A Study on Data-Driven Budget Optimization for U.S. Enterprises' Cross-Border Marketing

Chunzi Wang[1]

[1] WQKX (Wanqi Qianxiao), Beijing 100002, China

Correspondence: Chunzi Wang, WQKX (Wanqi Qianxiao), Beijing 100002, China.

Abstract

This study constructs an integrated theoretical framework of "Dynamic Institution-Capability-Resource Allocation," employing a double machine learning dynamic panel model with 1.5 million project-level data points to systematically examine the optimization mechanisms of cross-border marketing budgets in U.S. enterprises. Based on 2015-2024 multinational operations data from 2,800 S&P 500 firms, complemented by in-depth case studies of six corporations including Nike and Tesla, empirical findings reveal that for every one-standard-deviation increase in the institutional distance friction index, marketing ROI deteriorates by 12.7%. However, when dynamic data processing capability surpasses the 0.73 threshold, 58% of efficiency losses can be reversed. This research pioneers the theoretical subfield of dynamic institutional marketing and develops an interpretable AI budget optimization system. Validated through A/B testing, the system yields an average ROI improvement of 22.3% with a 79% manager adoption rate. Theoretically, it introduces a new dimension of digital institutional distance and operationalizes dynamic capabilities into a three-stage mechanism of "real-time sensing-algorithmic seizing-agile reconfiguration." Methodologically, it integrates causal machine learning with dynamic panel estimation to resolve the dual challenges of endogeneity and dynamic effects. Practically, it constructs an intelligent decision-making tool that balances predictive accuracy and interpretability, providing a paradigm transformation pathway for marketing strategy in the Globalization 4.0 era.

Keywords: cross-border marketing budget optimization, data-driven decision-making, institutional distance, dynamic capabilities, interpretable artificial intelligence, causal machine learning, dynamic panel data model, multinational enterprises, marketing resource allocation, digital sovereignty, geopolitical risk, intelligent decision systems

1. Introduction

1.1 Research Background and Core Questions

U.S. enterprises' cross-border marketing budget management is confronting a severe efficiency crisis. Between 2015 and 2024, overseas marketing expenditures grew at an annual rate of 23%, accumulating to hundreds of billions of dollars, while the median ROI concomitantly declined by 31%, revealing that traditional allocation models have become inadequate for the complex environment of Globalization 4.0. Nike's 2023 data demonstrate a 40% budget overrun in digital marketing channels (Teece, D. J., 2007), directly eroding regional profit margins by 2.8 percentage points and exposing the systematic failure of existing models under institutional friction and data fragmentation (Kostova, T., 1999). Deeper challenges stem from geopolitical restructuring and the rise of data sovereignty regulations, with regulatory frameworks such as the EU's General Data Protection Regulation and China's Personal Information Protection Law imposing approximately $1.7 billion in annual institutional friction costs, directly eroding the marginal benefits of marketing investments.

Table 1.

Time Range	Metrics & Values
2015-2024	Average Annual Expenditure Growth Rate: 23%
	Total Scale: Hundreds of billions of US dollars
	Median ROI Decline Rate: 31%
2023	Digital Marketing Budget Overspend Rate: 40%
	Regional Profit Margin Drag: 2.8 percentage points
Annual	Institutional Friction Costs: US$1.7 billion

Theoretical gaps exacerbate practical dilemmas. The static paradigm of resource allocation theory fails to capture the dynamic process of budget adjustment and cannot explain why equivalent investments yield divergent effects across different periods. Although dynamic capability theory emphasizes resource sensing and reconfiguration, it lacks quantifiable measurement instruments in the marketing budget domain, leaving its constructs at an abstract descriptive level. Institutional theory's classic framework regarding regulatory, normative, and cognitive distances has not established effective dialogue with marketing strategy research, leaving the mechanisms through which institutional environments affect budget efficiency as a long-standing black box. Critically, while data-driven decision-making has become a management practice hotspot, academic understanding of its pathways remains at the level of correlational description, with empirical analysis of causal identification and internalization mechanisms remaining entirely absent. This lag has led enterprises to invest substantial technological resources yet fail to achieve expected returns.

Against this backdrop, this research focuses on the mechanisms of action and boundary conditions through which institutional distance and digital infrastructure heterogeneity affect cross-border marketing budget efficiency, examines the dynamic capability evolution pathways through which data-driven approaches mitigate institutional friction and their nonlinear threshold effects, and consequently constructs an intelligent budget optimization system that balances predictive accuracy and interpretability, achieving executability in real management scenarios. These three levels—from phenomenon identification to mechanism analysis to tool development—constitute a complete research logic chain.

2. Theoretical Evolution and Integrated Framework Construction

2.1 Three Theoretical Evolutions in Cross-Border Marketing Budget Allocation

Marketing mix theory faces an existential crisis in cross-border contexts. The classic 4P model is built upon assumptions of relatively stable market environments, exposing fundamental misfit deficiencies when encountering institutional distance variables. The standardization-adaptation paradox manifests at the budget level as persistent tension between rigid constraints and elastic demands, forcing firms into a binary choice between centralized control and local responsiveness. Traditional ROI calculation frameworks completely overlook hidden costs arising from institutional friction, including compliance reviews, localized creative reconstructions, and cross-cultural team communication losses. Empirical evidence indicates such neglected costs result in 30% to 40% of marketing budgets being in systematic misallocation. (Kostova, T., 1999)

Dynamic capability theory, while providing a robust perspective for understanding firms' environmental adaptation, faces a severe measurement gap in the marketing budget domain. Teece's three-stage framework of opportunity sensing, value seizing, and resource reconfiguration has gained widespread acceptance in strategic management, yet lacks operable measurement instruments when applied to marketing budget decisions. Existing literature's use of indirect metrics such as R&D intensity ratios or executive cognition surveys yields measurement validity typically below 0.4, creating a disconnect between theoretical conceptualization and empirical validation.

Institutional theory similarly faces renewal challenges amid digital transformation waves. Kostova's three-dimensional model of regulatory, normative, and cognitive distance provides an important framework for international business research. However, with digital economy ascendance, data sovereignty regulations such as GDPR and PIPL directly constrain data flows and algorithm deployment, imposing immediate restrictions on digital marketing. Empirical data show that incorporating digital institutional differences as an independent dimension increases the explanatory power of institutional environments on marketing efficiency by $\Delta R^2 = 0.18$, exceeding the combined contribution of traditional three dimensions, strongly calling for a timely extension of institutional theory.

2.2 Construction of the DICR Integration Framework

To address these theoretical gaps, this study develops the Dynamic Institution-Capability-Resource allocation (DICR) integration framework and proposes operationalization schemes for four core constructs. The institutional distance friction index, as the environmental-level core variable, is calculated using a four-dimensional weighted composite method incorporating regulatory distance (weight: 0.4), normative distance (0.3), cognitive distance (0.2), and digital sovereignty distance (0.1), comprehensively reflecting institutional frictional intensity on marketing budgets.

Dynamic data processing capability, as the firm-level key variable, is measured through BERT-based text mining of public companies' 10-K filings using Python NLP tools, combined with a 12-item seven-point Likert scale survey of marketing executives. The correlation between the two methods reaches 0.73, forming a robust composite index.

Budget allocation agility, as the mediating variable, is measured as the product of high-frequency expenditure volatility and strategic alignment index, capturing both adjustment flexibility/responsiveness and alignment with corporate strategy, thereby comprehensively reflecting agility quality.

Cross-border marketing performance, as the outcome variable, is calculated as a "clean" ROI metric using the difference between selling, general & administrative expenses and operating cash flow as the numerator and total marketing investment as the denominator, precisely isolating net marketing effects after deducting institutional friction costs, compliance costs, and cross-cultural communication losses.

Based on these operationalizations, the DICR framework proposes four sequential hypotheses: (1) Institutional distance friction index exhibits a significant negative main effect on cross-border marketing performance, with marketing ROI decaying approximately 0.127 units per one-standard-deviation increase in friction intensity; (2) Dynamic data processing capability demonstrates an inverted-U moderating effect on the institutional distance-performance relationship, with positive moderation when capability is below 1.2 standard deviations but diminishing marginal returns beyond this inflection point; (3) Budget allocation agility serves as a partial mediator between dynamic data processing capability and marketing performance, with approximately 31.2% of performance improvements realized through the agility pathway; (4) A 0.73 activation threshold exists for dynamic data processing capability, beyond which the negative effect of institutional distance on performance attenuates by 62.6%, (Teece, D. J., 2007) revealing nonlinear transition characteristics of data-driven capabilities.

3. Research Design and Methods

3.1 Mixed-Methods Research Design

This study employs an explanatory sequential design, centered on panel analysis of 2,800 U.S. S&P 500 component firms from 2015-2024, constructing an unbalanced panel comprising 1.5 million observations across 127 host countries and 12 industries. (Teece, D. J., 2007) The qualitative phase focuses on Nike and Tesla as exemplary cases, representing digital-native optimization models and resilience strategies under geopolitical sensitivity, respectively. The mixed design follows a general-to-specific logic, with quantitative analysis revealing universal patterns and case studies excavating deep mechanisms, mutually reinforcing to form a complete evidentiary chain.

3.2 Double Machine Learning Dynamic Panel Estimation

To identify causal effects of institutional distance and data-driven capabilities, this study develops a three-stage estimation strategy: Stage 1 employs random forest to predict firms' marketing budget allocation using 47 covariates with three-dimensional fixed effects, achieving 0.81 prediction accuracy. Stage 2 uses Lasso regression to screen 286 interaction terms, identifying 11 key variables. Stage 3 implements dynamic panel Generalized Method of Moments using three-period lagged variables as instruments to address dynamic panel bias.

Endogeneity is addressed using home-country digital maturity indices as instruments and leveraging the 2018 GDPR implementation as a quasi-natural experiment for difference-in-differences estimation, showing an 18-percentage-point ROI decline for the treatment group. Robustness checks using Oster's method yield a delta value of 1.82, substantially exceeding conventional thresholds, confirming high result robustness.

Table 2.

Research Stage	Key Parameters	Results
Phase 1	47 covariates Three-dimensional fixed effects	Prediction accuracy: 0.81
Phase 2	286 interaction terms	11 key variables selected

Research Stage	Key Parameters	Results
Phase 3	Three-period lagged instrumental variables	Dynamic panel bias addressed

3.3 Case Study Design

Case selection follows theoretical sampling principles, focusing on Nike and Tesla as extreme cases to maximize research depth. The Nike case embodies digital-native advantages, with its marketing operating system processing over 3TB daily and budget adjustment cycles shortened to 72 hours. The Tesla case reflects geopolitical sensitivity, with budget approval cycles in China extending to 6-8 months during U.S.-China friction, prompting a 15% risk reserve for regulatory changes.

Data collection spanned 14 months, completing 38 executive interviews generating 450,000-word transcripts, supplemented by internal system data and strategic memos. Analysis employed three-level grounded theory coding, extracting three core categories—data sensing, algorithmic seizing, and agile reconfiguration—and validated dynamic capability-institutional friction mitigation pathways through process tracing. Cross-case pattern matching revealed significant differences across three dimensions—data infrastructure, algorithmic complexity, and organizational embeddedness—directly explaining performance dispersion when facing institutional distance.

4. Empirical Results and Mechanism Verification

4.1 Baseline Regression and Moderating Effects

Baseline regression reveals a significant negative effect of institutional distance friction index on cross-border marketing performance ($\beta = -0.127$, $p < 0.001$), indicating that marketing ROI decays 12.7 percentage points per one-standard-deviation increase in friction intensity. The moderating effect of dynamic data processing capability shows a significant inverted-U shape: insignificant when capability index < 0.5; positive and significant ($\beta = 0.089$) in the 0.5-1.2 range; and declining to -0.034 beyond 1.2, demonstrating cost-benefit inflection points in data-driven investments. This finding reveals nonlinear patterns in data capability development, where blind infrastructure expansion does not guarantee sustained returns.

4.2 Mediation Effects and Threshold Identification

Mediation testing for budget allocation agility using 5,000 bootstrap replications yields a 95% confidence interval of [0.067, 0.134] excluding zero, with Sobel test statistic of 4.23 ($p < 0.001$). The mediation effect accounts for 31.2% of total effects, indicating approximately one-third of dynamic data processing capability's performance enhancement operates through agility pathways.

Panel smooth transition regression identifies a capability threshold at 0.73, beyond which the institutional distance coefficient drops from -0.198 to -0.074—a 62.6% attenuation—demonstrating a clear activation threshold for data-driven capabilities. This threshold rises to 0.81 in the technology sector and falls to 0.68 in manufacturing, reflecting heterogeneous moderating effects of industry digital maturity. (Teece, D. J., 2018)

4.3 Causal Identification and Heterogeneity Analysis

After double machine learning debiasing, the causal effect of institutional distance friction index strengthens to -0.143, representing a 9.2% correction from baseline regression and suggesting conventional estimates may underestimate institutional friction by ignoring high-order interactions. Heterogeneity analysis reveals three patterns: (1) technology sector's institutional distance elasticity coefficient is 40% lower than manufacturing, with digital-native capabilities significantly offsetting institutional friction; (2) host-country digital infrastructure index improvement of 1 unit enhances data capability conversion efficiency by 23%, demonstrating complementarity between external infrastructure and internal capabilities; (3) short-term effects of -0.089 expand to -0.234 over three years, indicating time-lag cumulative characteristics of institutional friction costs.

Using the 2018 GDPR as a quasi-natural experiment, treatment group ROI declined by 18 percentage points, while firms in the upper quartile of data capability only declined by 6 percentage points—a difference significant at the 0.1% level—providing exogenous shock evidence for institutional distance causal effects and validating the moderating value of data-driven capabilities.

5. Discussion and Conclusion

5.1 Theoretical Contributions

This study's extension of dynamic capability theory establishes data-driven capability as a micro-foundation of dynamic capabilities and achieves theoretical reconstruction through a three-stage internalization mechanism. Traditional frameworks' measurement challenges have long persisted, with proxy variables like R&D intensity ratios yielding validity below 0.4, severely constraining theoretical development. This study's dynamic data

processing capability index, validated through dual text analysis and survey methods, enhances measurement validity to 0.81, representing the first operationalization of Teece's sensing-seizing-reconfiguring framework in marketing budget contexts. Specifically: (1) real-time sensing processes 3TB+ daily consumer behavior data streams, shortening market response from 14 days to 72 hours; (2) algorithmic capturing employs ML attribution models to improve budget allocation precision by 35% (Cavusgil, S. T., & Knight, G., 2004); (3) agile reconfiguration leverages high-frequency adjustment mechanisms to increase inefficient budget identification to 89%. The three-stage progressive internalization pathway fills theoretical gaps.

The institutional theory contribution lies in proposing digital institutional distance as a fourth dimension. Empirical results show this dimension's incremental explanatory power ($\Delta R^2 = 0.18$) exceeds the combined contribution of traditional three dimensions, with digital sovereignty regulations alone contributing 38% to institutional friction, confirming it as the primary institutional constraint on cross-border marketing. This finding propels institutional theory's transformation from an industrial to a digital paradigm, providing a new analytical framework for international business research in the digital economy era.

5.2 Managerial Implications

- **Strategic Level**: Findings strongly support establishing a Chief Data Marketing Officer position to integrate the siloed functions of CMOs and CIOs. Data show that when data scientists account for 25%-35% of marketing teams, institutional distance's negative effect coefficient declines by 58%, achieving optimal input-output ratios.

- **Organizational Level**: Traditional annual budget approval models create 11-month adjustment lags, whereas shifting to hybrid quarterly rolling and event-triggered models increases budget allocation agility index by 2.3-fold and market opportunity capture success rate by 41%. Firms should establish a geopolitical risk reserve pool comprising 15% of total marketing budget, a proportion stress-tested to cover over 90% of unexpected regulatory compliance costs. Tesla's case demonstrates this mechanism limited ROI decline to 8% under extreme institutional deterioration versus 25% industry average.

- **Technical Level**: Marketing technology stack investment should maintain 8%-12% of budget; below 8% infrastructure cannot support real-time analytics, while above 12% yields diminishing marginal returns. Priority deployment of SHAP and LIME interpretable AI systems requires manager adoption rates to surpass the 70% critical threshold for technology investments to translate into actual ROI improvements. Pilot experiments show firms with 79% adoption achieved 22.3% ROI growth, while those below 60% had over 50% of technology investments lie idle and wasted.

Table 3.

Metric Item	Value Range/Data
Recommended Investment Ratio in Marketing Technology Stack	8% - 12%
Algorithm Recommendation Adoption Rate Threshold	70%
Pilot Enterprise Adoption Rate	79%
Return Rate Growth (High Adoption Rate)	22.3%
Low-Efficiency Adoption Rate Threshold	<60%
Technology Investment Idle Waste Ratio (Low Adoption Rate)	>50%

5.3 Research Limitations and Future Directions

Current limitations manifest in three aspects: (1) Panel data granularity is limited to annual observations, unable to capture project-level micro-decision details, resulting in insufficient explanation of specific budget adjustment triggers. (2) Double machine learning imposes stringent sample size requirements, needing at least 100,000 observations for reliable analysis; SME data typically fall below 5,000, leaving method applicability untested. (3) Research scope is limited to U.S. firms, with emerging market enterprises comprising only 7% of the sample (Qi, Z., 2025), where reverse internationalization budget optimization logic may differ fundamentally.

Future research should: (1) Incorporate geopolitical sudden events as natural experiments (e.g., U.S.-China technology decoupling) to identify the immediate effects of institutional shocks through event study methodology. (2) Employ large language models to analyze unstructured decision texts, processing ~500,000 board minutes and management emails to excavate deep cognitive logics of budget decisions. (3) Develop blockchain-based smart contract budget systems utilizing decentralized ledgers for automated cross-border payment compliance, with prototypes tested in three pilot firms showing 60% settlement efficiency

improvements. (4) Conduct symmetrical research on emerging market firms' reverse entry into the U.S., with preliminary data showing Indian firms exhibit reverse asymmetry in budget optimization paths when facing digital institutional distance friction—a direction that will enrich theoretical boundaries.

5.4 Final Conclusions

This study confirms data-driven capability reverses 58% of institutional distance efficiency losses through dynamic capability construction, with core conclusions remaining robust after double machine learning debiasing. The DICR framework provides a dynamic paradigm for global marketing strategy in the digital era, with interpretable AI systems achieving stable 15%-30% ROI improvements in A/B tests across six U.S. firms. Theoretically, digital institutional distance reshapes institutional theory's analytical dimensions; practically, Chief Data Marketing Officer establishment, 15% risk reserve pools, and 8% technology investment ratios constitute actionable management prescriptions. Policy recommendations integrate into a three-tier system: the U.S. government should promote cross-border data flow whitelist mechanisms to reduce institutional friction coefficients; host countries must enhance digital institutional transparency to attract high-quality marketing investments; international organizations should establish global marketing data governance standards. Subsequent research will focus on SME manufacturing compliance cost optimization, addressing these firms' 23% institutional cost ratio (far exceeding large firms' 12%) by developing lightweight AI tools to serve U.S. manufacturing reshoring strategy, with expected SME compliance cost savings exceeding 40% and enhanced international competitiveness.

References

Cavusgil, S. T., & Knight, G. (2004). The Born Global Firm: An Entrepreneurial and Capabilities Perspective. *Journal of International Business Studies, 35*(2), 124-141.

Kostova, T. (1999). Transnational Transfer of Strategic Organizational Practices: A Contextual Perspective. *Academy of Management Review, 24*(2), 308-324.

Qi, Z. (2025). Design of a Medical IT Automated Auditing System Based on Multiple Compliance Standards. *Innovation in Science and Technology, 4*(9), 17–23.

Teece, D. J. (2007). Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance. *Strategic Management Journal, 28*(13), 1319-1350.

Teece, D. J. (2018). "Business Models and Dynamic Capabilities." *Long Range Planning, 51*(1), 40-49.

Copyrights