An LLM-Driven Agent-Based Framework for Cross-Regional Construction Policy Evaluation in the Greater Bay Area

Wentao Zhu, Xiao Li, Juan Wang, and Mingtian Xiong
The University of Hong Kong, Hong Kong SAR, China

Abstract

Cross-jurisdiction construction policies in the Guangdong–Hong Kong–Macao Greater Bay Area (GBA) aim to reduce institutional barriers and promote collaboration, yet their impacts are difficult to evaluate ex ante. This paper proposes an LLM-driven agent-based simulation framework integrating a stage-gated project lifecycle, heterogeneous stakeholder agents, and policy-parameterized institutional rules. Policy provisions are encoded into eligibility, cost–risk, and transition mechanisms, enabling traceable policy-to-outcome analysis. Validation is conducted through stylized-facts consistency, reproducing concentration growth and distance-decay collaboration patterns. The framework provides a reproducible testbed for analyzing cross-regional construction policies and their effects on collaboration networks and industry dynamics.

Introduction

The Guangdong–Hong Kong–Macao Greater Bay Area (GBA) is a “9+2” city-cluster initiative intended to strengthen regional connectivity and coordination. In construction, collaboration is increasingly cross-regional but remains shaped by differentiated city functions and institutional arrangements (Hui et al., 2020). To reduce these frictions, Mainland GBA jurisdictions have introduced record-filing and mutual-recognition arrangements, including pilot measures in Guangdong, Qianhai, and Hengqin (Guangdong Provincial Department of Housing and Urban-Rural Development, 2020). A central policy question is whether such measures can materially expand cross-border collaboration and reshape industry outcomes.

Ex ante assessment remains difficult. Macro approaches such as CGE and system dynamics provide transparent aggregate policy mappings, but are less suited to tracing heterogeneous behavior, repeated interaction, and institutionally differentiated responses (Shi et al., 2019; Shin et al., 2014). Agent-based modeling (ABM) addresses this limitation by representing heterogeneous actors and emergent system outcomes (Bonabeau, 2002). In construction, ABM has been used to study supply chains, labor productivity, and safety interactions (Min and Björnsson, 2008; Watkins et al., 2009; Zhang et al., 2019). However, most construction ABM still relies on manually specified heuristics or thresholds (Khodabandelu and Park, 2021). This becomes restrictive when decisions depend jointly on project stage, workload, partner history, policy friction, and compliance requirements, because rule sets grow quickly and remain brittle under newly combined conditions (Li et al., 2024; Wijermans et al., 2023).

Large language models (LLMs) provide a practical way to enrich ABM with more flexible and context-sensitive decision-making. Prior work shows that LLM agents can operate within structured environments under external constraints (Li et al., 2024; Dang et al., 2025), while memory-based LLM agents can support history-dependent adaptation (Park et al., 2023). This is useful in cross-jurisdiction construction settings, where firms must respond simultaneously to institutional heterogeneity, prior interactions, and changing project conditions. Table 1 summarizes the contrast among alternative approaches.

Two practical gaps remain. First, construction projects are process- and constraint-intensive: actions must satisfy stage gates, role dependencies, and compliance chains, so purely dialogue-oriented agents may not yield executable trajectories. Second, cross-jurisdiction policies act through computable mechanisms—entry friction, compliance burden, and scope constraints—so without explicit parameterization, simulations cannot trace effects from policy provisions to behavior and then to network and performance outcomes.

To address these gaps, we propose an LLM-driven ABM framework for evaluating cross-regional collaboration policies in the GBA construction sector. The framework combines an executable four-stage project lifecycle (INIT → SCHEME → AWARD → DELIVER), seven stakeholder roles operating through stage-feasible action sets, and a policy parameter vector embedded into gates and waiting times, cost–risk functions, and event/transition rules.

For validity, we adopt a stylized-facts consistency strategy. Rather than matching exact values, we test whether the simulation reproduces two robust empirical regularities of construction markets: (i) concentration with head-firm dominance and (ii) distance decay with local bias in cross-city collaboration. We then use the platform for counterfactual analysis of network stratification, diffusion paths, and directional asymmetries under alternative policy scenarios.

Methodology

Framework overview

We develop a micro–macro coupled agent-based simulation framework to characterize cross-regional collabora-

Table 1: Comparison of alternative simulation approaches for policy modeling

Model	Micro-level process-aware behavior	Policy heterogeneity	History / memory adaptation	Flexible under new contexts	Low manual rule burden
CGE / System Dynamics	✗	△	✗	✗	✓
Rule-based Agents	✓	△	△	✗	✗
LLM-driven Agents	✓	✓	✓	✓	△

Note: ✓ indicates strong support for the corresponding capability; △ indicates partial support, usually requiring simplified representations or additional manual assumptions; ✗ indicates that the capability is weakly supported or not explicitly represented.

Table 2: Fixed four-stage project lifecycle

Stage (s)	Purpose	Actions	Hard gates $g_s(\mathrm{x})$	Main outputs O_s
INIT	Formalize constraints and form the minimal team	OWN sets budget_ceiling; and schedule_target/risk_preference; QS derives cost_ceiling and cost_structure; ARC produces functional_brief; ENG produces feasibility_precheck.	brief_ready ∧ cost_ceiling_set ∧ team_ready (ARC,QS,ENG)	team_roster; budget_ceiling; functional_brief; risk_precheck.
SCHEME	Freeze scheme and close the cost loop	ARC generates scheme_package; ENG provides system_strategy; SUP provides price_band; and lead_time/availability; QS produces scheme_estimate; and budget_gap; OWN executes freeze_or_iterate.	scheme_frozen ∧ estimate ≤ budget_ceiling	frozen_scheme; tender_boundary; scheme_estimate; material_risk_signal.
AWARD	Mobilization gates: contract, safety, procurement	QS prepares BQ and contract_terms; CS submits bid (incl. risk premium); OWN awards and signs; SAFE outputs safety_plan; SUP confirms critical procurement terms.	contract_signed ∧ safety_plan_ready ∧ critical_supply_secured	signed_contract; safety_plan; procurement_terms; pre_start_delay.
DELIVER	Execute and close loops; handover and final account	CS executes construction; SAFE runs inspect→correct→reinspect; SUP triggers supply events; ARC/ENG initiate changes, QS prices, OWN approves/rejects; then commissioning, final account, and handover.	handover_complete ∧ final_account_closed	Outcome distributions; change_orders; compliance_log; supply_incidents; handover_docs.

tion in the GBA construction sector and evaluate policy impacts (Fig. 1). The model uses an executable, stage-gated project lifecycle as the process backbone, within which heterogeneous stakeholder decisions and interactions are embedded under hard gate conditions, producing auditable project trajectories and a time-varying collaboration network. A lightweight macro environment captures demand, capacity utilization, price pressure, and regulatory intensity, thereby coupling multi-project concurrency with micro execution outcomes. Cross-jurisdiction access and practice rules are encoded as a policy vector $\boldsymbol{\theta}(p)$ and embedded into (i) eligibility and waiting times, (ii) cost–risk functions, and (iii) event/transition mechanisms, yielding a traceable chain from policy provisions to behavioral change and system outputs, including network evolution, concentration trends, and distance-decay signatures.

Project lifecycle backbone: four-stage state machine

Each project follows a fixed four-stage state machine (INIT → SCHEME → AWARD → DELIVER), with transitions governed by hard gates defined on project states (cost, schedule, compliance, and supply). Agents may execute only stage-feasible actions, ensuring executability and direct logging of responsibilities and collaboration ties.

INIT sets budget, schedule, and team formation conditions. **SCHEME** freezes the scheme and closes the cost loop. **AWARD** finalizes mobilization prerequisites, including contract, safety, and procurement. **DELIVER** simulates execution through supply, inspection–correction, and change-management loops, propagating impacts to time–cost outcomes and network evolution.

LLM-driven agents and action execution

The framework includes seven heterogeneous, LLM-driven agent types: owner (OWN), architect (ARC), engineer (ENG), quantity surveyor (QS), contractor (CS),

Table 3: *Policy provisions, quantified model parameters, and the initial baseline state*

Region / regime	Entry friction	Compliance burden	Scope constraints	τ_p	π_p	σ_p	c_p	s_p
Initial state (2015 pre-policy baseline)	Long and uncertain cross-border approval/recognition procedures.	Higher recurring administrative burden and lower renewal predictability.	Narrow effective cross-border practice scope.	120	0.40	0.45	0.55	0.25
Guangdong provincial pilot (9 Mainland cities)	Online filing; registration/publication within **10 working days**; other procedural stages not uniformly specified.	Validity aligned with **Hong Kong registration validity**; change filing within **30 days**; renewal/annual reporting not explicitly specified.	Broad service coverage within the 9-city area, subject to qualification/capability.	20	0.71	0.20	0.25	0.85
Qianhai (Shen Qianhai Gui [2023] No. 6)	Preliminary review within **5 working days**; public notice **≥5 working days**; objection handling **30 working days** (extendable).	Validity typically **5 years**; renewal within **90 days before** expiry; change filing within **90 days**.	Moderate scope constrained by practice-category filing and qualification mapping.	10	0.82	0.20	0.35	0.60
Hengqin (Bureau Doc. No. 6/2024)	Fully online; supplementation within **10 working days**; public notice **≥5 working days**; objection review exists.	Validity aligned with **Hong Kong/Macao registration validity**; renewal within **90 days before** expiry; change filing within **30 days**; annual report by **Dec. 31**.	Broad scope covering construction and multiple professional services, subject to filing scope.	10	0.80	0.25	0.45	0.80
Hong Kong (AP/RSE/RGE registration)	Committee referral/scheduling within **2 months**; outcome notification within **3 months** after review/interview.	Validity **5 years**; renewal window from **4 months** to **28 days** before expiry; restoration within **2 years**.	Statutory scope strictly bounded by registration class.	150	0.85	0.10	0.30	0.20
Macao (urban architecture / engineering / planning qualification)	Hard entry gate requiring **≥2 years** full-time or **≥5 years** part-time internship before registration.	Renewal follows local rules; validity reported as **until the following January**.	Statutory scope bounded by recognized qualification and registration domain.	730	0.35	0.10	0.40	0.15

Note: τ_p is entry waiting time (days) and is set equal to the stated procedural time in the policy text, summing sequential stages where applicable. π_p is coded from normalized processing time, number of procedural stages, and whether the procedure contains unspecified or open-ended review components. σ_p is coded from the existence of extensions, objections, or discretionary review. c_p is coded from renewal frequency, annual reporting, and change-filing requirements. s_p is coded as the normalized breadth of recognized service categories and responsibility interfaces. All coded indices are scaled to $[0, 1]$. The initial state represents the pre-policy cross-border practice environment at the start of the simulation in 2015.

Table 4: *2015 empirical anchors and simulation initialization values of macro variables*

	D_t	K_t	u_t	P_t	M_t	R_t
2015 real-data anchor	RMB 40.81 tn new-start investment	168,222 firms / 51.15 m workers	12.43 bn m^2 / 51.15 m workers	98.2	98.9	442 accidents
Simulation init-value	40.81	168,222; 51.15	243	98.2	98.9	1.00

Note: tn = trillion; bn = billion; m = million. R_t is a constructed regulatory-intensity index, with the 2015 level set to 1.00.

safety/compliance entity (SAFE), and supplier (SUP). Each agent follows a role-consistent decision loop but acts only through the stage-feasible action set induced by the project state machine. The framework is therefore hybrid rather than purely generative: the LLM serves as a bounded behavioral policy module for context-sensitive action selection, while feasibility checks, state transitions, accounting updates, and regulatory enforcement are handled by the rule-based simulator. Agent calibration is implemented at the simulator level rather than through fine-tuning the LLM itself. Role populations, capacities, behavioral priors, and memory-update settings are initialized from empirical distributions or bounded assumptions and adjusted so that the baseline simulation reproduces key stylized collaboration patterns.

Memory mechanism. Each agent maintains two complementary memory stores. *Episodic memory* records short-horizon observations and time-stamped experience traces, including stage changes, gate outcomes, workload pressure, market or policy updates, and realized interaction outcomes such as invitations, bids, approvals, delays, violations, change orders, and partner-specific collaboration experiences. *Semantic memory* stores slower-changing summaries distilled from repeated episodes, including partner reliability, expected approval difficulty, coordination friction, effective bidding conditions, and self-assessed capability across jurisdictions.

Retrieval. At each decision step, the agent retrieves a compact subset of memories using recency–relevance scoring, with optional importance weighting for salient events.

Consolidation. Memory updating combines event-triggered and periodic summarization. Event-triggered consolidation follows high-impact milestones such as bid success or failure, project success or failure, major inspection or penalty events, and major policy changes. Periodic consolidation compresses recent episodic traces into higher-level summaries, allowing agents to revise beliefs about partners, jurisdictions, risks, and opportunities over time. After execution, realized outcomes are first written to episodic memory and may then update semantic memory through these mechanisms.

Workload modes. Agents operate in two modes: (i) *in-*

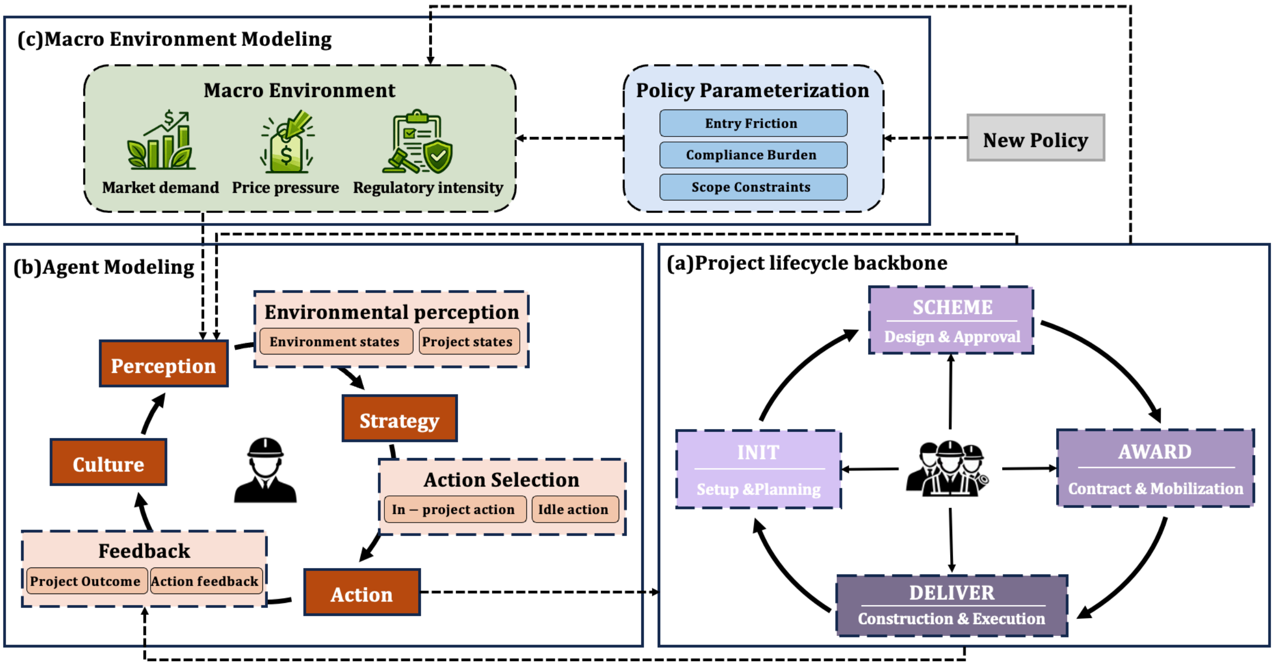

Figure 1: Overall micro–macro coupled agent-based simulation framework.

project, emphasizing resource lock-in, stage-dependent responsibilities, and reactive coordination; and (ii) *idle/out-of-project*, emphasizing opportunity search, teaming, eligibility maintenance, and bidding stance adjustment. When utilization exceeds capacity, agents enter a resource-constrained regime and apply interpretable load-shedding rules until capacity is released by project completion. This mechanism is important for reproducing concentration-upward dynamics (Fig. 3) through cumulative advantage under constrained capacity.

Decision pipeline. At each time step, the simulator builds the agent's decision context from three elements: **Perception**, **Culture**, and **Strategy**. **Perception** includes the observable state and retrieved memories, such as stage and gate status, workload and capacity, project states, macro signals, and partner reputation/availability. **Culture** captures stable value dispositions, such as baseline risk tolerance and cross-border inclination. **Strategy** captures short-horizon operating stances, such as expansion versus stability or profit-first versus scale-first. The simulator packages this context, together with role instruction and the admissible action set, into a structured prompt. The LLM selects one feasible action and returns it in a machine-parsable schema with a brief rationale. The simulator then executes the action and updates states, workloads, and memories through deterministic or stochastic transition rules.

Prompt structure and example. Each decision call is instantiated as a structured prompt with four components: (1) a *system layer* defining the agent's role, objectives, constraints, and behavioral style; (2) a *state layer* summarizing the current project stage, gate status, workload mode, key project indicators, and macro conditions; (3) a *memory layer* containing the retrieved episodic and semantic memories; and (4) an *action layer* listing the admissible actions and requiring one action with a brief rationale in a predefined output schema. Fig. 2 illustrates an example prompt template and decision interface for an LLM-driven agent.

```
[System Role]
You are a Contractor (CS) agent in a construction market simulation.
You must select exactly one feasible action.
Your action choice must comply with:
-  {Contractor Rules}, {Project State Machine}, {Jurisdiction Policy Constraints},
 {Workload Mode Constraints}

[Agent Profile]
Role: Contractor, Stable traits: {Risk Tolerance}, {Cross-border Inclination}
Current strategy stance: {Strategy Stance}, Home market: {Home Market}

[Current Decision Context]
Project stage: {Project Stage}, Gate status: {Gate Status}
Workload mode: {Workload Mode}, Utilization status: {Utilization Status}
Capacity condition: {Capacity Condition}, Observable state: {Current Project State}
Macro conditions: {Current Macro Conditions}

[Retrieved Memory]
Recent observations: {Recent World Observations}, Episodic memory: {Retrieved Episodic
Memory}
Semantic memory: {Retrieved Semantic Memory}

[Admissible Actions]
{Admissible Actions}

[Output Schema]
{ "selected_action": "...", "target": "...", "rationale": "...", "confidence": 0.00}

[Fallback Rule]
If the available information is incomplete or conflicting, choose the most conservative
feasible action consistent with {Contractor Rules} and current capacity constraints.
```

Figure 2: Example partially instantiated prompt template for a contractor agent.

Macro environment coupling

To couple micro execution with industry conditions, the macro state at time t is $z_t = \{D_t, K_t, u_t, P_t, M_t, R_t\}$, where D_t denotes market opportunity intensity, K_t role capacities, u_t utilization, P_t the general price level, M_t

the material price level, and R_t regulatory intensity. New projects are initiated endogenously as owners decide whether to launch projects under current macro conditions. Macro variables affect the simulation through three channels: *cost* (P_t, M_t), *delay/efficiency* (u_t), and *events* (M_t, u_t, R_t). Because the simulation starts in 2015, initial values are anchored to official data (Table 4) and then evolve through statistical inputs and endogenous feedback.

We use two parsimonious feedbacks:

$$P_{t+1} = P_t \left(1 + \kappa(\bar{u}_t - \bar{u}^\star)\right), \tag{1}$$

$$R_{t+1} = R_t + \eta \left(\mathrm{ViolationRate}_t - \rho^\star\right). \tag{2}$$

Policy parameterization and traceable embedding

For scenario evaluation, each policy regime p is represented by

$$\boldsymbol{\theta}(p) = \{\tau_p, \pi_p, \sigma_p, c_p, s_p\},$$

where τ_p denotes entry waiting time, π_p approval success probability, σ_p procedural uncertainty, c_p compliance burden, and s_p scope breadth. Table 3 reports both policy provisions and coded parameter values.

When a regional policy scenario is activated, the corresponding $\boldsymbol{\theta}(p)$ replaces the initial baseline state for that region. Specifically, τ_p is set from stated processing time; π_p is coded from normalized duration, procedural clarity, and review complexity:

$$\pi_p = \alpha_0 + \alpha_1 q_p - \alpha_2 \tilde{\tau}_p - \alpha_3 x_p,$$

procedural uncertainty is coded as

$$\sigma_p = \beta_1 u_p + \beta_2 x_p,$$

compliance burden is coded as

$$c_p = \gamma_1 \tilde{v}_p + \gamma_2 a_p + \gamma_3 f_p,$$

and scope breadth is coded as

$$s_p = \frac{m_p}{m_{\max}}.$$

These parameters enter the model through three channels: **(1) gates and waiting times,** where τ_p delays eligibility and cross-border readiness; **(2) cost and risk functions,** where π_p, σ_p, and c_p affect entry cost, compliance cost, and bid risk premia; and **(3) event and transition rules,** where c_p affects inspection, renewal, penalty, and disruption events, while s_p constrains the feasible cross-border action set.

This design yields a traceable mapping from *policy provisions → coded parameters → agent behavior → simulation outputs*.

Experiments and Analysis

The simulation is conducted in the GBA. The model includes seven heterogeneous agent types, with 100 agents per type (700 in total), allocated across cities according to observed firm distributions. Cross-jurisdiction policies are encoded as a parameter vector affecting entry friction, compliance burden, and scope constraints. The simulation spans 2015–2024 and generates multi-project concurrency and cross-city collaboration networks under a time-varying macro environment.

At the start of each run, agent populations, city allocation, macro initial values, and the pre-policy baseline are initialized using the calibration settings described earlier. A policy scenario is then activated by replacing the baseline policy state with the corresponding policy vector for the affected region. At each simulation step, owner agents decide whether to initiate new projects, while candidate partners make bounded LLM-driven decisions subject to workload, memory, policy constraints, and stage-feasible action sets. Feasible projects then advance through the lifecycle state machine, and repeated project-level interactions generate evolving collaboration networks and aggregate outputs for validation and policy analysis.

Simulation validity via stylized-facts consistency checks

We assess platform validity through *stylized-facts consistency*. Rather than targeting exact numerical matching, we test whether the model reproduces two robust empirical regularities of the construction industry: (i) market concentration with head-firm dominance and (ii) distance decay with local bias in cross-city collaboration.

Macro structure: concentration and head-firm dominance. Empirical evidence indicates rising concentration in the construction industry. For China, official statistics show that top-tier firms expanded their market share to 68.0% of output in 2021 (**?**), while the combined share of newly signed contracts held by the seven major central construction SOEs (CR7) reached 41.89% in 2023 (**?**). Consistent with these patterns, the simulation produces a clear concentration-upward trajectory (Fig. 3). We use the annual Top-10 share (CR10) as a regional proxy for GBA concentration and compare it with the national CR7 series. Both rise persistently, while the simulated CR10 exhibits a stable–accelerating–plateauing pattern driven by cumulative advantage under capacity, regulatory, and competitive constraints.

Spatial structure: distance decay and local bias. A second robust regularity is distance decay with strong local bias: cross-city collaboration becomes less frequent as distance increases, while same-city or nearby interactions remain dense. Similar patterns have been documented in large-scale procurement markets (Herz and Varela-Irimia, 2020). We conduct an analogous contract-level validation using simulated collaboration records. In Fig. 4, each point represents one collaboration contract, with inter-city distance on the x-axis and contract value on the y-axis. The scatter shows two stylized features: contracts become markedly sparser at longer distances, and contract values are heavy-tailed, with many small contracts and few large

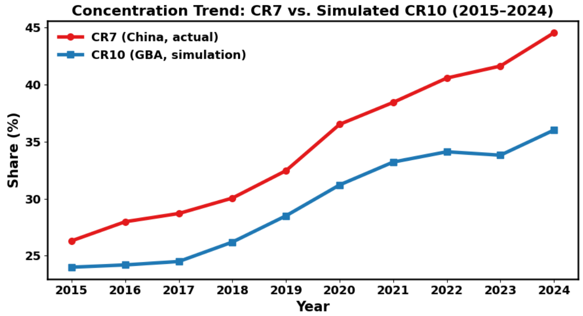

Figure 3: Concentration trend: national CR7 (China, actual) vs. simulated CR10 (GBA).

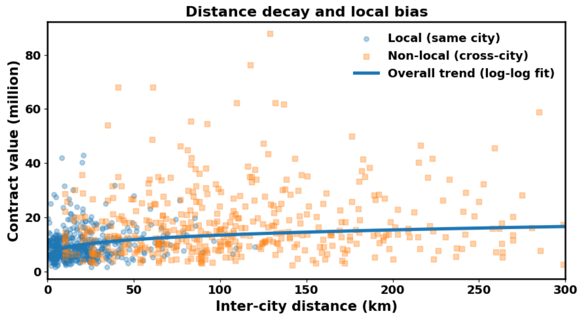

Figure 4: Distance decay and local bias in simulated collaboration contracts.

ones. Together, these results indicate that the platform reproduces representative industry regularities and therefore provides a credible baseline for counterfactual experiments.

Ablation study and stability. To assess the contribution of the proposed LLM-agent design, we compare three settings: (i) the full model with memory and intra-role heterogeneity, (ii) a variant without the memory module, and (iii) a variant without intra-role heterogeneity. Because the empirical benchmark is the national CR7 series whereas the simulation output is the GBA CR10 series, we compare normalized trajectories indexed to 2015 values. Similarity is evaluated using RMSE and trajectory correlation (ρ), and each setting is repeated three times. Table 5 reports mean values and relative variation across runs. The results show a stable ranking: the full model performs best, removing memory weakens performance, and removing intra-role heterogeneity weakens it further.

Cross-Regional Policy Analysis

After validating the platform against stylized facts, we use it to characterize the structure and city-level heterogeneity of cross-regional collaboration in the GBA, focusing on whether institutional differences yield interpretable spatial stratification and diffusion patterns.

Mainland-city collaboration network The simulated network among the nine Mainland GBA cities evolves into a clearer *core–semi-core–periphery* structure rather than expanding uniformly. Fig. 5 reports collaboration

Table 5: Ablation and stability results on concentration dynamics using normalized trajectories

Model setting	RMSE	Trajectory correlation (ρ)
Full model (memory + heterogeneity)	5.44 ± 28.47%	0.9913 ± 20.93%
Without memory module	9.62 ± 23.82%	0.9734 ± 16.39%
Without intra-role heterogeneity	11.44 ± 29.15%	0.9649 ± 23.68%

Note: The empirical benchmark is the national CR7 series, while the simulation output is the GBA CR10 series. Both series are indexed to their 2015 values before computing RMSE and trajectory correlation. Reported values are means over three repeated runs; percentages indicate relative variation across runs.

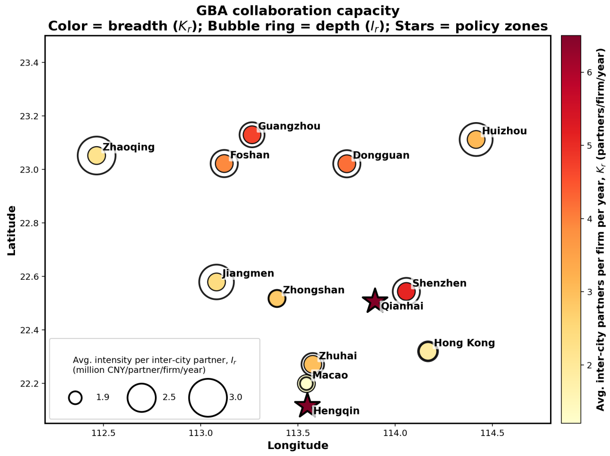

Figure 5: GBA collaboration capacity at the city level.

breadth K_r and collaboration depth I_r. Three patterns emerge: core hubs such as Guangzhou and Shenzhen exhibit high breadth; manufacturing and supply-chain cities such as Foshan and Dongguan form clustered collaboration anchored by stable division-of-labor ties; and peripheral cities such as Jiangmen, Zhaoqing, and Huizhou more often show lower breadth but higher depth, indicating fewer yet deeper channels.

Qianhai and Hengqin Qianhai and Hengqin function as *institutional gateways*. Their main role is to increase the arrival probability and diffusion speed of cross-border collaboration into the Mainland network by lowering institutional friction and improving procedural predictability. Two time-series signatures support this interpretation: (i) cross-border share rises earlier within the policy zones than in other cities (Fig. 6); and (ii) cross-border projects in other cities initially show a higher share linked to policy zones, which later declines as experience diffuses and projects shift from gateway-driven entry to broader localization (Fig. 7).

Hong Kong/Macao collaboration characteristics When institutional differences are explicitly modeled, Hong Kong/Macao shape the GBA collaboration landscape asymmetrically in both direction and intensity. Cross-border mobility becomes directional, with net inflow from Hong Kong/Macao to the Mainland—especially

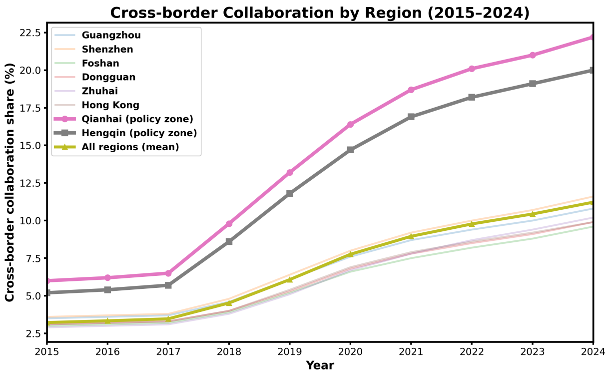

Figure 6: Cross-border collaboration by region (2015–2024).

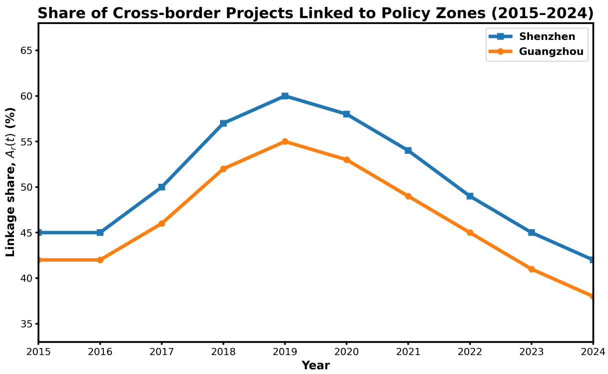

Figure 7: Share of cross-border projects linked to policy zones (2015–2024).

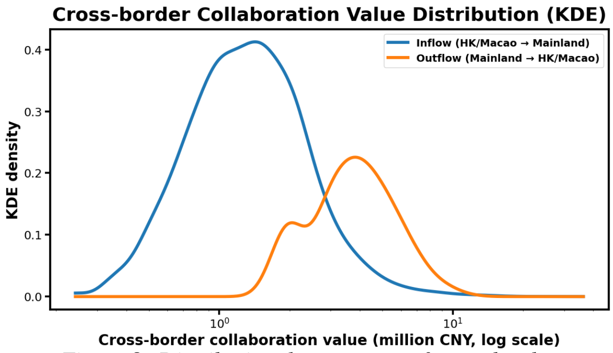

Figure 8: Distributional asymmetry of cross-border collaboration intensity/amount.

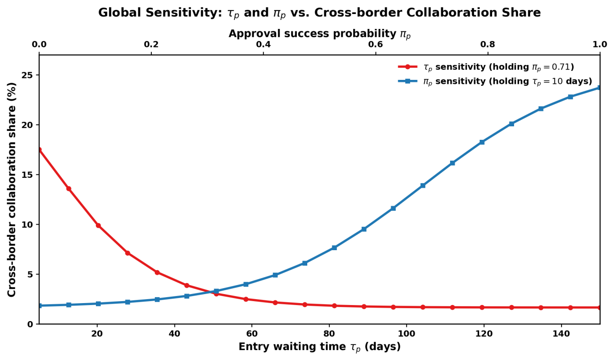

Figure 9: Global sensitivity of cross-border collaboration share.

via Qianhai and Hengqin—because Mainland-to-Hong Kong entry faces higher institutional costs and tighter constraints. Collaboration intensity also differs by direction: inflow is more concentrated in high-end professional and information/consulting services, whereas outflow is more constrained and tends to concentrate in price/efficiency-advantaged services and supply-chain cooperation.

Sensitivity analysis of key policy parameters

Global sensitivity of entry waiting time and approval success We first isolate entry waiting time τ_p and approval success probability π_p, holding other policy dimensions fixed, and record the resulting cross-border collaboration share (Fig. 9). The responses are nonlinear but smooth. For π_p, the curve flattens at both ends, indicating diminishing marginal effects. For τ_p, collaboration declines rapidly beyond the low-delay range and then approaches a floor. The model is therefore directionally sensitive but not dependent on knife-edge calibration.

Robustness of the Qianhai gateway effect under alternative friction gaps We further test whether Qianhai's gateway role remains robust when its friction advantage over the Guangdong baseline varies from 1% to 30%. For each setting, we compute the share of cross-border projects linked to Qianhai. As shown in Fig. 10, even a modest advantage increases Qianhai-linked projects, while marginal gains diminish as the advantage widens. The conclusion remains stable across the tested range: lower institutional friction systematically strengthens Qianhai's gateway role.

Conclusions

This paper proposes an LLM-driven agent-based simulation framework for evaluating cross-regional construction policies in the GBA. The framework combines an executable stage-gated project lifecycle, role-consistent agent action sets, and a lightweight macro environment, while encoding cross-jurisdiction practice rules as a policy vector embedded into gates, cost–risk functions, and event/transition rules. Stylized-facts consistency checks show that the platform reproduces two robust market regularities—increasing concentration with head-firm dominance and distance decay with strong local bias—thereby providing a credible baseline for counterfactual policy experiments. We further show how institutional differences and policy-zone mechanisms generate interpretable signatures in network stratification, diffusion paths, and directional asymmetries of cross-border collaboration intensity.

Beyond the GBA case, the framework is extensible and conditionally adaptable. While the core process backbone can be retained, applying it to other contexts would require context-specific adjustment of agent configuration parameters, macro-level parameters, and policy-vector settings to reflect local institutional arrangements, market conditions, and project structures. The framework should there-

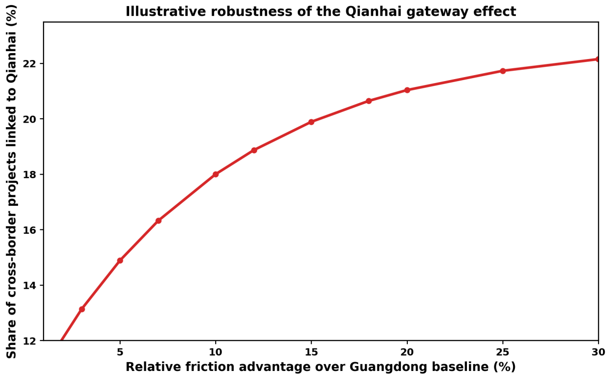

Figure 10: Qianhai gateway effect under alternative friction advantages.

fore be understood as a reproducible methodological template for comparative policy analysis, rather than a directly transferable model.

Acknowledgments

The work described in this paper was substantially supported by a grant from the Natural Science Foundation of China / Research Grants Council of the Hong Kong SAR of China (RGC Project No. CRS_PolyU505/23), and a grant from the Strategic Public Policy Research (SPPR) Funding Scheme (Project No. S2023.A6.023.23S).

References

Bonabeau, E. (2002). Agent-based modeling: methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences of the United States of America, 99(suppl. 3):7280–7287.

Dang, P., Zhu, J., Li, W., Xie, Y., and Zhang, H. (2025). Large-language-model-driven agents for fire evacuation simulation in a cellular automata environment. Safety Science, 191:106935.

Guangdong Provincial Department of Housing and Urban-Rural Development (2020). Interim administrative measures for the pilot program for hong kong engineering construction consultancy enterprises and professionals to establish practice and provide services in the mainland cities of the guangdong–hong kong–macao greater bay area. Guangdong Provincial Department of Housing and Urban-Rural Development.

Herz, B. and Varela-Irimia, X.-L. (2020). Border effects in european public procurement. Journal of Economic Geography, 20(6):1359–1405.

Hui, E. C. M., Li, X., Chen, T., and Lang, W. (2020). Deciphering the spatial structure of china's megacity region: a new bay area—the guangdong–hong kong–macao greater bay area in the making. Cities, 105:102168.

Khodabandelu, A. and Park, J. (2021). Agent-based modeling and simulation in construction. Automation in Construction, 131:103882.

Li, N., Gao, C., Li, M., Li, Y., and Liao, Q. (2024). Econagent: large language model-empowered agents for simulating macroeconomic activities. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, pages 15523–15536. Association for Computational Linguistics.

Min, J. U. and Björnsson, H. C. (2008). Agent-based construction supply chain simulator (cs2) for measuring the value of real-time information sharing in construction. Journal of Management in Engineering, 24(4):245–254.

Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., and Bernstein, M. S. (2023). Generative agents: interactive simulacra of human behavior. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, pages 1–22. Association for Computing Machinery.

Shi, Q., Ren, H., Cai, W., and Gao, J. (2019). How to set the proper level of carbon tax in the context of chinese construction sector? a cge analysis. Journal of Cleaner Production, 240:117955.

Shin, M., Lee, H.-S., Park, M., Moon, M., and Han, S. (2014). A system dynamics approach for modeling construction workers' safety attitudes and behaviors. Accident Analysis & Prevention, 68:95–105.

Watkins, M., Mukherjee, A., Onder, N., and Mattila, K. G. (2009). Using agent-based modeling to study construction labor productivity as an emergent property of individual and crew interactions. Journal of Construction Engineering and Management, 135(7):657–667.

Wijermans, N., Scholz, G., Chappin, É., Heppenstall, A., Filatova, T., Polhill, J. G., Semeniuk, C., and Stöppler, F. (2023). Agent decision-making: the elephant in the room—enabling the justification of decision model fit in social-ecological models. Environmental Modelling & Software, 170:105850.

Zhang, P., Li, N., Jiang, Z., Fang, D., and Anumba, C. J. (2019). An agent-based modeling approach for understanding the effect of worker-management interactions on construction workers' safety-related behaviors. Automation in Construction, 97:29–43.