\[CrossRef\]](#)
41. Koyanagi, S.; Yokura, Y. Comparative study on the spatial characteristics of university-industry collaboration between Japan and the U.S.: Empirical analysis using patent data. *Evol. Institutional Econ. Rev.* **2025**, *22*, 655–683. [\[CrossRef\]](#)
42. Merlo, L.; Petrella, L.; Tzavidis, N. Quantile mixed hidden Markov models for multivariate longitudinal data: An application to children's Strengths and Difficulties Questionnaire scores. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2022**, *71*, 417–448. [\[CrossRef\]](#)
43. Mews, S.; Koslik, J.O.; Langrock, R. How to build your latent Markov model: The role of time and space. *Stat. Model.* **2025**, *25*, 481–507. [\[CrossRef\]](#)
44. Hewamalage, H.; Ackermann, K.; Bergmeir, C. Forecast evaluation for data scientists: Common pitfalls and best practices. *Data Min. Knowl. Discov.* **2023**, *37*, 788–832. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Amati, V.; Lomi, A.; Snijders, T.A.B. A goodness of fit framework for relational event models. *J. R. Stat. Soc. Ser. A Stat. Soc.* **2024**, *187*, 967–988. [\[CrossRef\]](#)
46. Buckby, J.; Wang, T.; Fletcher, D.; Zhuang, J.; Takeo, A.; Obara, K. Finding the number of latent states in hidden Markov models using information criteria. *Environ. Ecol. Stat.* **2023**, *30*, 797–825. [\[CrossRef\]](#)
47. Chen, Y.; Fuh, C.D.; Kao, C.L.M. Determine the number of states in hidden Markov models via marginal likelihood. *J. Mach. Learn. Res.* **2025**, *26*, 1–59.
48. Ricciotti, L.; Picone, M.; Pollice, A.; Maruotti, A. A zero-inflated hidden semi-Markov model with covariate-dependent sojourn parameters for analysing marine data in the Venice lagoon. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2025**, *74*, 506–529. [\[CrossRef\]](#)
49. Michelot, T. hmmTMB: Hidden Markov models with flexible covariate effects in R. *J. Stat. Softw.* **2025**, *114*, 1–45. [\[CrossRef\]](#)
50. Lagona, F.; Mingione, M. Nonhomogeneous hidden semi-Markov models for toroidal data. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2025**, *74*, 142–166. [\[CrossRef\]](#)
51. Lin, L.I.K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **1989**, *45*, 255–268. [\[CrossRef\]](#)

52. Wadoux, A.M.J.C.; Minasny, B. Some limitations of the concordance correlation coefficient to characterise model accuracy. *Ecol. Inform.* **2024**, *83*, 102820. [[CrossRef](#)]
53. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2018.
54. Eimer, T.; Lindauer, M.; Raileanu, R. Hyperparameters in reinforcement learning and how to tune them. In Proceedings of the 40th International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023; Volume 202, pp. 9104–9149.
55. Hayes, C.F.; Rădulescu, R.; Bargiacchi, E.; Källström, J.; Macfarlane, M.; Reymond, M.; Verstraeten, T.; Zintgraf, L.M.; Dazeley, R.; Heintz, F.; et al. A practical guide to multi-objective reinforcement learning and planning. *Auton. Agents Multi-Agent Syst.* **2022**, *36*, 26. [[CrossRef](#)]
56. Zhu, G.; Sai, F.; Qin, Z. How technological relatedness influences inter-firm collaborative performance: An empirical study on five high-tech fields of scientific and technological achievements. *Eur. J. Innov. Manag.* **2024**, *27*, 1496–1520. [[CrossRef](#)]
57. Keller, A.; Lumineau, F.; Mellewigt, T.; Ariño, A. Alliance governance mechanisms in the face of disruption. *Organ. Sci.* **2021**, *32*, 1542–1570. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.