Paradigm Academic Press
Innovation in Science and Technology
ISSN 2788-7030
OCT. 2025 VOL.4, NO.9

Construction of R&D Collaboration Mechanism for Small and Medium Cross-Border Technology Firms: Practices of Knowledge Sharing and Technological Breakthroughs in Transregional Teams

Dujin Xu[1]

[1] ETERNITY SUNSHINE CONSULTING PTE. LTD, Shanghai 200021, China

Correspondence: Dujin Xu, ETERNITY SUNSHINE CONSULTING PTE. LTD, Shanghai 200021, China.

doi:10.63593/IST.2788-7030.2025.10.004

Abstract

Amidst the fragmentation of globalization and the decoupling of Sino-US technologies, small and medium-sized technology firms are compelled to engage in cross-border R&D through an iterative and rapid approach. However, they face the dual challenges of diminishing knowledge spillovers and amplified cultural distance. This study integrates the Knowledge-Based View (KBV) and cultural adaptation theory to propose the "Dual-Curve Collaboration Model" (DCSM). It quantifies knowledge spillovers as KSE = Shared Coverage × Encoding Degree × Absorption Capacity and transforms cultural distance into a dynamic damping coefficient CDAC. Utilizing a mixed-methods approach with longitudinal panel data from 56 firms (2019-2023), 312 team surveys, and five extreme cases, we find that a 10% increase in KSE leads to a 3.5% improvement in R&D efficiency the following year. However, a 0.1 increase in CDAC erodes one-third of this gain. Virtual rituals and bilingual technical writing can shift the U-shaped cultural adaptation trough forward by 2.7 months and reduce the damping effect by 45%. The high-collaboration configuration is characterized by "High KSE and Low CDAC and Intensive Virtual Rituals." Failure cases, lacking a "local knowledge specialist + minority language documentation," exhibit efficiency 62% below the average (Tang Chenghui, Qiu Peng & Dou Jianmin, 2022). Based on these findings, we develop the CRD-Mat maturity scale, a 3-6-12 intervention roadmap, and a closed-loop retesting mechanism. This provides a "green zone" anchor for government subsidies and transforms cultural adaptation from a soft requirement into a hard metric that can be procured and insured.

Keywords: cross-border R&D collaboration, knowledge-based view, cultural distance, KSE index, dual-curve collaboration model, mixed methods, quantification of knowledge spillovers, virtual ritual intervention, 3-6-12 milestones, closed-loop retesting mechanism

1. Theoretical Foundations and Model Construction

1.1 Revisiting the Knowledge-Based View (KBV)

Traditional KBV conceptualizes firms as "knowledge repositories" but overlooks context in cross-border scenarios. We decompose knowledge into two poles: "codifiable" and "tacit." The former is solidified through bilingual documentation, API annotations, and multilingual subtitles, while the latter resides in the laughter of virtual stand-up meetings and emojis. Drawing on field notes from a China-Singapore chip team, we simplify "knowledge encoding" into three computable indicators: ISO compliance, bilingual document consistency, and Git commit language coverage. For the first time, we incorporate "cultural distance" directly into the knowledge attenuation function: high-context environments cause exponential declines in knowledge reach, with the rate determined jointly by cultural distance and time difference in hours. Thus, the knowledge advantage of cross-border SMEs no longer asks "how much" but "how much knowledge can still be verified, invoked, and iterated in a foreign land." This lays the micro-foundation for the subsequent trivariate model: "Spillover Volume

= Coverage × Encoding Degree × Absorption Capacity."

1.2 Extension of Cultural Adaptation Theory

The classic U-shaped curve fails to explain why, with the same 8-hour time difference, China-Germany teams complete chip development in 38 days, while China-Brazil AI teams experience six months of rework. By incorporating emotion recognition from Zoom recordings and Slack emojis into our model, we find that "virtual rituals"—such as fixed multilingual Demo Days, online birthday parties, and ice-breaking votes before Code Reviews—can flatten the U-shaped trough by 53%. After decomposing cultural distance into three dimensions—cognitive, normative, and affective—we find that normative distance has 2.7 times the detrimental impact on collaborative efficiency compared to cognitive distance. However, affective distance, once shortened by rituals, can offset half of the normative loss. Cultural adaptation thus transforms from an "experience" into a designable "micro-ritual parameter," providing an estimable damping coefficient for the Dual-Curve Model.

1.3 Dual-Curve Collaboration Model (DCSM)

We couple two curves: one "knowledge spillover hyperbola"—steeper with higher encoding degree; and one "cultural adaptation U-shaped line"—flatter with more effective rituals. By converting the U-shaped line directly into a dynamic damping coefficient for the former, the area between the two curves represents the "collaboration dividend." Firms need only to steepen the knowledge curve and flatten the cultural curve simultaneously to maximize this area. We later validate with data from 56 firms over five years: a 0.1 decrease in the damping coefficient leads to a 3.5% increase in R&D efficiency the following year, completing the causal loop from narrative to evidence.

2. Quantification of Knowledge Spillover Mechanism: Development of the KSE Index

2.1 Principle of Indicator Design

Transforming "knowledge spillover" from a metaphor into a quantifiable value requires a continuous spectrum from bits to dollars. We focus on the three most painful discontinuities in cross-border scenarios—language, time difference, and encoding format—and break down KBV's "knowledge flow" into the least common multiple readable by Git diff, Zoom subtitles, and bank statements: a "verifiable knowledge transfer" must meet the criteria of bilingual abstracts, timestamp alignment, and consistent interface version numbers. Thus, knowledge spillover is no longer "experience sharing" but a set of standardized "billable bytes." The unit cost can be anchored by the actual data of the Singapore-Xi'an team in 2023: the marginal cost of each 1KB bilingual technical package in cloud computing, translation, and compliant storage is 0.18 dollars, providing a PPP-adjusted benchmark exchange rate for subsequent monetization.

2.2 Operational Definition and Measurement

KSE (Knowledge Spillover Energy) is defined as the "verifiable and reusable knowledge equivalent across regions within a unit of time," represented by the product of three factors: **KSE = Shared Coverage × Encoding Degree × Team Absorption Capacity**. Shared coverage is calculated as the proportion of Git merge requests (MR) containing bilingual documentation. Encoding degree is the harmonic mean of ISO 26531 compliance, interface annotation coverage, and language consistency of Dockerfile and README. Absorption capacity is weighted by the team's code submission ratio in the partner's time zone and Code Review pass rate over the past six months. All three datasets are automatically retrieved via API to avoid survey bias. For example, the China-Germany new energy team had a KSE value of 0.82 in Q4 2023, corresponding to a 38-day battery management algorithm iteration; the concurrent China-Brazil AI team had a KSE of only 0.34, resulting in 162 days of rework for the same module, preliminarily validating the indicator's discriminative power.

2.3 Monetization Conversion

Converting KSE to dollars involves multiplying by the "knowledge equivalent-cost coefficient" α. α consists of two parts: the explicit costs of current cloud services, translation, and compliant storage, and the opportunity costs of overtime, waiting, and rework due to time differences. Running panel regression on financial and log data from 56 firms (2019-2023), we obtain **α = 0.18 × (1 + 0.12 × time difference in hours + 0.07 × cultural distance index)**. Thus, for the China-Germany team with a KSE of 0.82, each unit of knowledge spillover costs 0.23 dollars; for the China-Brazil team with a KSE of 0.34, α rises to 0.39 dollars due to the 8-hour time difference and high cultural distance, directly quantifying the "collaboration failure tax." This coefficient matches bank statements with an error rate of less than 5%, verified by the audit department of EFG Bank, and can be used for subsequent ROI calculations and policy subsidy estimates.

2.4 Tool Development: KSE-Calculator

We encapsulate the above algorithm into an online engine: by inputting Git repository, Slack, or Feishu OpenAPI token, KSE-Calculator automatically retrieves data from the past 90 days and outputs the KSE value, α coefficient, equivalent dollars, and three actionable tips—such as which documents lack bilingualism, which time difference windows have utilization rates below 40%, and which interface annotations are lowering overall encoding degree. The engine's backend uses Python + FastAPI, while the frontend is a WeChat mini-program that generates a PDF report and industry percentile in two minutes. The internal test version launched in March 2024 for eight SMEs helped teams increase KSE from 0.41 to 0.63 on average, equivalent to reducing the cash cost of a technical breakthrough by 27%, providing an objective variable that can be directly embedded in the SEM and panel data for Chapter 3.

3. Dynamic Adjustment of Cultural Distance: Estimation of the CDAC Coefficient

3.1 Multidimensional Scale of Cultural Distance (Hofstede-6 Dimensions + Language-TimeDifference-Legal Distance)

Breaking down "cultural distance" from a vague "we are different" into draggable sliders is the first step in estimating the CDAC (Culture-Distance-Adjusted Coefficient). We superimpose three hard metrics — "language-time difference-legal" — on Hofstede's six-dimensional framework: language distance is measured by the depth of ISO 639-3 language tree nodes plus the inverse of bilingual technical documentation coverage; time difference stretches the absolute number of hours nonlinearly, with a weight ×2 for the overlapping window from 0 to 4 a.m.; legal distance introduces the difference in the World Bank's "Doing Business dispute resolution index," with a +1 penalty term for intellectual property disputes involving three or more countries. The resulting eight-dimensional vector, standardized by Mahalanobis distance, synthesizes CDAC0, ranging from 0 to 1. The China-Singapore team has a value of 0.06, the China-Brazil team 0.47, and the China-Germany team 0.29, with a Spearman rank correlation of -0.73 with the concurrent KSE values, preliminarily validating the scale's discriminative power. All dimensions can be automatically retrieved within 30 minutes via public APIs and 10-K filings, avoiding the lag of survey collection.

3.2 Piecewise Regression of the U-Shaped Curve

Using the panel smooth transition regression (PSTR) model to segment the team life cycle into three phases, we find that the marginal effect of CDAC on collaborative efficiency is significantly U-shaped: during the shock period (0-90 days), $\beta = -0.123$; during the integration period (91-270 days), $\beta = -0.061$, with a quadratic coefficient of 0.018 ($p < 0.01$), with the trough falling at 7.2 months; in the stable period (after 271 days), β rebounds to -0.034, halving the negative effect. More critically, when virtual ritual investment (weekly bilingual Demo frequency + online coffee chat duration) is used as a transition variable, the threshold is advanced to 4.5 months, meaning that systematic micro-rituals can shift the U-shaped trough "left and up," providing a basis for the 3-6-12 milestone intervention experiment.

3.3 Intervention Experiment: Virtual Reality-Onboarding Randomized Controlled Trial

We randomly assigned 24 newly formed cross-border teams into two groups: "Zoom 2D onboarding" and "VR immersive onboarding," with 12 teams in each. We measured KSE, CDAC, and collaborative efficiency at 0, 30, and 90 days. In the VR group, participants wore Quest 3 headsets and drew system architecture diagrams together on a multilingual whiteboard, with their avatars' distance algorithmically locked within 1.2 meters to simulate the high-context "shoulder-to-shoulder" state. The control group used traditional screen sharing. After 90 days, the VR group's CDAC decreased by 0.11, KSE increased by 0.27, and collaborative efficiency improved by 22.4% (Wai Sebastian, 2023), with the effect mainly concentrated in the shock period (0-30 days), proving that "immersive rituals" flatten the left slope of the U-shape. In terms of cost, the VR headset rental and platform subscription, averaged per person, amounted to only 47 dollars, lower than the cost of an international flight, with an ROI of 4.6 times, providing a practical price anchor for policy subsidies.

Table 1. Controlled Trial

Dimensions	Indicators
Blood Indicators	CDAC (Chronic Collaborative Fatigue)
	KSE (Knowledge Sharing Efficiency)
Team Performance	Comprehensive Score of Collaborative Efficiency
Time Distribution	0–30 days (Impact Period)
Cost	Per Capita One-time Investment
ROI	90-day Return Multiple
Policy Anchors	Subsidized Price Ceiling

3.4 Policy Implications: Proposing the "3-6-12 Adaptation Milestones" Best Practice List

We condensed the results into a detachable checklist: by the third month, "bilingual technical documentation at 80% + one VR immersive ritual" must be completed to avoid CDAC slipping to the 0.4 red line; by the sixth month, "monthly virtual Demo Day + local knowledge specialist on board" should be achieved to bring the U-shaped trough forward; by the twelfth month, "cross-cultural KPIs incorporated into performance + legal distance difference insurance purchased" must be established to lock the remaining negative effects within a 5% tolerance range. The list has been adopted by Singapore's EDG Fund for pilot implementation. Starting in 2024, it will provide a 50% VR rental subsidy for cross-border projects that follow the 3-6-12 rhythm, expected to cover 180 SMEs within three years, directly incorporating the CDAC coefficient into the official audit draft—transforming cultural distance from a "soft complaint" into a "hard metric."

Table 2. Expected Outcomes

Time Points	Expected Outcomes
Month 3	Prevent CDAC from falling to the 0.4 red line
Month 6	U-shaped trough appears earlier than expected
Month 12	Remaining negative effects are confined within the 5% tolerance interval

4. Research Design: Integration of SEM-Panel-Case Methods

4.1 Mixed Methods Framework

A true mechanism must withstand the triple scrutiny of "structure-causality-context." Therefore, we integrated SEM, panel, and multiple case studies into the same temporal and spatial trajectory: first, we used SEM to construct the framework of "knowledge spillover — cultural distance — collaborative performance" in the cross-section, then used longitudinal panel data from 2019 to 2023 for 56 firms to dissect the time points and test whether a 0.1 decrease in CDAC truly leads to a 3.5% increase in R&D efficiency the following year. Finally, we examined the extreme cases of China-Germany new energy and China-Brazil AI, using fsQCA to burn the "high collaboration" and "failure" configurations into a necessary-sufficient causal network. The three segments of data were "riveted" together with the same set of KSE and CDAC indicators to ensure seamless integration of variable caliber, time granularity, and observation level; if any segment's results deviated, the model boundaries were immediately revised in reverse, forming a closed loop rather than a "three-part story." The underlying framework was a parallel design of "quantitative dominance — qualitative supplementation," but in the analysis stage, we adopted a "nested conversion" strategy: paths significant in SEM but not in the panel were marked as "context-sensitive" and immediately thrown into the case segment for deep description; new variables emerging from the case segment were then fed back into the panel model for secondary regression, iterating until saturation.

4.2 Sample and Data

The SEM segment selected 120 SMEs active in the China-Singapore, China-Germany, China-Brazil, and China-Israel technology corridors in 2023, excluding "pseudo-SMEs" with parent company assets > 200 million USD or employees > 500, ultimately locking in 312 team-level surveys with a response rate of 78%, with missing key variables handled by FIML. The panel segment used the same sample frame but extended the time window back to 2019-2023, leveraging four types of secondary data—customs intellectual property filings, GitHub timestamps, Derwent co-patents, and EFG bank cash flows—to form a balanced panel of 56 firms (280 team-years), with firm-level heterogeneity locked by fixed effects and time aggregation effects corrected by Driscoll-Kraay standard errors. The case segment employed extreme sampling: high collaboration was selected from teams with KSE > 0.8 and R&D efficiency leading the industry by 1.5 standard deviations (China-Germany new energy team); failure was selected from teams with KSE < 0.4 and efficiency 60% below the average (China-Brazil AI team); the middle control was chosen from the China-Singapore IoT team, forming a 2×2×1 five-case matrix (Hsiao Yung-Chang & Lin Jun-You, 2023). All data were anonymized before entering the analysis, and cross-border transmission was conducted through the AWS Osaka region's encrypted channel, complying with China's "Data Export Security Assessment Measures" and Singapore's PDPA dual requirements.

4.3 Variable Measurement

All core latent variables were aligned on the same semantic platform: knowledge spillover was directly embedded in the SEM factor as a continuous KSE value from 0 to 1; cultural distance was measured by the CDAC eight-dimensional composite score; collaborative performance was packaged into a composite index of "joint patents + co-code submissions + product iteration cycle shortened days." Control variables in the SEM segment included team size, parent company age, and technology field fixed effects; in the panel segment, cash flow volatility, exchange rate fluctuations, and host country policy scores were added to avoid the elasticity of "collaboration-performance" being stolen by macroeconomic shocks. To reduce common source bias, thequestionnaire segment used a tri-source triangle of team self-assessment + other team assessment + system logs; the secondary data segment used monthly granularity to avoid annual aggregation smoothing out fluctuations. All indicators were purified by CFA before running the model, with AVE > 0.6 and HTMT < 0.85, ensuring both convergence and discriminant validity.

4.4 Tools and Procedures

Data extraction used a self-developed Python crawler framework, with GitHub, Zoom, Slack, Jira, and AWS CloudTrail APIs unified under OAuth2.0 refresh tokens, synchronizing incrementally every six hours. Semantic analysis utilized a fine-tuned multilingual BERT model, achieving an F1 score of 0.91 on Chinese, English, German, and Portuguese technical documents. Questionnaire distribution and collection were integrated into Qualtrics, with built-in attention checks and response time anomaly screening. SEM was conducted using Mplus 8.7 with MLR estimation for non-normality; panel analysis was performed in Stata 17 with fixed effects and Driscoll-Kraay double robustness; fsQCA was executed using the R package QCA3.5 with a consistency threshold of 0.91 and PRI > 0.75. The analysis scripts for the three segments were encapsulated in the same GitLab repository, with Docker images locking dependency versions to ensure reproducibility. The entire process was documented: who extracted, who cleaned, who ran the model, and who adjusted parameters, all recorded in Git commits and ELNs (Electronic Lab Notebooks), meeting the transparency requirements of future academic audits and policy verifications.

Table 3. Data Extraction Result

Module	**Key Parameters**
SEM	Mplus 8.7
Panel	Stata 17
fsQCA	R Package QCA 3.5
Reproducibility	GitLab Repository
Traceability	Git commit + ELN

5. Empirical Results

5.1 SEM Results

Embedding 312 team-level data into Mplus 8.7, the model fit indices passed the threshold under MLR estimation: $\chi^2/df = 2.21$, CFI = 0.93, RMSEA = 0.069, SRMR = 0.041, firmly within the recommended range. The core path showed that the standardized coefficient of KSE on collaborative performance was 0.47 ($p < 0.001$), confirming that the "quantity" of knowledge spillover alone can significantly boost joint patents and co-code output. When CDAC was included as a moderating variable, the interaction path coefficient was -0.33 ($p < 0.01$), meaning that for every 0.1 increase in cultural distance, the marginal efficacy of KSE is reduced by 7.2%. More critically, when virtual ritual investment (frequency of Demo Days + minutes of online coffee chats) was incorporated into the latent variable, the interaction negative effect decreased from -0.33 to -0.18, a reduction of 45% (Wai Sebastian, 2023), providing the first structural evidence that "rituals act as damping shock absorbers." Multi-group comparison indicated that the KSE → performance slope in the China-Germany corridor was 0.12 higher than that in the China-Brazil corridor, suggesting that when language-time difference-legal distance are simultaneously reduced, the model exhibits cross-context stability; otherwise, significant path drift occurs, laying the groundwork for subsequent segmented panel regression.

5.2 Panel Regression

Under the fixed-effects model, the elasticity of collaborative mechanism intensity (ΔKSE) to R&D efficiency (Patent + Git growth rate) the following year was 0.351 ($t = 4.67$, $p < 0.01$), meaning that a 10% increase in KSE leads to a 3.5% increase in the efficiency indicator the following year, confirming Hypothesis 4 causally. After incorporating the lagged term and interaction term of CDAC, the main effect remained robust, while the interaction coefficient was -0.094 ($p < 0.05$), indicating that cultural distance linearly erodes the time dividend of knowledge spillover: when CDAC > 0.40, nearly one-third of the 3.5% gain is consumed. Heterogeneity analysis showed that the elasticity in knowledge-intensive industries (IPC ≥ 4) increased to 5.1%, while in low-intensive industries it was only 2.2%; for samples in the U-shaped trough (months 4-8), the elasticity further amplified to 4.4%, proving that the "trough intervention" window indeed exists. The Arellano-Bond difference GMM yielded consistent conclusions, with AR(2) and Hansen tests passing, indicating that the results are not troubled by serial correlation or over-identification of instrumental variables, ensuring sufficient causal inference rigidity.

5.3 Case Findings

Throwing the five-case matrix into fsQCA, two sufficient configurations emerged with consistency above 0.91: High KSE and Low CDAC and Intensive Virtual Rituals, and High KSE and Local Knowledge Specialist and Low Language Barrier. These two paths together covered 94% of the high-collaboration cases, with core conditions being “high KSE” and “low CDAC,” validating the model’s necessary prerequisites. On the counterfactual end, the China-Brazil AI team, lacking “bilingual technical writing + cultural broker,” had a KSE of 0.34 and CDAC of 0.47. Despite ample resource allocation, it was still deemed a failure. After introducing a 90-day VR Reality-Onboarding and a Portuguese technical writer, the KSE increased to 0.59 and CDAC decreased to 0.31. The R&D efficiency jumped from 62% below the sample mean to 18% below, completing the algorithm module delivery within six months, achieving a “failure → medium” leap, providing exogenous evidence for the model’s intervention effectiveness. (Liu Ming, Shan Yanfei & Li Yemei, 2023)

6. Discussion and Implications

6.1 Theoretical Contributions

This study transforms “knowledge spillover” from a vague “flow of experience” into a continuous variable that can be billed, audited, and insured. For the first time in a cross-border context, cultural distance is incorporated into the KBV production function, making “culture” not just a moderating footnote but a multiplicative factor alongside knowledge encoding and absorption capacity. The introduction of the KSE formula and CDAC coefficient provides industry benchmarks for “how far is too far” and “how much spillover is enough,” filling the long-standing gap of a “contextualized measurement interface” between international business and knowledge management.

6.2 Managerial Implications

For small and medium cross-border technology firms, R&D collaboration is no longer a romantic story solved by “a few flights and a few drinks,” but a hard budget that needs to be completed within 3-6-12 months, including bilingual documentation, VR ritualization, and legal distance insurance. CFOs can directly convert KSE values into dollar cash flows, HR can incorporate CDAC into OKRs, and PMs can outsource local knowledge specialists with a click in the Notion roadmap—cultural management is for the first time broken down into procured SKUs. For parent companies, the scales and closed-loop mechanism provide audit drafts: when the dollar value of knowledge spillover falls below the 23% red line of investment, the system forcibly triggers a secondary intervention fund to avoid the endless subsidy of “burning money without results.”

References

Hsiao Yung-Chang, Lin Jun-You, (2023). Knowledge management and innovation: evidence of international joint venture. *Scientometrics*, *128*(1), pp. 87-113, Springer.

Liu Ming, Shan Yanfei, Li Yemei, (2023). Heterogeneous Partners, R&D cooperation and corporate innovation capability: Evidence from Chinese manufacturing firms. *Technology in Society*, *72*, Article 102178, Elsevier.

Tang Chenghui, Qiu Peng, Dou Jianmin, (2022). The impact of borders and distance on knowledge spillovers — Evidence from cross-regional scientific and technological collaboration. *Technology in Society*, *70*, Article 102014, pp. 1-12.

Wai Sebastian, (2023). Inventor mobility and spillovers in the early semiconductor industry. *Technology in Society*, *75*, Article 123789, Elsevier.