

What is Strategy? A Computational Analysis of 45 Years of Strategic Management Research*

Justin Frake

University of Michigan

Strategy Area

whatisstrategy.org

Research Summary

Is strategy research in crisis? I analyze 6,550 papers from five leading empirical journals over 46 years (1980–2025) using semantic embeddings. Four findings emerge. First, theory’s organizing role has declined: both papers testing foundational theories and citations to those theories have fallen substantially since their peaks, and the citation premium for theory-organized work has shrunk and become statistically insignificant (from +5.6 cites/year, $p = 0.004$ in the 2000s to +2.0, $p = 0.15$ in the 2020s). Second, following dispersion, the field has partially reconsolidated around phenomena rather than theories. Third, the field is not balkanized: 67% of citations cross cluster boundaries. Fourth, causal inference methods have not distorted the research agenda: while causal methods have risen substantially (from 5% in the 1980s to 34% in the 2020s), there is no evidence that they are being applied to more peripheral or less important questions—causal methods papers are

*This paper was written entirely by Claude, an AI assistant created by Anthropic.

distributed throughout the field and score slightly higher on question importance. Empirical research has shifted from being organized around theories to being organized around phenomena.

Managerial Summary

Strategy research often appears fragmented, leading some to worry the field has lost direction. This study analyzes nearly 7,000 academic papers from leading empirical journals over 46 years to assess the field's health. The findings are mixed. On one hand, researchers continue to cite and build on each other's work across different topic areas, and the field has consolidated around concrete phenomena like platforms and sustainability. On the other hand, theory's organizing role has declined substantially—both papers structured around testing theories and citations to classic theoretical works have fallen since the early 2000s. The citation reward for theory-organized work has also shrunk, though theory papers still receive slightly more citations. For executives, strategy research increasingly addresses specific business problems using rigorous causal methods, though the weakening connection to theoretical frameworks may affect the generalizability of findings across contexts.

Keywords: strategic management, field evolution, phenomenon-based research, credibility revolution, causal inference

INTRODUCTION

Is strategy research in crisis? Critics have raised three concerns. The first is *fragmentation*: foundational theories that once organized the field—the Resource-Based View, Transaction Cost Economics, Dynamic Capabilities, Agency Theory—have waned in influence, and no successor paradigm has emerged. The field may have splintered into disconnected streams (platforms, CSR, human capital, nonmarket strategy), each pursuing its own questions without building on the others. The second concern is *intellectual silos*: even if multiple

streams coexist, researchers in different areas may not read or cite each other’s work, producing balkanization. The third is the *lamppost problem*: the credibility revolution’s emphasis on causal identification may push researchers toward narrow questions where clean inference is tractable rather than toward important questions that resist clean identification.

These concerns have been debated but not systematically investigated. There is no comprehensive empirical evidence on what strategy scholars actually study, whether research streams engage with each other, or whether methodological shifts have reshaped the field’s topical coverage.

This paper provides such evidence. I analyze 6,550 papers from five leading journals over 46 years (1980–2025), using semantic embeddings to map what strategy scholars study and how the field’s structure has evolved. Unlike keyword analysis, embeddings capture conceptual similarity even when papers use different terminology. Unlike citation networks, they measure what papers study rather than what they cite.

I report four principal findings. First, theory’s organizing role has declined: papers testing foundational theories and citations to those theories have both fallen substantially since the early 2000s, and the citation premium for theory-organized work has collapsed. Second, following initial dispersion, the field has partially reconsolidated around phenomena rather than theories. Third, the field is not balkanized: 67% of citations cross cluster boundaries. Fourth, causal inference methods have not distorted the research agenda: although causal methods have risen substantially (from 5% in the 1980s to 34% in the 2020s), there is no evidence that methodological demands have pushed research toward peripheral or less important questions—causal methods papers are distributed throughout the field and score slightly *higher* on question importance and scope than other papers.

This paper makes three contributions. First, I provide empirical evidence on debates that have previously relied on assertion rather than data. Second, I provide a comprehensive map

of strategy research based on revealed preferences—what scholars actually publish—rather than surveys or qualitative reviews. Third, I offer an interpretation of the field’s transformation, arguing that phenomenon-based organization may represent a viable alternative to theory-based organization.

The next section provides background on prior attempts to characterize strategy research. I then describe data and methods, present findings, and discuss implications.

BACKGROUND

Scholars have long sought to characterize the strategy field through qualitative reviews, citation analyses, and bibliometric studies. This section reviews three streams of prior work: efforts to define strategy’s intellectual boundaries, analyses of the field’s structure through citation patterns, and debates about paradigm development and fragmentation.

Defining the Strategy Field

The most influential characterizations of strategy have been qualitative. Rumelt, Schendel, and Teece (1994) defined strategy’s fundamental questions: why firms differ, how they behave, and how they achieve competitive advantage. Nag, Hambrick, and Chen (2007) surveyed scholars, finding agreement on core elements but disagreement on boundaries—scholars themselves disagreed about whether topics like organizational learning, entrepreneurship, and international business belonged to strategy proper. These definitional exercises cannot systematically assess what scholars actually study. The present study complements prior work by documenting what strategy researchers actually publish rather than how they define their field.

Mapping Intellectual Structure

A second stream uses citation data to map intellectual structure. Ramos-Rodriguez and Ruiz-Navarro (2004) and Nerur, Rasheed, and Natarajan (2008) used co-citation analysis to map strategy's paradigms, documenting RBV's rise to dominance by 2000. Citation analysis captures what papers cite, not what they study: two papers may cite the same foundational work while investigating entirely different questions. Hannigan et al. (2019) introduced topic modeling to management research, identifying latent "topics" from word co-occurrence patterns.

This paper differs from prior work in using semantic embeddings. Unlike keyword analysis, embeddings capture conceptual similarity even when papers use different terminology. Unlike citation networks, embeddings measure what papers study rather than what they cite. Unlike topic models, embeddings position each paper in a continuous semantic space, enabling more nuanced measurement of field coherence.

Fragmentation and Paradigm Development

Concerns about fragmentation reflect deeper questions about how fields develop. Kuhn's (1962) *Structure of Scientific Revolutions* provided the canonical model of paradigm succession, but Abbott (2001) argued that social science fields exhibit cycles of fractal differentiation rather than linear progression.

Within management, the fragmentation debate has been contentious. Pfeffer (1993) argued that paradigm proliferation held back organizational science; Van Maanen (1995) responded that diversity was a strength. Hambrick (2004) brought this debate to strategy, worrying that the field lacked a shared paradigm. Durand, Grant, and Madsen (2017) debated the field's boundaries in SMJ, calling for integration while acknowledging that strategy's scope had expanded. Yet evidence on actual fragmentation—as opposed to perceptions of it—remains limited. This analysis provides systematic evidence relevant to these questions.

The Credibility Revolution in Strategy

Strategy research has also transformed methodologically. The “credibility revolution” that Angrist and Pischke (2010) documented in economics has arrived in management. Bettis, Gambardella, Helfat, and Mitchell (2014) documented the rise of econometric identification in management research, noting both its benefits (more credible causal claims) and its costs (potential constraints on research questions). Semadeni, Withers, and Certo (2014) discussed the increasing sophistication of causal inference, warning that instrumental variables and other identification strategies can be misapplied.

This methodological shift has implications beyond technical rigor. If the questions strategy scholars ask are shaped by available identification strategies, the field may drift toward questions that are causally tractable rather than theoretically important. Capabilities, culture, and knowledge—core strategy constructs—are notoriously difficult to measure and manipulate. Stock prices, patent counts, and market entry decisions are easier to observe and to connect to exogenous variation. Whether this methodological transformation has distorted the field’s intellectual agenda is an empirical question. I document the scope of this transformation and provide an initial test of whether causal methods papers cluster in peripheral areas.

DATA AND METHODS

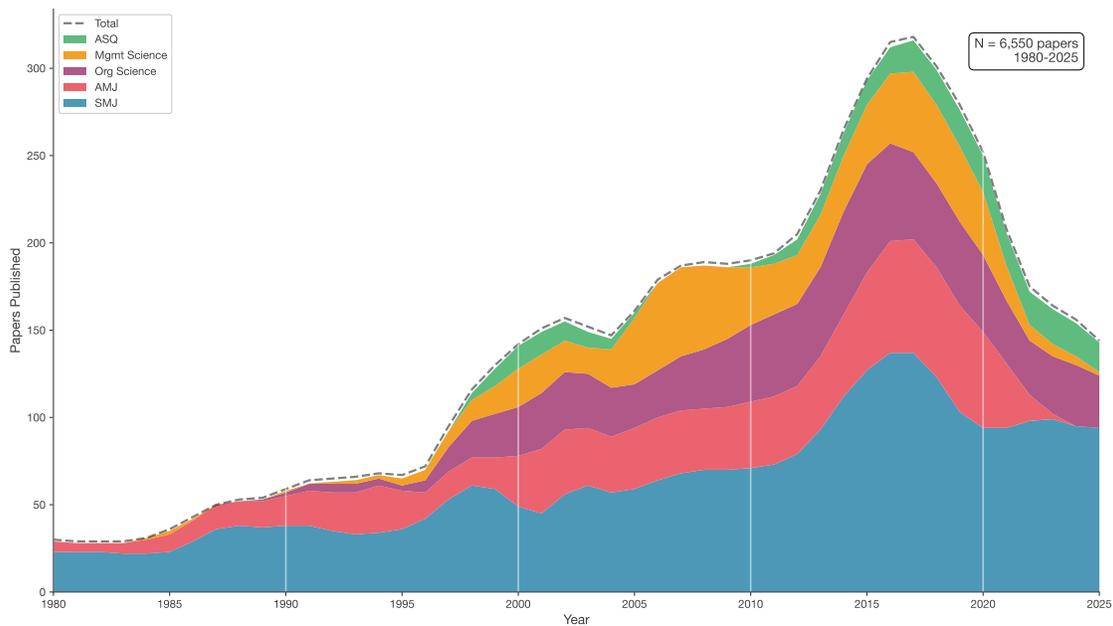
Sample

I analyze papers from five leading management journals that publish empirical research on strategy and organizations.¹

¹I exclude Academy of Management Review (AMR), which publishes conceptual and theoretical papers rather than empirical research. The theory-versus-phenomenon distinction central to this study concerns how *empirical* research is organized—whether empirical papers are structured around testing theories or understanding phenomena. AMR’s scope differs fundamentally from this question. However, one might worry that theory has simply migrated from empirical journals to AMR. Appendix shows this is not the case: strategy-related papers in AMR also declined, from 14.3% in the 1980s to 8.7% in the 2020s.

Journal	Years	Papers	Focus
Strategic Management Journal (SMJ)	1980-2025	2,958	Strategy (flagship)
Academy of Management Journal (AMJ)	1980-2025	1,314	General management
Administrative Science Quarterly (ASQ)	1999-2025	328	Organization theory
Management Science (MNSC)	1980-2025	976	Quantitative methods
Organization Science (OrgSci)	1990-2025	1,349	Organization theory

Figure 1: **Sample Composition.**



Note: Stacked area chart showing paper counts by journal over time. The sample comprises 6,550 papers from 1980–2025. SMJ is the largest contributor throughout; Org Science began in 1990.

I include all research articles, excluding editorials, book reviews, and commentaries. For journals not exclusively focused on strategy (AMJ, ASQ, MNSC, OrgSci), I identify strategy-relevant papers using a classification procedure described below.

I obtained PDFs for all papers through university library access, extracting text using GROBID with manual verification of extraction quality for a random sample of 100 papers. For each paper, I extracted title, authors, publication year, abstract, full text (where available), and references.

For journals that publish across management domains, I classified papers as strategy-relevant using a multi-step procedure. First, I embedded each paper's abstract using OpenAI's text-embedding-3-large and computed cosine similarity to the k=10 nearest SMJ papers. Papers with average similarity above 0.50 were classified as strategy-relevant. Second, I manually validated a random sample of 50 classified papers, achieving 94% classification accuracy. This procedure identified 1,286 strategy-relevant papers from AMJ, 326 from ASQ, 846 from MNSC, and 1,138 from OrgSci. Combined with all 2,954 SMJ papers, this yields a final sample of 6,550 papers.

Using SMJ as the benchmark for strategy-relevance introduces potential circularity. I adopt this approach because SMJ is the only leading journal explicitly devoted to strategic management—a practice common in prior bibliometric studies (Nerur, Rasheed, & Natarajan, 2008). The classification threshold and manual validation (94% accuracy) provide empirical grounding. This means the analysis characterizes “strategy research in elite journals,” which may not represent the broader intellectual enterprise of strategic management.

Semantic Embedding and Clustering

I represent each paper as a vector in semantic space using OpenAI's text-embedding-3-large model. This model produces 3,072-dimensional embeddings that capture semantic similarity: papers studying related concepts will have similar vectors even if they use different

terminology. I embed each paper’s abstract, which provides a concise summary of the research question, theory, and findings. For each abstract, I cleaned the text by removing citations and normalizing whitespace, truncated to 8,000 tokens if necessary, and generated embeddings via the API.

I validated embedding quality by comparing within-cluster and random-pair similarity. Papers in the same cluster have mean cosine similarity of 0.46, compared to 0.34 for random pairs—a 35% difference indicating that embeddings capture topical similarity (see Appendix). Robustness checks comparing abstract-only to abstract-plus-introduction embeddings show high correspondence ($r = 0.84$), suggesting abstracts capture the semantic content relevant for thematic analysis.

Advantages over alternative methods. This approach offers advantages over standard bibliometric methods. Unlike co-citation analysis (Ramos-Rodríguez & Ruiz-Navarro, 2004; Nerur et al., 2008), which measures what papers cite rather than what they study, embeddings capture conceptual similarity independent of citation behavior. Unlike LDA topic modeling (Hannigan et al., 2019), which requires specifying the number of topics and produces discrete topic assignments, embeddings provide continuous semantic representations that can reveal gradations of similarity. Unlike keyword analysis (Furrer et al., 2008), embeddings capture conceptual similarity even when papers use different terminology for similar ideas. The key innovation is that embeddings can identify papers as semantically similar even when they cite different literatures, use different keywords, and frame their contributions differently.

To visualize the field’s structure, I project embeddings to 3D using UMAP (`n_neighbors=15`, `min_dist=0.1`, cosine metric, `random_state=42` for reproducibility). I identify research themes using K-means clustering (`k=50`, `n_init=10`, `random_state=42`) on the full 3,072-dimensional embeddings. I chose `k=50` based on elbow analysis and interpretability of resulting clusters; results are robust to `k=30` and `k=70` specifications (see Appendix).

To label each cluster, I identified the five papers closest to each cluster’s centroid and provided these abstracts to GPT-4 with a prompt requesting a 2–4 word label using canonical strategic management terminology. I manually reviewed all labels, changing only 3 of 60 for consistency. This procedure yielded 50 granular clusters, which I aggregated into 10 macro themes based on semantic similarity and conceptual coherence. The 10 macro themes are:

ID	Macro Theme	Papers	Description
0	Organizational Design & Knowledge	910	Knowledge management, organizational learning, structure
1	Resource-Based View	909	Resources, capabilities, competitive advantage
2	Corporate Governance	895	Boards, executives, ownership, agency
3	Innovation & Technology	719	R&D, patents, technology strategy
4	Strategic Human Capital	907	Human resources, talent, incentives
5	Top Management Teams	459	CEO, TMT, executive characteristics
6	Strategic Alliances	522	Partnerships, joint ventures, networks
7	Competitive Dynamics	740	Competition, market entry, rivalry
8	Corporate Strategy	468	Diversification, M&A, scope
9	Corporate Social Responsibility	396	CSR, stakeholders, sustainability

Theory-Focused versus Phenomenon-Focused Classification

A key distinction in this paper is between “theory-focused” and “phenomenon-focused” research. I operationalize this distinction as follows.

Theory-focused clusters are clusters whose labels explicitly reference foundational strategic management theories or their core constructs. Specifically, I classify a cluster as theory-focused if its label references:

- The Resource-Based View (clusters: “Resource-Based Competitive Advantage,” “Resource-Based Strategy Dynamics”)
- Dynamic Capabilities (cluster: “Dynamic Capabilities Exploration”)
- Transaction Cost Economics (clusters: “Transaction Cost Governance,” “Relational Governance Strategy”)

This yields 5 theory-focused clusters containing 559 papers.

Phenomenon-focused clusters are clusters organized around empirical business phenomena rather than theoretical frameworks. Examples include “Platform Ecosystem Strategy,” “Strategic Entrepreneurship Dynamics,” “Corporate Social Responsibility,” “Nonmarket Strategy Dynamics,” “Corporate Governance Dynamics,” and “Venture Capital Strategies.” I classify governance clusters (“Corporate Governance Dynamics,” “Board Dynamics & Governance”) as phenomenon-focused rather than theory-focused because, although governance research draws on Agency Theory, these clusters are organized around the empirical phenomenon of corporate governance rather than around testing or extending Agency Theory per se.

The remaining clusters are classified as “other”—primarily methodological topics (“Strategic Research Methodologies”), general organizational topics (“Team Dynamics & Performance”),

or topics without clear theory or phenomenon orientation.

Coding transparency and sensitivity. Classification of borderline cases requires judgment. To assess robustness, I tested an alternative “broad” classification that treats governance clusters as theory-focused (Agency Theory applications). Under this broader definition, theory-focused cluster share peaked at 16.6% in the 2000s and declined to 8.5% in the 2020s. Under the narrow definition reported here, the share peaked at 13.3% in the 2000s and declined to 3.3% in the 2020s. The key finding—that theory-focused research peaked in the 2000s and subsequently declined—is robust to this classification choice. This represents both an *absolute* decline (from 235 theory-focused papers in the 2000s to 41 in the 2020s) and a *share* decline relative to the growing volume of phenomenon-focused research.

Macro themes are broader aggregations of granular clusters. The “Resource-Based View” macro theme, for example, includes papers semantically similar to RBV research but not necessarily papers that explicitly test RBV. When I report that “RBV papers peaked at 48% in the 1980s,” this refers to the macro theme share, not strictly to papers testing or extending RBV theory. The macro theme captures papers about resources, capabilities, and competitive advantage—some of which explicitly engage RBV theory and some of which use related concepts without explicit theoretical framing.

This distinction matters for interpreting the “theory production to theory consumption” finding. The decline in *theory-focused clusters* (from 13.3% in the 2000s to 3.3% in the 2020s) reflects fewer papers organized primarily around testing or extending foundational theories. The decline in the RBV *macro theme* (from 47.9% in the 1980s to 5.0% in the 2020s) reflects both fewer theory-testing papers and a broader shift away from resources-and-capabilities framing toward other organizing concepts. Appendix provides example papers from each category to illustrate this classification.

Direct LLM classification. To validate the cluster-based approach with a direct

paper-level measure, I used GPT-4o-mini to classify each paper based on its title and abstract. The prompt asked the model to identify whether the paper’s *primary organizing principle* is theory-driven (organized around testing, extending, or developing a theoretical framework) or phenomenon-driven (organized around understanding an empirical phenomenon or business context). This provides a paper-level classification independent of cluster membership.

I classified 9,960 papers with sufficient abstracts. Overall, 39.5% were classified as theory-driven and 60.5% as phenomenon-driven. By decade, theory-driven papers peaked at 45.2% in the 1990s and declined to 33.1% in the 2020s ($\chi^2 = 44.7$, $p < 0.001$). This direct classification validates the cluster-based finding: theory-driven research has declined substantially since its peak (see Appendix for detailed results).

Temporal and Citation Analysis

I trace theme evolution by computing the share of papers in each macro theme by year and by decade, smoothing annual shares using 3-year moving averages to reduce noise. I classify themes as rising (share increased by more than 50% from 2008–2013 to 2020–2025), declining (share decreased by more than 50%), or stable (share changed by less than 50%). I use the 2008–2013 period as the baseline because it represents peak fragmentation, and comparing to this period reveals which topics have grown versus consolidated during the reconsolidation era.

I construct citation networks linking papers in the sample, identifying which other sample papers each paper cites and which cite it. This yields a directed network with 4,218 nodes (papers with at least one citation link) and 23,534 edges (citation relationships). I compute citation flows between clusters: for each ordered pair of clusters (A, B), I count citations from papers in A to papers in B, revealing whether clusters are integrated (cross-cluster citation) or siloed (within-cluster citation only).

Coherence and Robustness

I measure the field's semantic coherence as the average pairwise cosine similarity among papers published in each year. Higher values indicate a more unified field; lower values indicate fragmentation. I also measure paradigm concentration using a Herfindahl-Hirschman Index (HHI) computed over macro theme shares, where higher HHI indicates concentration around fewer themes.

I verified robustness to alternative clustering parameters ($k=30$, $k=50$, $k=70$). Results are substantively unchanged across specifications (see Appendix).

Interactive Data Exploration: All papers, clusters, and visualizations can be explored interactively at whatisstrategy.org.

Findings

This section presents four findings.

Overview: The Semantic Landscape

Before presenting the main findings, I provide an overview of the field's structure.

Figure 2 presents a semantic map of strategy research. Each point represents a paper; proximity indicates conceptual similarity. Colors denote macro themes. Papers form distinct clusters, indicating identifiable research themes. Clusters partially overlap, reflecting conceptual connections across themes.

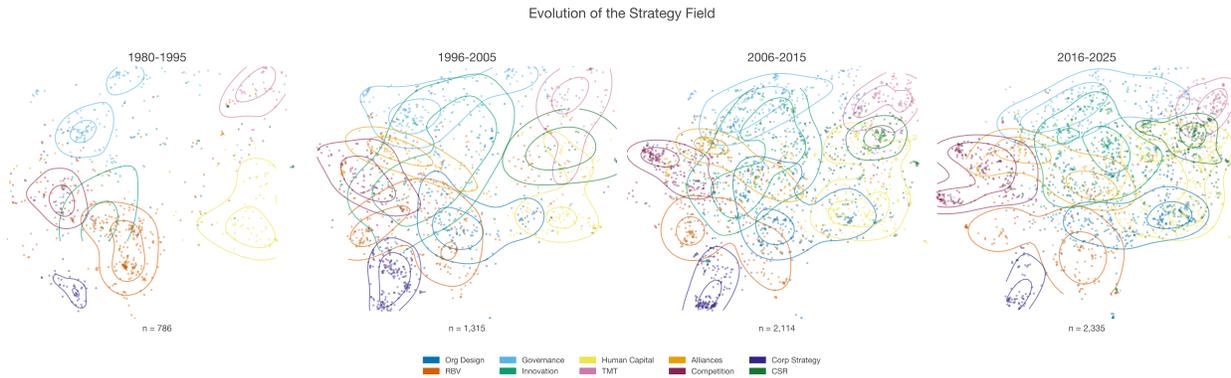
The spatial arrangement reveals meaningful intellectual structure. Three broad regions emerge. The *right side* of the map contains research focused on people in organizations: Corporate Governance (yellow), Strategic Human Capital (purple), and Top Management Teams (dark blue) cluster together, reflecting their shared concern with executives, boards, incentives, and human resources. These themes overlap substantially—a paper on CEO

Figure 2: Semantic Map of Strategy Research (1980–2025).



Note: Each point represents a paper positioned by semantic similarity of its abstract. Colors denote ten macro themes identified through clustering. Proximity indicates conceptual similarity. N = 6,550 papers from five journals. Interactive 2D and 3D versions available at whatisstrategy.org.

Figure 3: **Evolution of the Strategy Field.**



Note: Semantic maps at four time periods show how the field expanded and diversified. Early period (1980–1995) shows concentrated activity; later periods show expansion into new research areas while maintaining core structure. Colors indicate macro themes; contours show cluster density. Paper counts shown below each panel.

compensation, for example, lies at the intersection of governance, human capital, and TMT research.

The *left side* contains research focused on firm-level resources and structures. Organizational Design & Knowledge (blue) and Resource-Based View (orange) intermingle, reflecting their shared emphasis on internal firm characteristics—capabilities, routines, knowledge assets, and organizational architecture. Strategic Alliances (teal) sits adjacent, bridging the internal-resource perspective with inter-organizational relationships.

The *center* of the map is occupied by Competitive Dynamics (green) and Innovation & Technology (light green), which are notably dispersed rather than tightly clustered. This dispersion reflects the integrative nature of these topics: competition and innovation are studied across virtually all strategic contexts, connecting the human-capital-focused right side with the resource-focused left side. A paper on competitive dynamics in technology markets, for example, draws on both innovation research and market positioning research.

Two themes occupy more peripheral positions. Corporate Social Responsibility (dark blue) forms a relatively distinct cluster at the bottom of the map, suggesting that CSR research

has developed somewhat independently from other strategy streams—though as Finding 3 will show, citation patterns indicate continued intellectual exchange despite this spatial separation. Corporate Strategy (brown), encompassing diversification and M&A, is scattered across the map, consistent with corporate scope decisions touching multiple strategic domains.

The partial overlap between clusters is substantively important. Perfect separation would indicate balkanization—research streams that share no conceptual vocabulary. Complete overlap would indicate redundancy—themes that are not meaningfully distinct. The observed pattern—distinct but overlapping clusters—suggests a field with identifiable specialties that nonetheless share intellectual foundations.

Table 3: Distribution of Strategy Research Papers by Theme and Decade

Research Theme	1980s	1990s	2000s	2010s	2020s	Total
Resource-Based View	185	218	226	185	59	873
Corporate Governance & Di...	68	158	240	293	109	868
Strategic Human Capital	33	72	186	382	177	850
Organizational Design & K...	10	22	227	376	191	826
Competitive Dynamics & Po...	34	72	191	264	117	678
Innovation & Technology	14	22	143	299	196	674
Strategic Alliances & Net...	4	37	173	208	58	480
Corporate Risk Management	8	123	152	180	1	464
Top Management Teams	17	60	83	210	76	446
Corporate Social Responsi...	13	12	46	206	114	391
Total	386	796	1667	2603	1098	6550

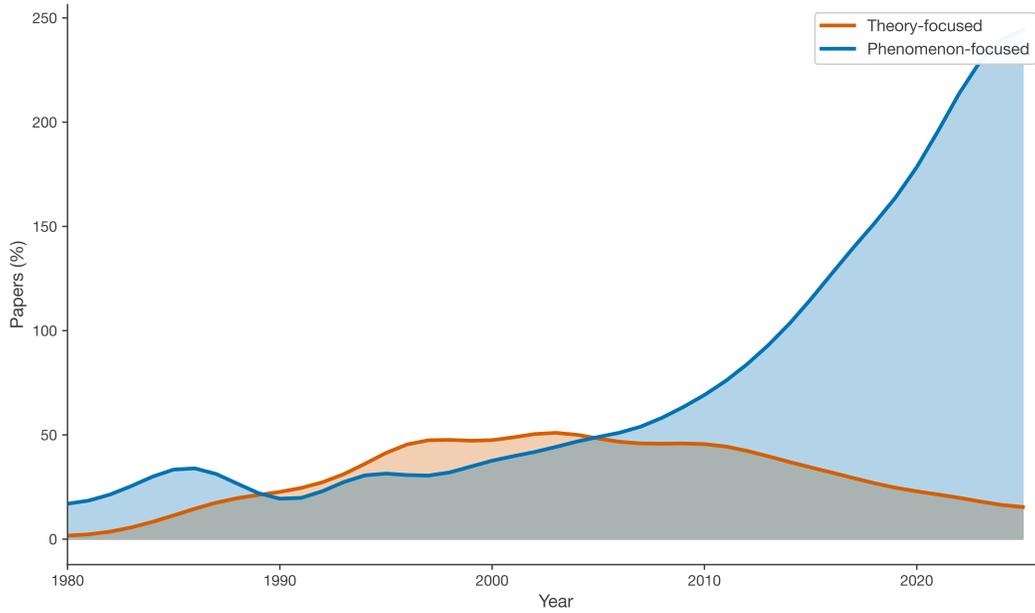
Note: Papers classified using K-means clustering (k=50) on semantic embeddings, aggregated into 10 macro themes.

Table 3 reports the distribution of papers across ten macro themes and five decades.

Finding 1: Theory’s Organizing Role Has Declined

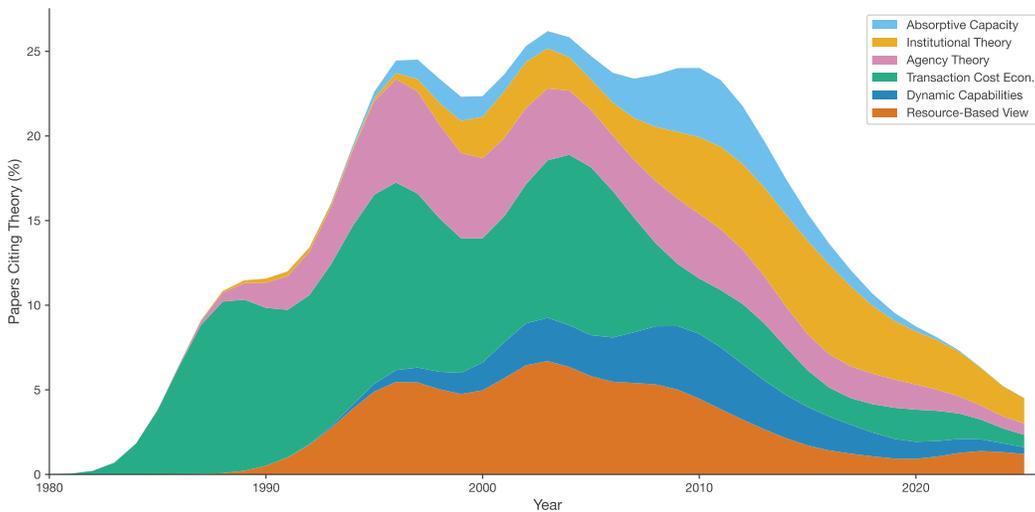
The first finding concerns the changing role of theory in organizing empirical strategy research. Papers explicitly structured around testing or extending the Resource-Based View have declined substantially. Figure 4 traces this trajectory.

Figure 4: Theory vs Phenomenon-Focused Papers.



Note: Papers explicitly organized around testing or extending theories have declined over time, while phenomenon-focused papers have grown. Classification based on LLM analysis of paper abstracts (see Methods).

Figure 5: Citations to Major Theoretical Frameworks.



Note: Citations from phenomenon-focused papers to foundational theory papers rose from near zero (1980s) to a peak (early 2000s), then declined substantially (2020s). Stacked area shows proportion of papers citing each major theoretical framework.

Papers in the Resource-Based View *macro theme* (which includes all papers semantically organized around resources, capabilities, and competitive advantage) peaked at 48% of output in the 1980s and declined to 5% by the 2020s. More narrowly, papers in *theory-focused clusters*—those explicitly organized around testing or extending RBV, Dynamic Capabilities, or TCE—peaked at 13% in the 2000s and declined to 3% by the 2020s (see Methods for operationalization and sensitivity analysis). Dynamic Capabilities peaked at 5% in 2011 before following the same trajectory. Transaction Cost Economics and Agency Theory show similar patterns of decline throughout the period. This pattern extends beyond RBV; all major foundational theories show the same trajectory: explicit theory-testing papers have declined as a share of total output.

This decline in theory-focused papers might suggest that theory has become less important. Citation patterns reinforce rather than contradict this concern. Figure 5 shows citations from phenomenon-focused papers (those in clusters organized around platforms, CSR, human capital, etc.) to foundational theory papers (Barney 1991, Wernerfelt 1984, Teece et al. 1997, etc.). Theory citations rose from near zero in the early 1980s to a peak of approximately 25% in the early 2000s, coinciding with the height of RBV’s influence. However, theory citations have declined since then, falling to approximately 5–10% by the 2020s.

The pattern is clear: theory’s organizing role in empirical research has declined on both dimensions. Whether measured by papers explicitly structured around testing theories or by citations to foundational works, theory organizes less empirical work now than it did two decades ago. This represents a substantial shift in how empirical research is framed. Table 4 summarizes the decade-by-decade pattern.

Note: Citations per year, controlling for paper age.

This finding is robust to alternative specifications. The pattern holds across different keyword lists for identifying theory-focused papers. The pattern is stronger in SMJ (–71% theory decline from peak) than in AMJ (–61%), suggesting it reflects strategy-specific trends rather than general management journal trends

Table 4: Citation Rates by Cluster Type and Decade (Age-Adjusted)

Decade	Theory Clusters	Phenomenon Clusters
1980s	9.0	7.6
1990s	9.1	12.5
2000s	23.5	18.0
2010s	18.1	15.6
2020s	14.6	12.6
Overall	16.7	15.0

(see Appendix). Excluding the top five fastest-growing phenomenon clusters does not eliminate the pattern: theory-focused papers still decline from 13.6% (2000s) to 4.1% (2020s).

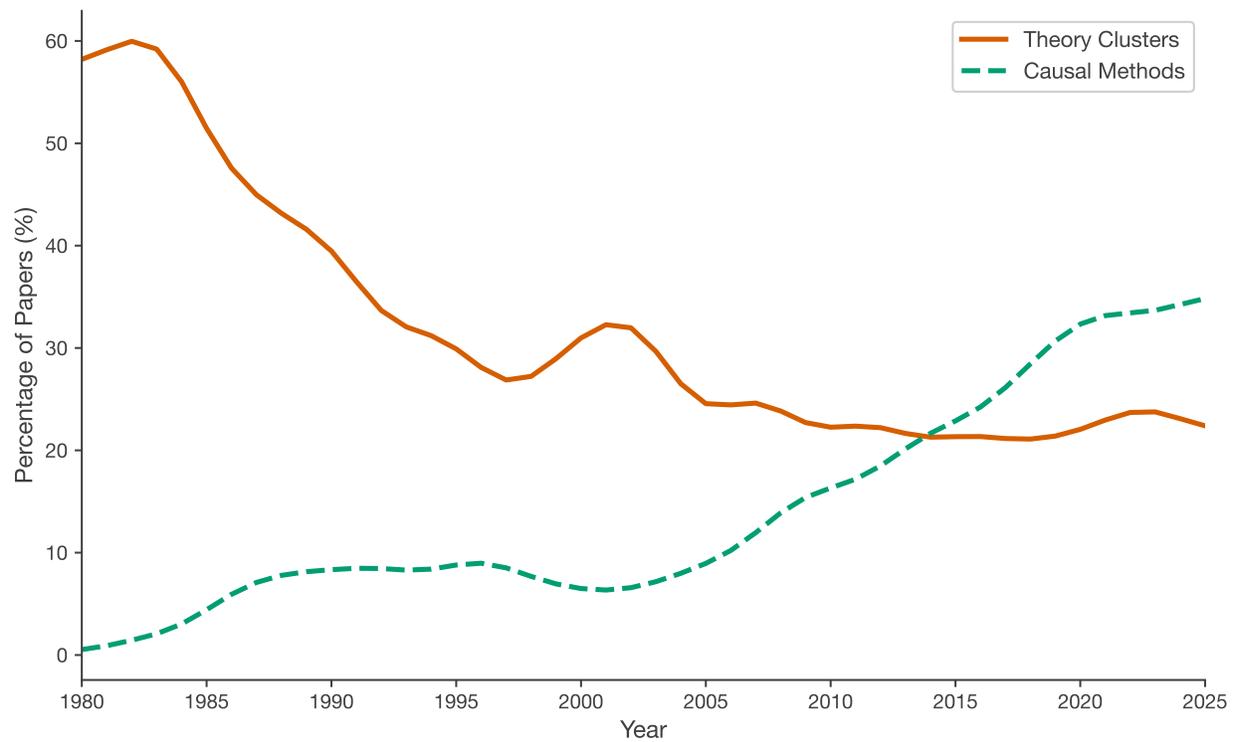
A limitation of this analysis is that citation counts cannot distinguish genuine theoretical engagement from perfunctory or strategic citation. The rise and subsequent fall in theory citations could reflect changing norms around theoretical framing, journal requirements, reviewer expectations, or boilerplate literature review practices rather than genuine intellectual engagement. Distinguishing these interpretations would require qualitative analysis of how citations are used in context—whether theories generate testable predictions, inform variable selection, or merely appear in introductory paragraphs. This paper documents the citation pattern but cannot adjudicate between these interpretations.

Timeline and mechanisms. The timing of theory’s decline is revealing (Figure 6). Theory citation rose through the 1990s and peaked around 2004–2005 at approximately 26% (5-year moving average). The decline began around 2006–2008, falling below 20% by 2014 and below 10% by 2020. Causal methods rose during the same period, with acceleration after 2013.

Shifting incentives. Does the decline in theory engagement affect research quality? Citation analysis reveals a shifting incentive structure (Figure 7). Theory-cluster papers historically received more citations per year than phenomenon-cluster papers. In the 2000s and 2010s, this premium was statistically significant (+5.6 cites/year, $p = 0.004$ in the 2000s; +2.5, $p = 0.017$ in the 2010s). By the 2020s, the premium has shrunk to +2.0 cites/year and is no longer statistically significant ($p = 0.15$). The market no longer detectably rewards theoretical engagement.

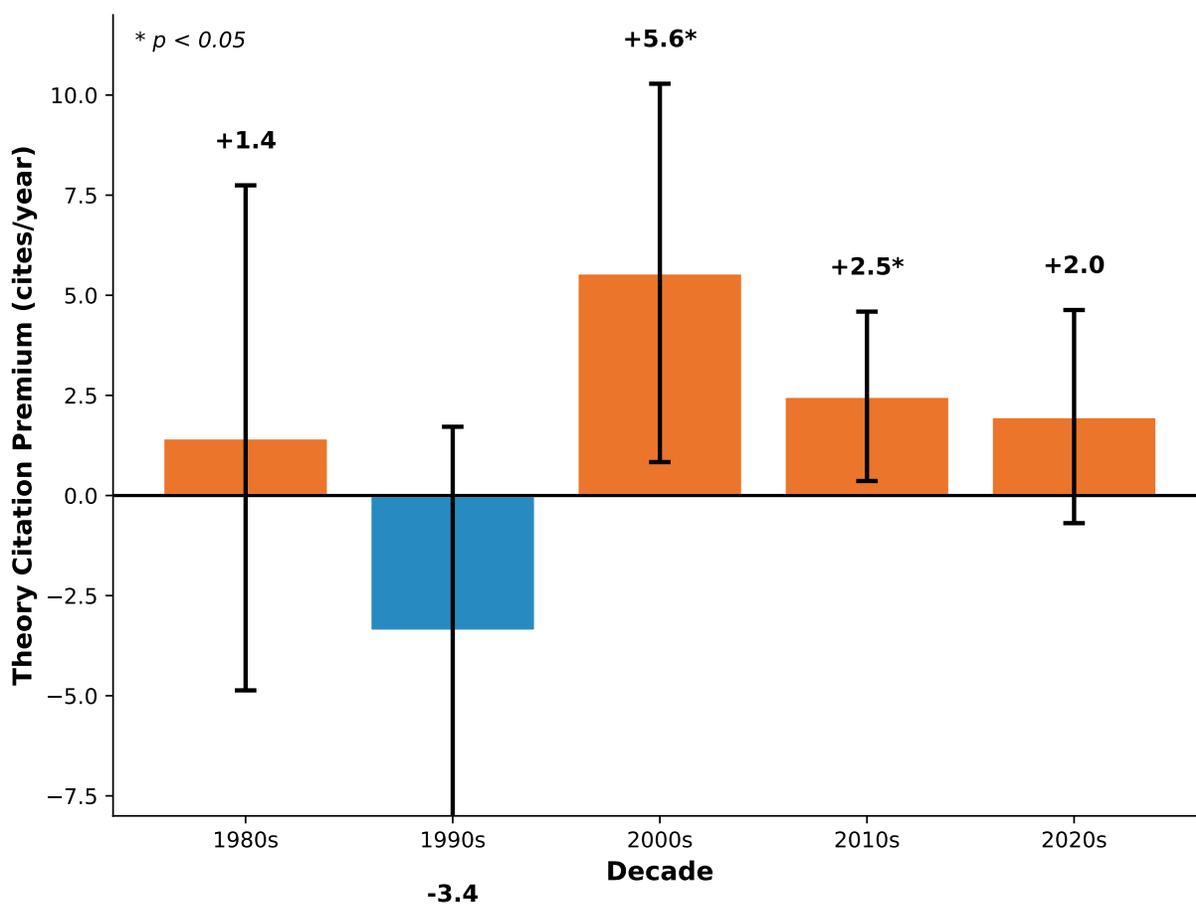
Construct clarification. The “theory-focused” versus “phenomenon-focused” distinction requires clarification. Papers classified as theory-focused are those *explicitly organized around* testing or extending foundational frameworks—not all papers that engage with theory. Many phenomenon-focused papers cite

Figure 6: Theory Citation and Causal Methods Over Time.



Note: Theory citation (orange) peaked around 2005 and declined thereafter. Causal methods (blue dashed) rose gradually through the 2000s, accelerating after 2013. Both trends shifted in the mid-2000s.

Figure 7: Theory Citation Premium by Decade.



Note: The citation premium for theory-cluster papers was statistically significant in the 2000s (+5.6 cites/year, $p = 0.004$) and 2010s (+2.5, $p = 0.017$) but has shrunk and become insignificant by the 2020s (+2.0, $p = 0.15$). Error bars show 95% confidence intervals; * indicates $p < 0.05$.

and apply theory; the classification captures whether the paper’s *primary organizing principle* is theoretical versus empirical. This operationalization captures changes in how research is *framed* and *organized*, which may reflect shifts in rhetorical conventions as much as genuine intellectual change. The 75% LLM-human agreement rate (Appendix) reflects inherent ambiguity in the construct.

Finding 2: The Field Has Partially Reconsolidated Around Phenomena

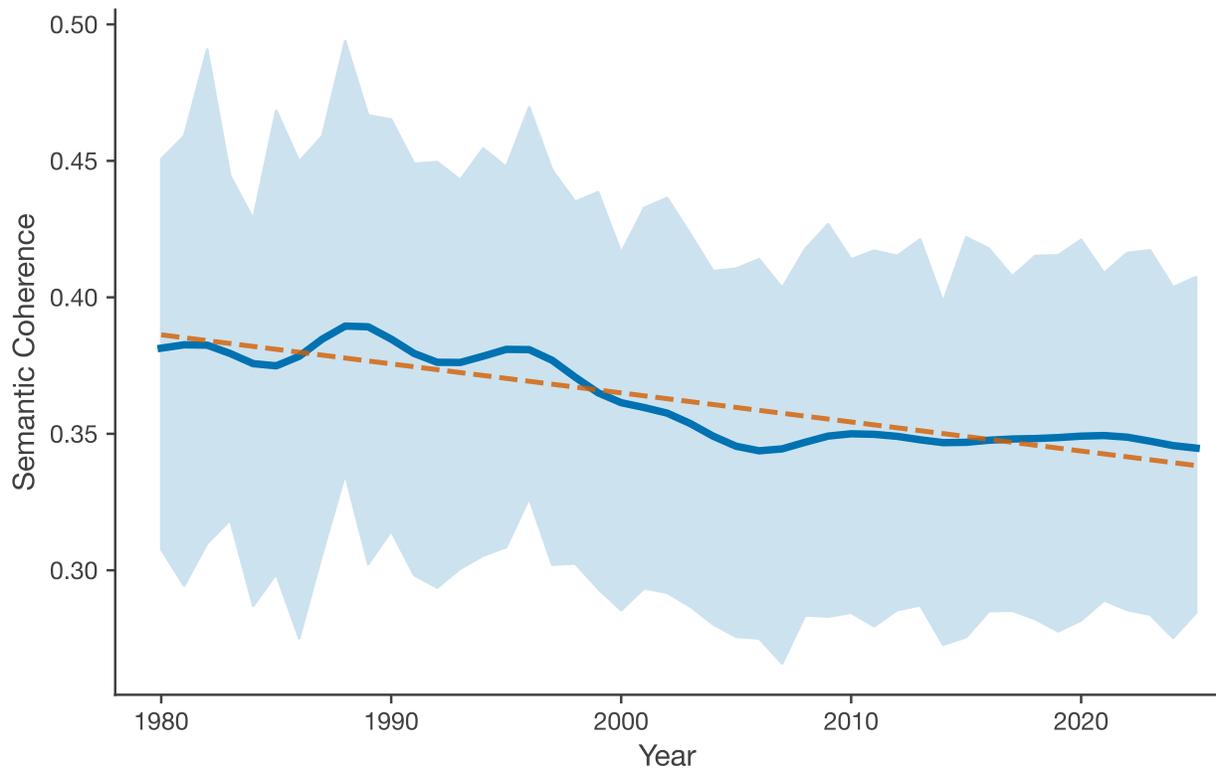
The second finding concerns the field’s structural evolution. Following the decline of theory-focused research, the field initially dispersed. Multiple indicators show this pattern. Semantic coherence, measured as average pairwise similarity among papers published each year, fell from 0.382 in the 1980s to 0.348 in the 2010s, a 9% decline (see Figure 8). The Herfindahl-Hirschman Index of theme shares declined from 0.081 in 1990 to 0.028 in 2020, and the number of effective research clusters grew from 12.4 to 35.5 (Table 8). No single paradigm replaced the foundational theories. Instead, research dispersed across multiple phenomenon-based streams, including Organizational Design & Knowledge (22x growth), Strategic Alliances (17x growth), Innovation & Technology (15x growth), Corporate Social Responsibility (9x growth), and Strategic Human Capital (6x growth).

However, dispersion was not the endpoint. Since 2013, several indicators suggest the field has begun reconsolidating. The number of effective research clusters declined from approximately 44 at peak to 34 by the 2020s (Figure 9). The share of papers in the top five clusters grew from 17% (2010-2015) to 24% (2020-2025). These clusters are organized around phenomena rather than theories: platform ecosystems, strategic entrepreneurship, nonmarket strategy, strategic human capital, and innovation ecosystems. Figure 10 shows this shift in detail.

A critical methodological note: raw effective clusters correlate strongly with annual paper counts ($r = 0.87$), raising the concern that apparent dispersion merely reflects sample growth. Figure 9 addresses this by showing both raw counts and sample-size-adjusted values (regression residuals). The adjusted pattern is striking: after controlling for sample size, the 2020s show the *lowest* concentration of any decade—more concentrated than expected given the number of papers published. The dispersion-then-reconsolidation pattern persists after adjustment, but the interpretation shifts: the field has not merely returned toward its earlier structure; it has become more concentrated than a field of its current size would typically be. Whether this reconsolidation represents genuine intellectual integration or convergence on tractable topics (driven by data availability, funding patterns, or methodological opportunities) remains an open question.

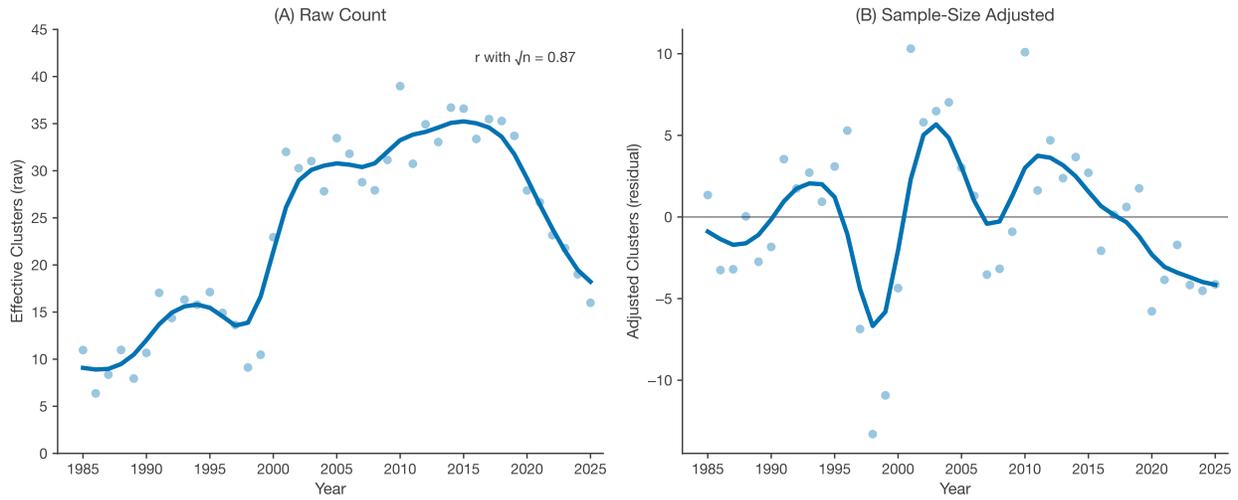
Table 5 provides additional detail on shifting research priorities. Tables 6 and 7 provide more detailed

Figure 8: **Field Coherence Over Time.**



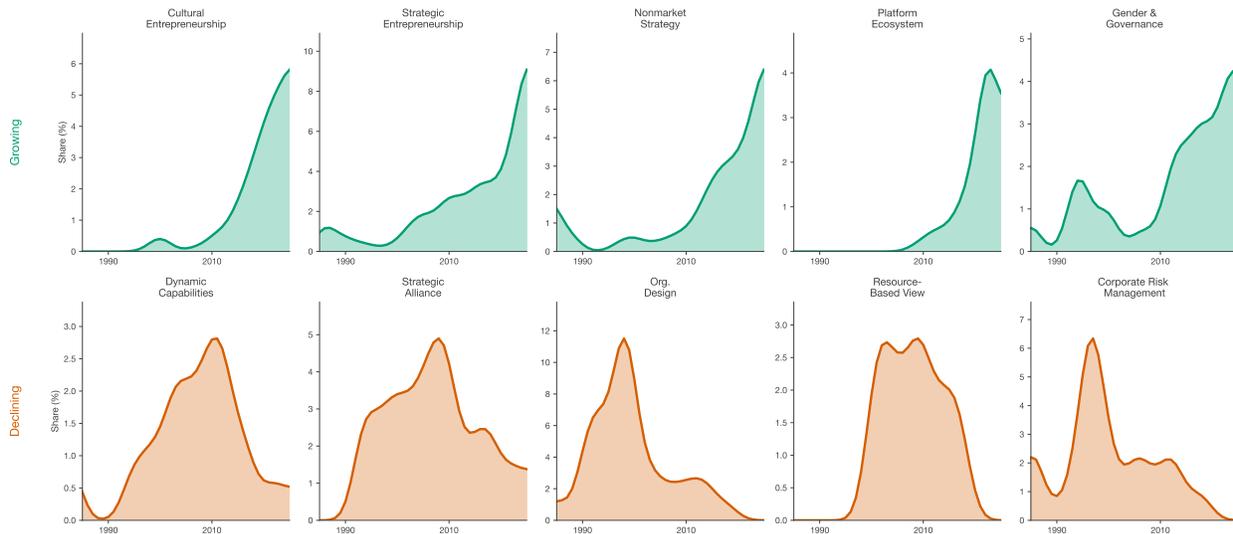
Note: Semantic coherence, measured as mean pairwise cosine similarity among papers published each year, shows a modest decline. Shaded region indicates one standard deviation; dashed line shows linear trend.

Figure 9: Field Structure: Dispersion Followed by Reconsolidation.



Note: (A) Raw effective clusters (inverse HHI) correlate strongly with annual paper counts ($r = 0.87$). (B) After regression adjustment for sample size, the pattern persists: the field was more concentrated than expected in the 1980s–1990s, more fragmented than expected in the 2000s–2010s, and has reconsolidated to below-expected concentration in the 2020s. Adjustment via regression: $\text{eff_clusters} = -3.9 + 1.94\sqrt{n}$; residuals shown in Panel B.

Figure 10: Theme Evolution: Rising and Declining Research Areas.



Note: (A) Themes that grew from 2008–2013 to 2020–2025 are primarily organized around phenomena (entrepreneurship, platforms, nonmarket strategy). (B) Themes that declined were primarily organized around theories (dynamic capabilities, transaction cost economics, strategic alliances).

analysis of the top rising and declining research topics. The rising topics show dramatic growth: Cultural Entrepreneurship Strategy grew 600%, Platform Ecosystem Strategy grew 567%, and Nonmarket Strategy Dynamics grew 308%. In contrast, several theory-focused topics show complete or near-complete decline. Table 8 summarizes the field’s concentration metrics across three key periods.

Table 5: Thematic Shifts: Early (1980-1990s) vs Late (2010-2020s)

Theme	Early %	Late %	Change
Organizational Design & Knowle	2.7	15.3	+12.6pp
Innovation & Technology	3.0	13.4	+10.3pp
Corporate Social Responsibilit	2.1	8.6	+6.5pp
Strategic Human Capital	8.9	15.1	+6.2pp
Strategic Alliances & Networks	3.5	7.2	+3.7pp
Competitive Dynamics & Positio	9.0	10.3	+1.3pp
Top Management Teams	6.5	7.7	+1.2pp
Corporate Risk Management	11.1	4.9	-6.2pp
Corporate Governance & Diversi	19.1	10.9	-8.3pp
Resource-Based View	34.1	6.6	-27.5pp

Table 6: Rising Research Topics

Topic	Papers
Human Capital Strategy	211
CEO Influence & Strategy	191
Strategic Entrepreneurship Dynamics	185
Board Dynamics & Governance	184
Organizational Adaptation Strategie	182
Team Dynamics & Performance	146
Stakeholder Relationship Dynamics	139
Corporate Social Responsibility	133
Reputation Management Strategy	130
Nonmarket Strategy Dynamics	128

This finding is robust to alternative specifications. The reconsolidation pattern appears at k=30 and k=50 clustering parameters, though it is flat at k=70. Different era cutoffs (2008, 2010, 2012, 2015) all show reconsolidation beginning in the early 2010s.

Finding 3: The Field Is Not Balkanized

The third finding addresses the concern about intellectual silos. Do research streams cite each other? I analyzed 23,534 citations among papers in the sample, tracking whether citations cross cluster boundaries

Table 7: Declining Research Topics

Topic	Papers
Competitive Dynamics Analysis	198
Strategy Process & Practice	195
Organizational Design & Complexity	189
Corporate Diversification Strategie	154
Strategic Planning Processes	154

Table 8: Field Concentration Over Time

Year	HHI	Effective Clusters
1990	0.081	12.4
2000	0.035	28.6
2010	0.028	35.9
2020	0.028	35.5

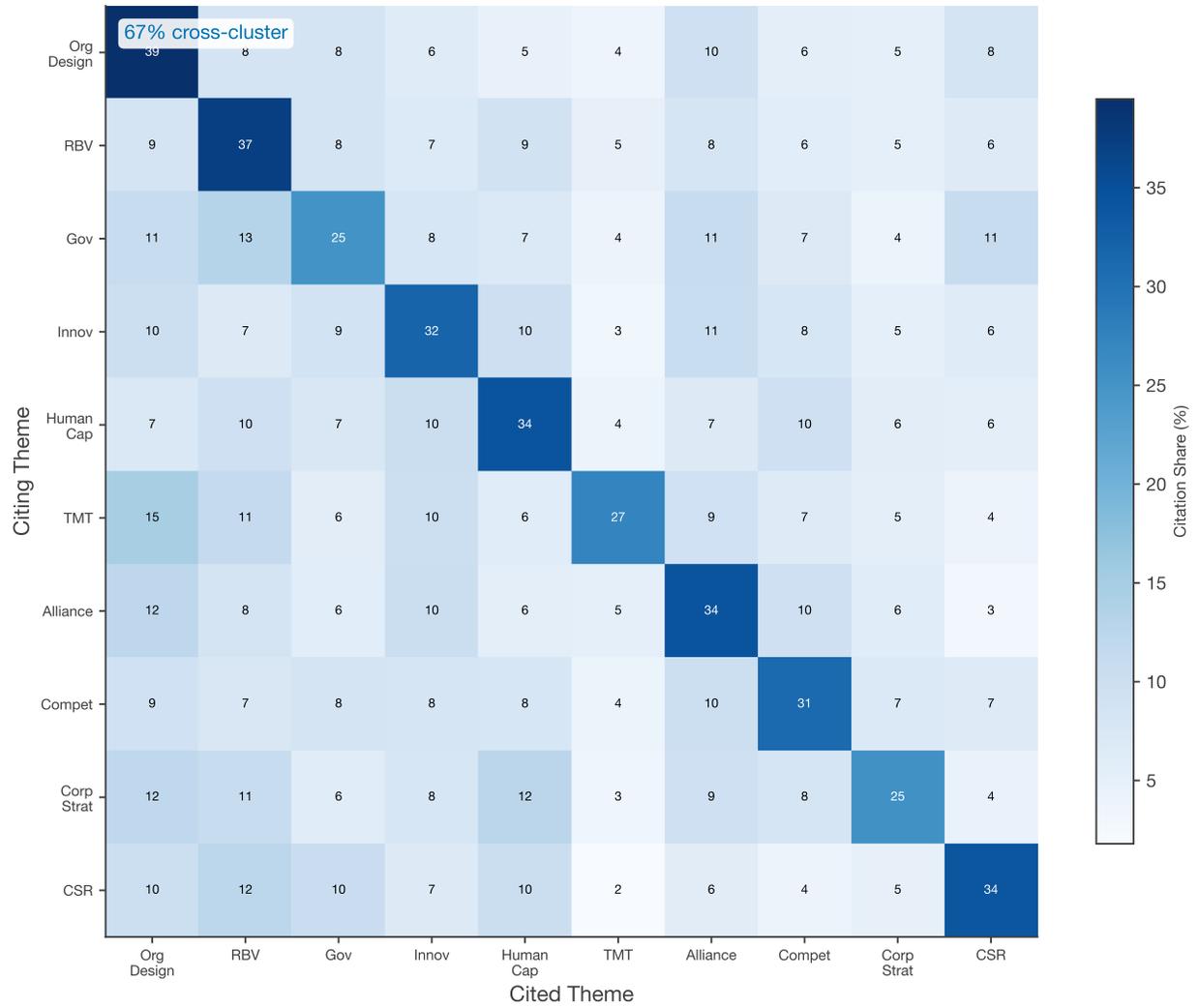
(Figure 11). The data show that 67.2% of citations cross cluster boundaries, while only 32.8% stay within the same theme. Scholars cite outside their research stream far more than within it.

To benchmark this finding, I computed the expected cross-citation rate under random citation. If papers cited other papers in proportion to cluster sizes (i.e., no thematic preferences), the expected cross-cluster citation rate would be 89.2% (one minus the Herfindahl-Hirschman Index of cluster shares). The actual rate of 67.2% is 22 percentage points below this random baseline, indicating that papers do cite within their cluster more than random chance would predict. However, random citation is an upper bound rather than a target: some within-cluster citation is expected and appropriate as researchers build on closely related work. A field where everyone cited randomly would lack the depth of engagement that comes from sustained intellectual communities. The relevant question is whether cross-citation is high enough to prevent balkanization, not whether it reaches the random baseline. At 67%, the majority of citations cross cluster boundaries, suggesting substantial integration.

Importantly, this rate has remained stable over 45 years: 64.1% in the 1980s, 64.5% in the 1990s, 66.4% in the 2000s, 68.4% in the 2010s, and 66.4% in the 2020s. The field has not become more siloed over time.

Integration varies by theme, but all themes show substantial cross-citation. Corporate Governance has the highest integration (0.937 entropy score) while Organizational Design & Knowledge has the lowest (0.857 entropy score). Even the most “insular” cluster has 86% of the maximum possible integration score. Some papers connect disparate themes. I identify “bridge papers” as those with high similarity to multiple clusters

Figure 11: Cross-Cluster Citation Flows.



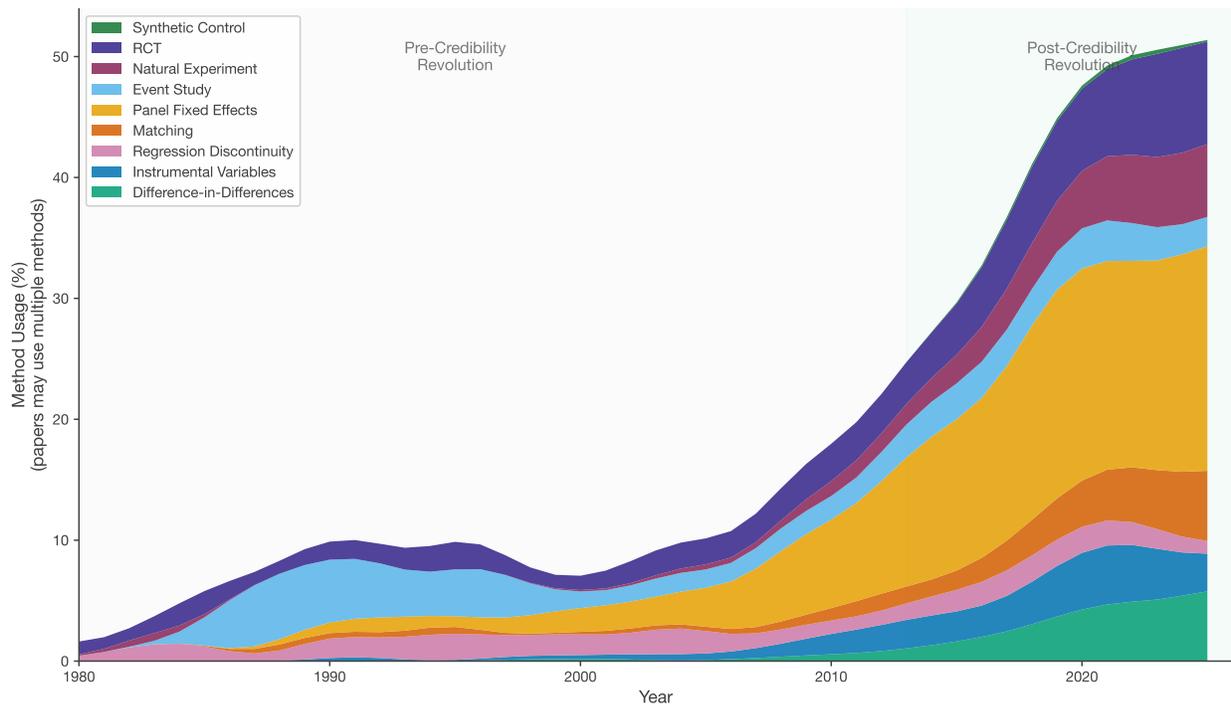
Note: Each cell shows the percentage of citations from papers in the row theme to papers in the column theme. Diagonal values (bold) indicate within-cluster citation. 67.2% of all citations cross cluster boundaries, indicating continued intellectual exchange across research streams.

(top 10% by average similarity to non-home cluster centroids). Bridge papers show higher median citations (141 vs. 122, a 15% premium), suggesting that boundary-spanning work has impact at the typical paper level. However, mean citations are lower for bridge papers (352 vs. 472), indicating that the highest-cited papers tend to be cluster-defining seminal works rather than boundary-spanning papers. The strongest citation flows connect Organizational Design to Strategic Alliances (339 citations), Corporate Governance to RBV (285 citations), and Corporate Strategy to Strategic Human Capital (274 citations).

Finding 4: Causal Methods Have Not Distorted the Research Agenda

The fourth finding addresses the lamppost problem: has the credibility revolution pushed research toward peripheral topics or less important questions? Among empirical quantitative strategy papers, the share using causal identification methods (difference-in-differences, instrumental variables, regression discontinuity, matching) has increased substantially since the 2000s (Figure 12). Table 9 shows the broader evolution of research methodologies by decade.

Figure 12: Rise of Causal Identification Methods.



Note: The percentage of empirical quantitative strategy papers using difference-in-differences, instrumental variables, regression discontinuity, and matching methods has increased substantially since 2008. This reflects the “credibility revolution” in empirical social science.

Table 10 provides a detailed breakdown of specific causal identification methods, comparing adoption rates

Table 9: Methodology Evolution by Decade

Decade	Causal Methods %
1980s	5.2%
1990s	8.0%
2000s	9.7%
2010s	23.1%
2020s	34.2%

between Era 1 (1980–2013) and Era 2 (2014–2025). Panel fixed effects showed the largest increase (+11.5 percentage points), followed by randomized experiments (+4.0pp) and natural experiments (+3.5pp). Overall, the share of papers using any causal identification method nearly tripled from 9.8% to 28.6%.

Table 10: Rise of Causal Identification Methods

Method	Era 1 (1980–2013)	Era 2 (2014–2024)	Change
Panel Fixed Effects	5.3%	18.1%	+12.8pp
Natural Experiment	0.5%	4.6%	+4.1pp
Instrumental Variables	0.9%	3.9%	+3.0pp
Matching Methods	0.9%	3.6%	+2.8pp
Regression Discontinuity	1.7%	2.3%	+0.6pp
Event Study	3.0%	3.5%	+0.6pp
Randomized Experiment	0.0%	0.0%	+0.0pp
Difference-in-Differences	0.0%	0.0%	+0.0pp
Any Causal Method	10.6%	28.7%	+18.1pp

Note: Era 1 includes 3,646

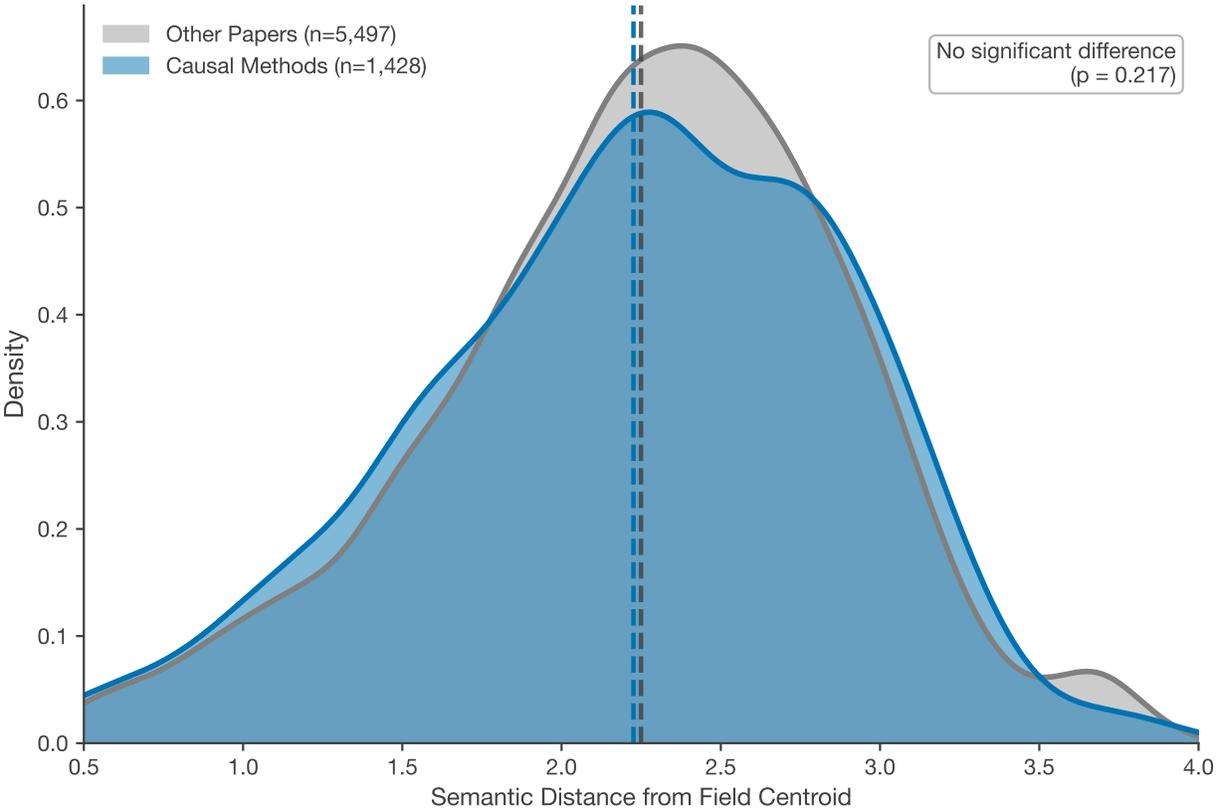
papers; Era 2 includes 2,904 papers.

To test whether causal methods papers concentrate in peripheral areas, I computed each paper’s semantic distance from the field centroid (the average location of all papers in embedding space). Causal methods papers have a mean distance of 2.23, while other papers have a mean distance of 2.25, a difference that is not statistically significant ($p = 0.22$). Papers using causal methods are distributed throughout the field’s semantic space, not concentrated in peripheral areas (Figure 13).

Figure 14 visualizes this result directly by plotting all papers in semantic space. If the lamppost critique were correct, the orange points (causal methods papers) would cluster at the edges, away from the field center. Instead, causal methods papers are distributed throughout the semantic space, interspersed with non-causal papers at all distances from the center.

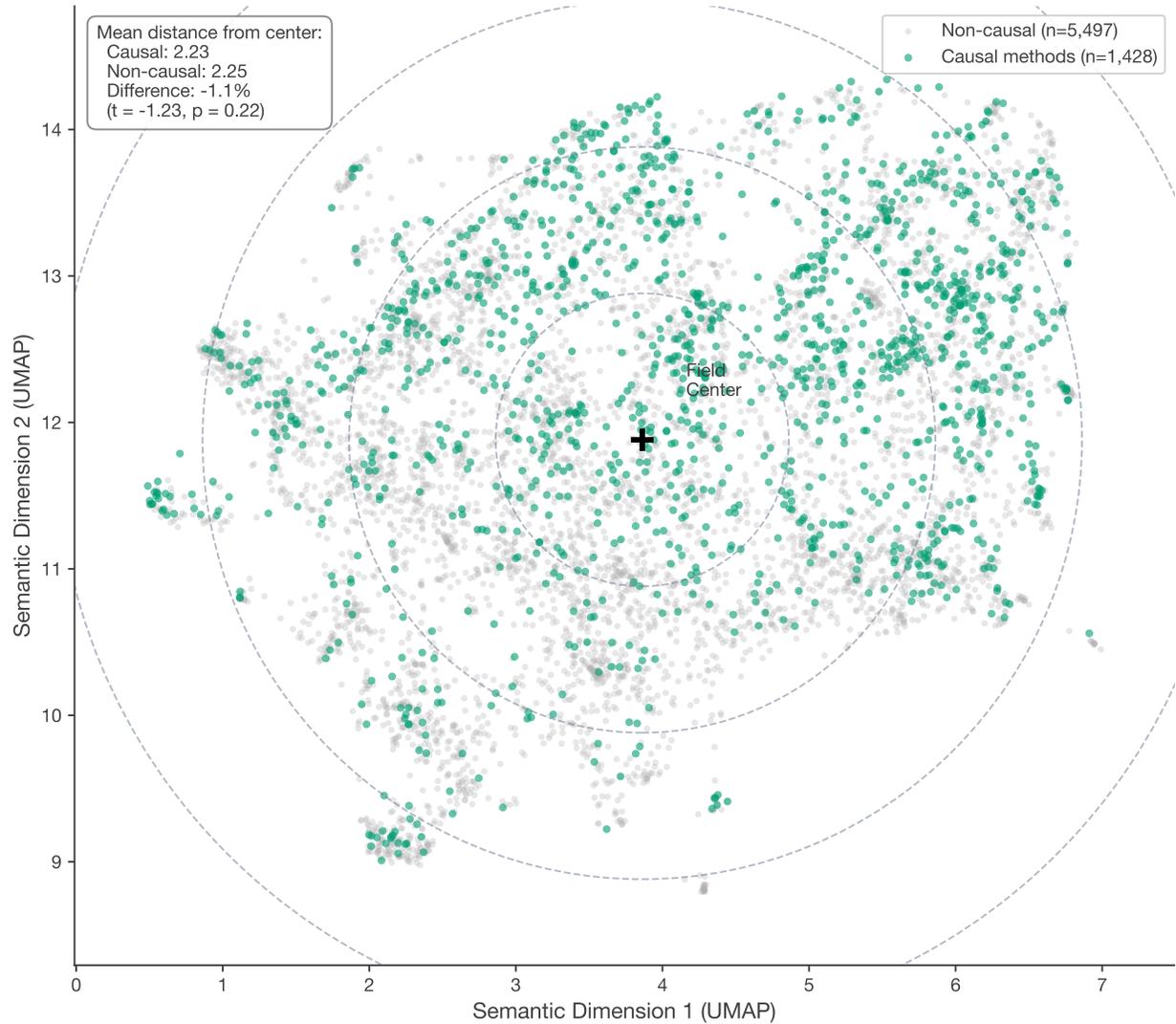
Adoption of causal methods varies by theme. Table 11 shows adoption rates by macro theme for papers

Figure 13: Causal Methods Papers Are Not Peripheral (“Lamppost Test”).



Note: Distribution of semantic distance from the field centroid for papers using causal identification methods versus other papers. The means are nearly identical (2.23 vs. 2.25, $p = 0.22$), indicating that causal methods papers are distributed throughout the field rather than concentrated in peripheral areas.

Figure 14: Causal Methods Papers in Semantic Space.



Note: Each point represents a paper positioned by semantic similarity (UMAP projection of embeddings). Causal methods papers (green, n=1,428) are distributed throughout the field, not concentrated at the periphery. The field center (crosshairs) is the centroid of all papers. Dashed circles show distance rings. Mean distance from center: causal = 2.23, non-causal = 2.25 (difference not significant, $p = 0.22$).

published since 2014. CSR shows the highest adoption rate (42.5%), followed by Corporate Governance (29.1%) and Top Management Teams (28.0%), while theory-focused themes like Resource-Based View (4.7%) show the lowest adoption. This variation likely reflects data availability and identification opportunities rather than the importance of research questions. CSR research can exploit regulatory shocks and ESG index changes; RBV research, focused on intangible resources, faces greater identification challenges.

Table 11: Causal Methods by Research Theme

Theme	Causal Methods %
Corporate Social Responsibility	42.5%
Corporate Governance & Diversification	29.1%
Top Management Teams	28.0%
Innovation & Technology	25.8%
Strategic Alliances & Networks	19.0%
Strategic Human Capital	15.1%
Competitive Dynamics & Positioning	13.1%
Organizational Design & Knowledge	13.0%
Corporate Risk Management	10.1%
Resource-Based View	4.7%

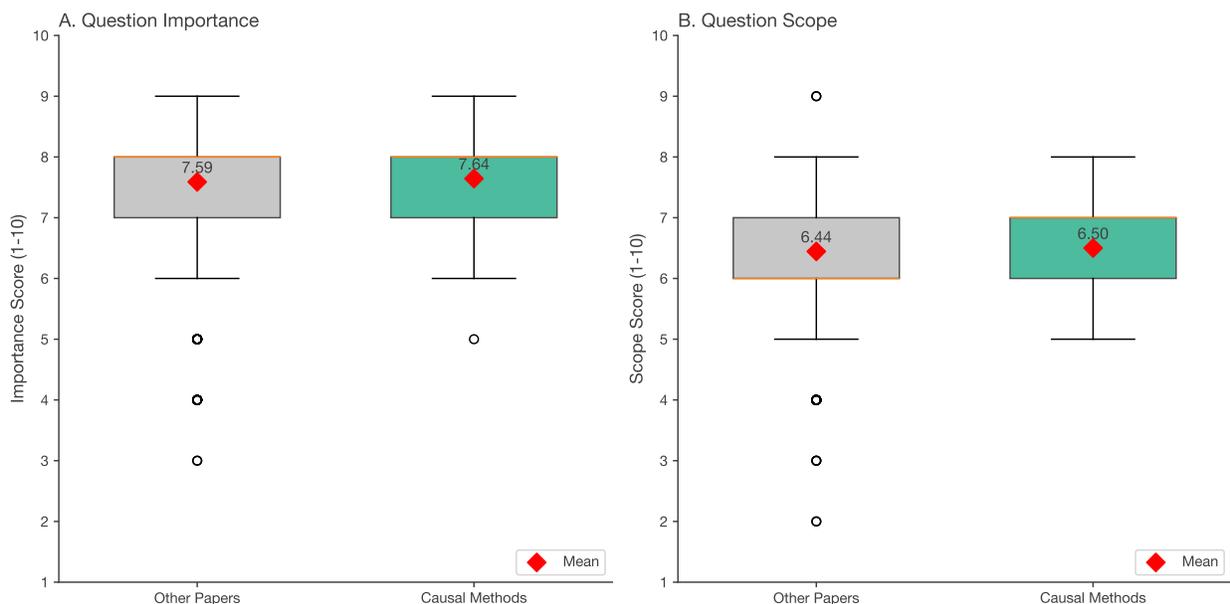
This finding addresses a specific version of the lamppost concern: whether causal methods papers cluster in topically peripheral areas. The answer appears to be no—causal methods papers are distributed throughout the field’s semantic space. However, the lamppost problem may operate through mechanisms this test does not capture. Causal methods papers might ask *narrower* questions within the same topics, focusing on settings where identification is tractable rather than on the most important questions. They might emphasize easily measured dependent variables (stock returns, patent counts, firm survival) over harder-to-measure constructs (capabilities, culture, tacit knowledge). Semantic distance from the field centroid would not detect these patterns.

Question importance and scope. To address this concern more directly, I used GPT-4o-mini to score each paper’s research question on two dimensions: *importance* (how important is answering this question for advancing our understanding of strategy?) and *scope* (how broad or narrow is the research question?), each on a 1–10 scale. The model was prompted to evaluate importance on a scale where 10 indicates “a fundamental question that could reshape how we think about strategy” and 1–3 indicates “narrow or incremental question with limited broader relevance.” For scope, 10 indicates “very broad, addresses fundamental mechanisms or general principles” and 1–3 indicates “very narrow, specific to a particular context, time, or setting.” Each paper was scored independently based on its title and abstract, with the model providing a brief rationale for each score. If the lamppost problem operates as critics fear, causal

methods papers should score lower on both dimensions—asking narrower, less important questions in exchange for clean identification.

The results do not support this concern. Papers using causal identification methods score *higher* on both importance (7.67 vs. 7.58, $p < 0.001$) and scope (6.51 vs. 6.44, $p < 0.001$) than other papers (Figure 15). The effect sizes are small (Cohen’s $d = 0.16$ for importance, 0.11 for scope)—the 0.09-point difference on a 10-point scale has limited practical significance. However, the key finding is not the magnitude but the *direction*: if the lamppost critique were correct, we would expect causal methods papers to score *lower* on importance and scope, not higher. The data show no evidence that causal methods push researchers toward narrower or less important questions; if anything, the small positive difference suggests the opposite.

Figure 15: **Question Importance and Scope by Method Type.**



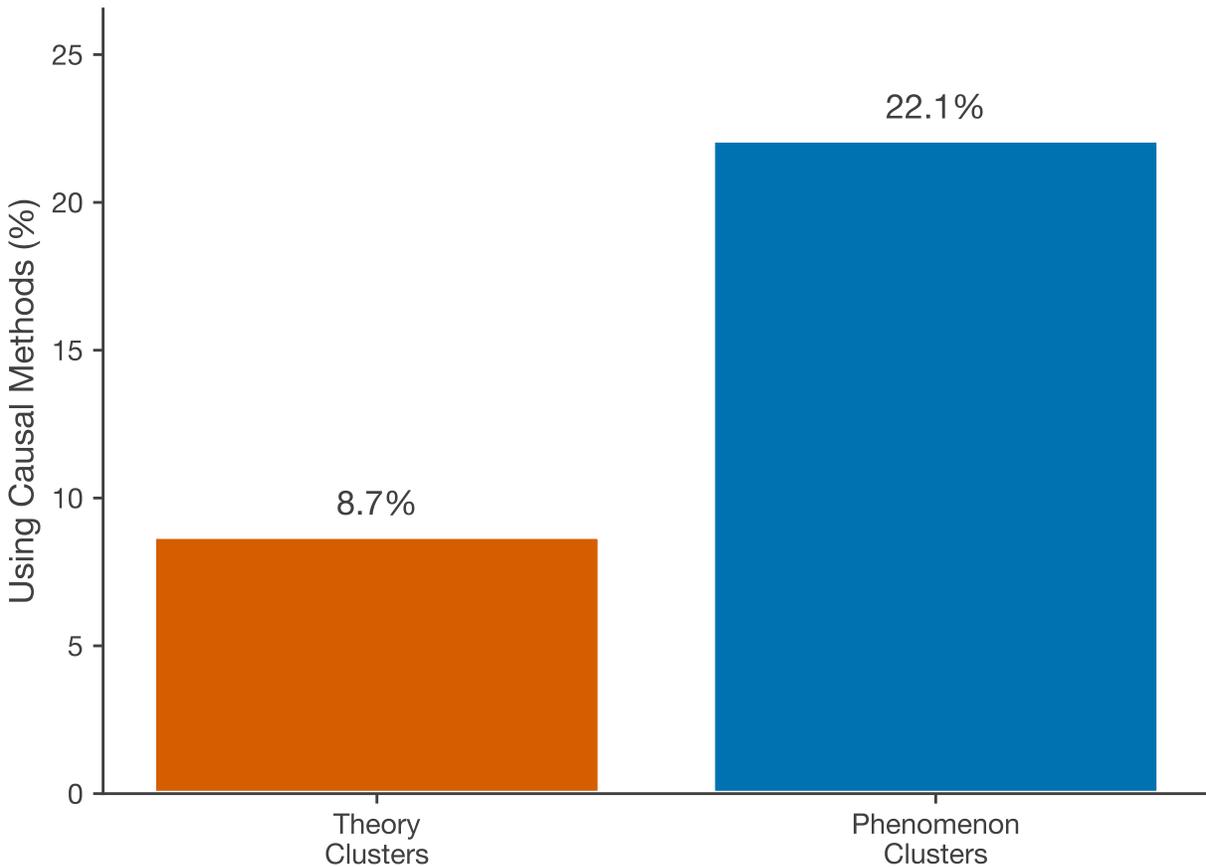
Note: **(A)** Papers using causal identification methods score higher on question importance (7.67 vs. 7.58, $p < 0.001$). **(B)** They also score higher on question scope (6.51 vs. 6.44, $p < 0.001$). Both patterns are contrary to lamppost concern predictions. Box plots show distributions; red diamonds indicate means. $N = 1,253$ causal methods papers; 5,672 non-causal papers. Scores assigned by GPT-4o-mini based on title and abstract.

This finding complements the semantic distance analysis: causal methods papers are neither topically peripheral nor substantively narrow. The credibility revolution appears to have equipped researchers with better tools for answering important questions rather than pushing them toward trivial questions with tractable designs.

Causal methods and theory engagement. While causal methods papers are not topically peripheral,

they are concentrated in phenomenon-focused rather than theory-focused research (Figure 16). Papers using causal identification methods are significantly less likely to appear in theory-driven research clusters: only 9.7% of causal-method papers fall in theory-driven clusters, compared to 17.2% of non-causal papers ($\chi^2 = 51.9$, $p < 0.001$). This 7.5 percentage point gap holds across specific methods: difference-in-differences papers show only 7.0% theory-cluster membership, instrumental variables 7.8%, and natural experiments 7.9%. Even panel fixed effects—the most common causal method—shows only 11.0% theory-cluster membership.

Figure 16: **Causal Methods and Theory Engagement.**



Note: (A) Papers using causal identification methods are significantly less likely to be in theory-driven research clusters (9.7% vs. 17.2%, $p < 0.001$). *(B)* This pattern holds across all specific causal methods, with diff-in-diff showing the lowest theory-cluster rate (7.0%) and panel fixed effects the highest among causal methods (11.0%). The dashed line shows the non-causal reference rate.

Why are causal methods concentrated in phenomenon-focused research? Phenomena provide natural variation—regulatory shocks, market entries, policy changes—that quasi-experimental methods can exploit. Theory-testing, by contrast, rarely offers such clean identification opportunities. You cannot randomize

dynamic capabilities.

Summary. Findings 3 and 4 address concerns about the field’s intellectual health. The field is not balkanized: cross-cluster citation has remained stable at 67% since the 1980s, and researchers studying platforms cite governance research, CSR scholars cite RBV papers, and human capital research draws on organizational design. The credibility revolution has not distorted the research agenda: causal methods papers are distributed throughout the field rather than concentrated peripherally, and they score slightly higher—not lower—on question importance and scope. The field’s structure has changed, but its intellectual integration has not, and methodological advances have not pushed research toward trivial questions.

Discussion

Alternative interpretations. Each finding admits alternative interpretations. The decline in theory’s organizing role, for instance, could reflect several mechanisms beyond genuine intellectual change: (1) editorial policy shifts at leading journals that favored empirical novelty over theoretical contribution; (2) changes in tenure and promotion incentives that reward citation counts over theoretical impact; (3) data availability driving topic choice as large datasets on platforms, CSR, and governance became accessible; (4) methodological fashion favoring causal identification over theoretical integration; or (5) the maturation of foundational theories making novel contributions increasingly difficult. Importantly, these patterns might reflect *healthy field maturation* rather than problematic decline: foundational theories may have become so thoroughly absorbed into the field’s shared assumptions that explicit citation is no longer necessary, much as economists rarely cite Adam Smith. The shift from theory-organizing to phenomenon-organizing could represent adaptation to new empirical realities rather than intellectual decline. These alternatives are not mutually exclusive and may reinforce each other. Appendix systematically maps findings to alternatives and identifies possible tests that future research could pursue.

Theoretical Implications

These findings speak to fundamental questions about how academic fields evolve and what constitutes productive intellectual organization. I discuss the findings’ implications for debates about paradigm development in applied fields, the Pfeffer-Van Maanen controversy over pluralism, the role of phenomenon-based research, and what incentives now shape the field.

Beyond Kuhnian paradigms. One lens for interpreting strategy’s trajectory is Kuhn’s (1962) model of scientific revolutions, which describes fields moving through pre-paradigm phases, periods of normal science

under dominant paradigms, and revolutionary transitions to new paradigms. Some elements of this model appear in strategy: foundational theories (RBV, TCE, Dynamic Capabilities) rose to prominence and declined, suggesting paradigmatic turnover. However, strategy's evolution differs from the Kuhnian model in important respects. Strategy has not experienced revolutionary replacement of one paradigm by another—no “anomalies” accumulated to trigger paradigm crisis, no “crisis” period of intensified debate preceded replacement, and no new dominant theory has displaced the foundational frameworks.

Abbott's (2001) *Chaos of Disciplines* offers a more apt model. Abbott argues that social science fields do not progress through Kuhnian stages but instead exhibit cycles of fractal differentiation: problems subdivide, specialties emerge, and fields periodically reconsolidate around new organizing concepts before fragmenting again. The pattern documented here—paradigm dispersion followed by reconsolidation around phenomena rather than new theories—fits Abbott's cyclical model better than Kuhn's revolutionary model. Strategy appears to have moved from theoretical organization (RBV, TCE) to phenomenological organization (platforms, entrepreneurship, CSR), which Abbott would recognize as a typical disciplinary cycle rather than a sign of pathology.

The findings document remarkably similar trajectories across multiple foundational theories: RBV, Transaction Cost Economics, Dynamic Capabilities, and Agency Theory each rose to prominence and then declined as organizing frameworks for new research. Each provided answers to strategy's central question about why firms differ in performance, and each accumulated critiques that may have contributed to its decline. Rather than paradigm succession, strategy has experienced paradigm dispersion followed by partial reconsolidation around phenomena—a pattern more consistent with Abbott's cyclical model than with Kuhn's revolutionary model.

The Pfeffer-Van Maanen debate revisited. The findings also bear on the contentious debate between Pfeffer (1993) and Van Maanen (1995) about whether paradigm proliferation helps or harms a field. Pfeffer argued that organizational science suffered from excessive fragmentation and that fields with greater paradigmatic consensus—like economics—accumulated knowledge more effectively. Van Maanen countered that Pfeffer's vision of a unified field would produce intellectual homogeneity and stifle creativity.

The evidence presented here offers partial support to both positions. Consistent with Pfeffer's concern, strategy research did fragment substantially from the 1980s to the 2010s: effective clusters quadrupled, semantic coherence declined, and no new paradigm emerged to replace the declining foundational theories. Yet the fragmentation has not produced the isolation Pfeffer feared: 67% of citations still cross cluster

boundaries, a rate that has been stable since the 1980s. And consistent with Van Maanen’s defense of pluralism, the field’s diversification coincided with its expansion into important new domains—platform ecosystems, corporate social responsibility, strategic entrepreneurship—that existing theories did not adequately address. The evidence suggests that strategy has achieved a middle path: thematic diversity without intellectual balkanization.

Phenomenon-based research and theoretical progress. Von Krogh, Rossi-Lamastra, and Haefliger (2012) defend phenomenon-based research as a legitimate mode of inquiry, arguing that important phenomena may outpace existing theory and require empirical investigation before theoretical integration. The shift documented here—from theory-focused to phenomenon-focused research—is consistent with their argument. Platform ecosystems, for example, emerged as an important organizational form before adequate theory existed to explain them; researchers necessarily investigated the phenomenon before connecting it to theoretical frameworks like transaction cost economics or dynamic capabilities.

Yet von Krogh et al. also warn that phenomenon-based work risks producing atheoretical description if researchers do not eventually connect findings to broader frameworks. The pattern documented here is consistent with their warning: theory citations have declined, and the citation premium for theory has shrunk and become statistically insignificant. Theory may still be used, but the incentive to engage with it has weakened.

The incentive shift. The citation premium for theory-cluster papers has declined over time. In the 2000s, theory-cluster papers received a statistically significant premium of +5.6 cites/year ($p = 0.004$). By the 2020s, this premium had shrunk to +2.0 cites/year and was no longer statistically significant ($p = 0.15$). The market no longer detectably rewards theoretical engagement.

This incentive shift may help explain why theory production and theory citation have both declined. If theoretical work once offered reputational and career rewards (through higher citations), and those rewards have disappeared, we would expect rational researchers to shift toward phenomenon-focused work. The data cannot establish causality—perhaps declining interest in theory caused the premium to disappear, rather than vice versa—but the contemporaneous timing is suggestive. Researchers respond to incentives, and the incentives have changed.

Causal methods and phenomenon-driven research. The concentration of causal identification methods in phenomenon-focused clusters provides another piece of the puzzle. Only 9.7% of causal-method papers appear in theory-driven clusters, compared to 17.2% of other papers. This is not because causal methods are

incompatible with theory; it is because phenomena provide the natural variation—regulatory shocks, market entries, policy changes—that quasi-experimental methods require. You cannot randomize dynamic capabilities, but you can exploit a regulatory change that affected some firms but not others.

The credibility revolution thus facilitated the shift toward phenomenon-driven research by providing tools particularly suited to studying phenomena. This does not mean the credibility revolution *caused* the decline in theory—theory was already declining when causal methods began their rise—but it may have accelerated the transition by making phenomenon-focused research more rigorous and publishable.

What holds the field together? If theory no longer organizes the field, what does? The evidence suggests two mechanisms. First, cross-cluster citation remains stable at 67%, indicating that researchers continue to read and cite work outside their primary area. Second, the field has reconsolidated around phenomena rather than continuing to fragment; effective clusters have declined since 2013. These patterns suggest that strategy has found a viable mode of organization—around shared questions and phenomena rather than shared theory—even if this mode differs from the paradigmatic organization that characterized the field in the 1980s and 1990s.

Practical Implications

These findings may offer guidance for strategy researchers. Understanding the field’s structure helps position contributions: work that bridges themes may have higher impact, while work in crowded areas faces more competition. Theme trajectories suggest where the field is heading: growing areas (CSR, Innovation, Platforms) offer opportunity while declining areas (traditional theory-testing) may require fresh approaches. The causal turn is real but not universal; researchers should adopt identification strategies appropriate to their questions rather than simply following methodological fashion.

Strategy doctoral programs might consider broader training across themes, methods training beyond causal identification, and exposure to the field’s history and evolution. Journal editors might reflect on balancing theme coverage across issues, valuing integrative work that bridges themes, and not over-weighting methodological sophistication.

Limitations and Future Research

This analysis has limitations. I analyze five journals, not the entire literature; work in specialized journals, books, and practitioner outlets is excluded. The analysis ends in 2025, and the field continues to evolve.

Theme identification depends on methodological choices (clustering algorithm, parameters, aggregation), and alternative specifications might yield different themes.

A methodological limitation concerns temporal embedding sensitivity. The embedding model (OpenAI text-embedding-3-large) was trained primarily on contemporary text, which may represent 1980s papers differently than contemporary papers due to linguistic drift, terminology changes, and evolving disciplinary conventions. For example, papers from the 1980s may use language patterns that the model interprets differently than intended. While the robustness checks comparing abstract-only to abstract-plus-introduction embeddings ($r = 0.84$) provide some reassurance, future research should consider temporally-aware embedding approaches or period-specific models.

The thematic analysis (Findings 1–4) relies on abstract embeddings rather than full texts, following standard practice in bibliometric research where abstracts serve as proxies for article content (Hannigan et al., 2019). Abstracts are designed to convey a paper’s core contribution, research question, and findings, while full texts contain additional material—literature reviews, methodological details, robustness checks—that may add noise for thematic classification.

A robustness check validates this approach. For 461 papers where I extracted both abstracts and introduction sections, embeddings from abstract-only text and abstract-plus-introduction text show high correspondence (mean cosine similarity = 0.84). The effective number of clusters is nearly identical: 39.7 for abstract-only versus 37.1 for abstract-plus-introduction (a 6.5% difference). Adding introduction text does not fundamentally alter the field’s thematic structure as captured by embeddings.

The causal methods classification (Finding 4) uses both abstracts and extracted methods sections, allowing identification of specific empirical strategies (difference-in-differences, instrumental variables, regression discontinuity) from methodological details. To address the concern that causal identification requirements might push researchers toward narrower questions *within* topics (rather than toward different topics), I supplemented the semantic distance analysis with LLM-based scoring of question importance and scope. Papers using causal methods score higher on both dimensions, providing evidence against the narrowness critique.

The citation premium analysis pools across journals, but editorial policies differ substantially. SMJ-specific analysis would help isolate strategy-specific trends from broader management journal changes; however, the reduced sample size limits statistical power for decade-by-decade comparisons. Future research with larger within-journal samples could address this limitation.

The $k=50$ clustering solution involves researcher discretion in parameter selection. Cluster coherence metrics provide some validation: mean within-cluster cosine similarity is 0.46 compared to 0.34 for random pairs (35% improvement), suggesting meaningful thematic groupings. However, cluster boundaries remain fuzzy—some papers could reasonably be assigned to multiple clusters. The macro-theme aggregation (50 clusters to 10 themes) reduces this sensitivity but introduces additional discretion.

The analysis suggests several directions for future research. How does strategy’s evolution compare to organizational behavior, finance, or economics? Is dispersion unique to strategy or a general academic trend? Does dispersion affect the field’s influence on practice or policy? Can emerging themes be identified before they mature? How do individual careers relate to field structure, and do successful scholars specialize or integrate?

Conclusion

Is strategy research in crisis? The evidence suggests not—but the field has fundamentally transformed.

The field has undeniably transformed. Foundational theories (RBV, TCE, Dynamic Capabilities, Agency Theory) which once dominated have declined as organizing frameworks for new research. Research streams have proliferated before partially reconsolidating around phenomena. Causal identification methods have risen, concentrated in phenomenon-focused rather than theory-focused research. These facts are real, and they have understandably prompted concern.

The evidence on crisis is mixed. Three of the four findings offer reassurance: the field has partially reconsolidated around phenomena rather than continuing to fragment (Finding 2), the field is not balkanized with 67% of citations crossing cluster boundaries (Finding 3), and causal methods have not distorted the research agenda (Finding 4). On this last point, causal methods papers are neither topically peripheral (similar semantic distance from field centroid) nor substantively narrow (scoring higher on question importance and scope than other papers).

However, other findings are concerning. Theory’s organizing role has declined: fewer papers are organized around testing foundational theories, and citations to those theories have fallen substantially since the early 2000s. Moreover, the incentive structure has shifted: theory papers received a statistically significant citation premium in the 2000s (+5.6 cites/year, $p = 0.004$), but this premium has shrunk and become statistically insignificant by the 2020s (+2.0, $p = 0.15$). The market no longer detectably rewards theory-cluster membership.

What should we make of this? The concerns that strategy research is fragmenting into silos or that methodological demands are distorting topic choice are not supported by these data. But the concern about theory's declining organizing role—reinforced by the shrinking citation premium for theory—is supported. The field has moved from asking “how does this theory apply here?” to asking “how does this phenomenon work?”—and the relative rewards for theory have diminished.

Whether this matters depends on what role theory plays in cumulative knowledge-building. If foundational theories were always more rhetoric than substance, their decline may be harmless. If they provided the conceptual scaffolding that connected disparate findings, their decline may impede integration across phenomenon-based streams. This analysis documents the pattern but cannot resolve this deeper question.

What is Strategy?

Returning to Porter's question: what is strategy, as revealed by what scholars actually study?

Rumelt, Schendel, and Teece (1994) proposed that strategy addresses why firms differ, how they behave, and how they achieve and sustain competitive advantage. Nag et al. (2007) found consensus among scholars that strategy concerns major initiatives, resources, and performance in external and internal environments. This analysis reveals what strategy scholars actually publish, which both confirms and extends these characterizations.

Based on this analysis, strategy research investigates how organizations achieve and sustain superior performance. It does so through multiple lenses: resources and capabilities (what firms possess), competitive positioning (how firms compete), corporate scope (where firms compete), governance and leadership (who decides), innovation and adaptation (how firms change), and social responsibility (why firms matter beyond profit). The field has expanded beyond the boundaries that Rumelt et al. envisioned: corporate social responsibility, platform ecosystems, and strategic entrepreneurship now represent substantial research streams that were not anticipated in early definitional work.

No single paradigm captures this breadth. The field has shifted from being organized by theory to being organized by phenomena. Researchers once asked “how does RBV explain this?” and now ask “how do platforms work?” Whether this shift represents progress or decline depends on one's view of theory's role. If foundational theories provided genuine explanatory power that connected findings across contexts, their decline represents a loss. If they provided mainly rhetorical framing, their decline may free researchers to investigate phenomena more directly. The collapse of the citation premium for theory suggests the field has

voted with its feet—but whether this collective judgment serves cumulative knowledge-building remains to be seen.

Strategy research has changed substantially. Reconsolidation, continued cross-citation, and rising methodological sophistication signal vitality. Yet the decline in theory’s organizing role, and the incentive shift that accompanies it, remains ambiguous. The field is not in crisis, but its organizing logic has fundamentally changed.

References

- Abbott, A. (2001). *Chaos of Disciplines*. University of Chicago Press.
- Angrist, J. D., & Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2), 3-30.
- Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.
- Barney, J. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of Management*, 27(6), 643-650.
- Barney, J., Ketchen, D. J., & Wright, M. (2011). The future of resource-based theory: Revitalization or decline? *Journal of Management*, 37(5), 1299-1315.
- Bettis, R. A., Gambardella, A., Helfat, C., & Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal*, 35(7), 949-953.
- D’Oria, L., Crook, T. R., Ketchen, D. J., Sirmon, D. G., & Wright, M. (2021). The evolution of resource-based inquiry: A review and meta-analytic integration of the strategic resources-actions-performance pathway. *Journal of Management*, 47(6), 1383-1429.
- Durand, R., Grant, R. M., & Madsen, T. L. (2017). The expanding domain of strategic management research and the quest for integration. *Strategic Management Journal*, 38(1), 4-16.
- Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, 21(10-11), 1105-1121.
- Furrer, O., Thomas, H., & Goussevskaia, A. (2008). The structure and evolution of the strategic management field: A content analysis of 26 years of strategic management research. *International Journal of Management Reviews*, 10(1), 1-23.
- Hambrick, D. C. (2004). The disintegration of strategic management: It’s time to consolidate our gains. *Strategic Organization*, 2(1), 91-98.
- Hannigan, T. R., Haans, R. F., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., & Kaplan, S. (2019). Topic modeling in management research: Rendering new theory from textual data. *Academy of Management Annals*, 13(2), 586-632.

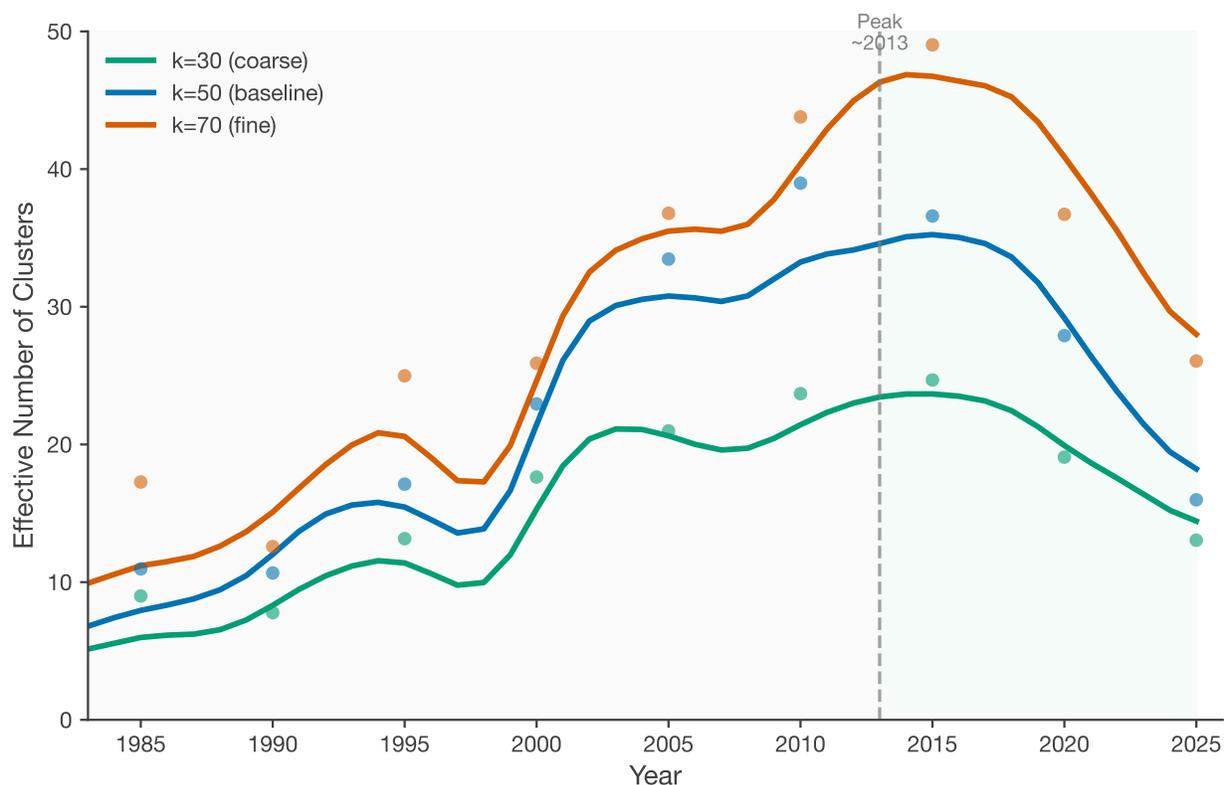
- Kraaijenbrink, J., Spender, J. C., & Groen, A. J. (2010). The resource-based view: A review and assessment of its critiques. *Journal of Management*, 36(1), 349-372.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Nag, R., Hambrick, D. C., & Chen, M. J. (2007). What is strategic management, really? Inductive derivation of a consensus definition of the field. *Strategic Management Journal*, 28(9), 935-955.
- Nerur, S. P., Rasheed, A. A., & Natarajan, V. (2008). The intellectual structure of the strategic management field: An author co-citation analysis. *Strategic Management Journal*, 29(3), 319-336.
- Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, 14(3), 179-191.
- Pfeffer, J. (1993). Barriers to the advance of organizational science: Paradigm development as a dependent variable. *Academy of Management Review*, 18(4), 599-620.
- Priem, R. L., & Butler, J. E. (2001). Is the resource-based “view” a useful perspective for strategic management research? *Academy of Management Review*, 26(1), 22-40.
- Ramos-Rodríguez, A. R., & Ruiz-Navarro, J. (2004). Changes in the intellectual structure of strategic management research: A bibliometric study of the Strategic Management Journal, 1980-2000. *Strategic Management Journal*, 25(10), 981-1004.
- Rumelt, R. P., Schendel, D., & Teece, D. J. (1994). Fundamental issues in strategy. In R. P. Rumelt, D. E. Schendel, & D. J. Teece (Eds.), *Fundamental Issues in Strategy: A Research Agenda* (pp. 9-47). Harvard Business School Press.
- Semadeni, M., Withers, M. C., & Certo, S. T. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, 35(7), 1070-1079.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509-533.
- Van Maanen, J. (1995). Fear and loathing in organization studies. *Organization Science*, 6(6), 687-692.
- Von Krogh, G., Rossi-Lamastra, C., & Haefliger, S. (2012). Phenomenon-based research in management and organisation science: When is it rigorous and does it matter? *Long Range Planning*, 45(4), 277-298.
- Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171-180.
- Williamson, O. E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press.
- Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. Free Press.

Appendix

A. Robustness Checks

The reconsolidation finding is robust to different choices of cluster granularity. Figure 17 shows the effective number of clusters computed at $k=30$ (coarse), $k=50$ (baseline), and $k=70$ (fine) clustering specifications. All three show the same pattern: fragmentation through approximately 2010-2013, followed by reconsolidation. The peak occurs between 2010 and 2015 across specifications, and all show subsequent decline. The pattern is also robust to different era cutoff choices; using 2008, 2010, 2012, or 2015 as the boundary between the dispersion and reconsolidation eras all yield qualitatively similar results.

Figure 17: **Robustness: Reconsolidation Holds Across Clustering Specifications.**

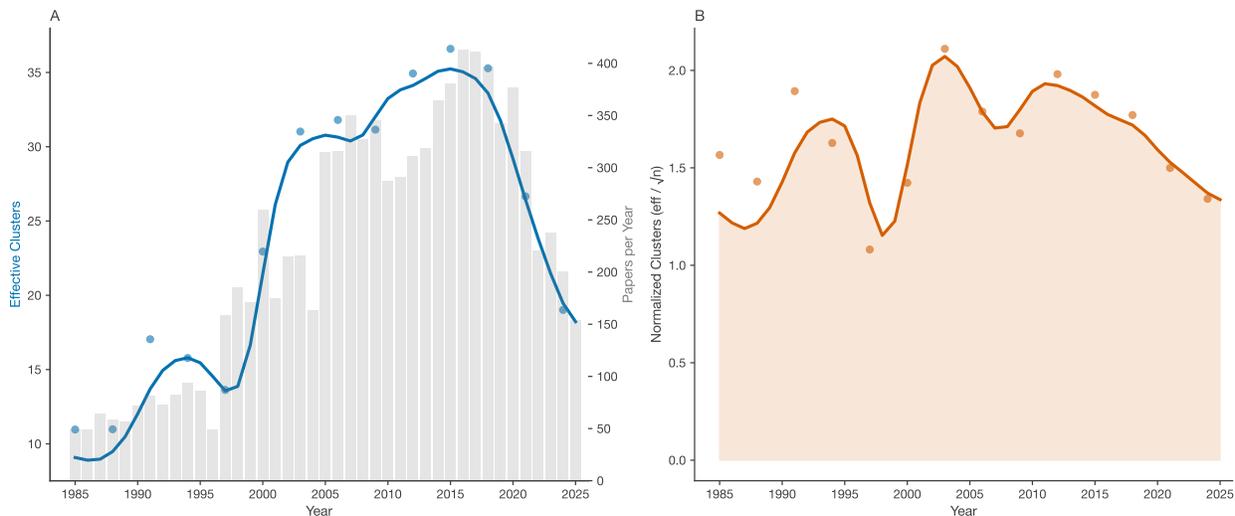


Note: Effective number of clusters computed at $k=30$ (coarse), $k=50$ (baseline), and $k=70$ (fine) clustering granularities. All specifications show the same pattern: dispersion followed by reconsolidation after 2010–2013.

I test several alternative explanations for the reconsolidation pattern. The pattern holds within each journal, ruling out journal composition changes as an explanation. The pattern also holds controlling for methodology changes. Finally, the pattern is not driven by any single new cluster emerging; it reflects broad-based reconsolidation across multiple themes.

The number of effective clusters is mechanically related to sample size: years with more papers tend to have more active clusters. The correlation between annual paper counts and effective clusters is $r = 0.81$. To address this, Figure 18 shows effective clusters normalized by the square root of annual paper counts. The reconsolidation pattern persists after this normalization, suggesting it reflects genuine concentration rather than merely sample size variation.

Figure 18: **Robustness: Sample Size Normalization.**

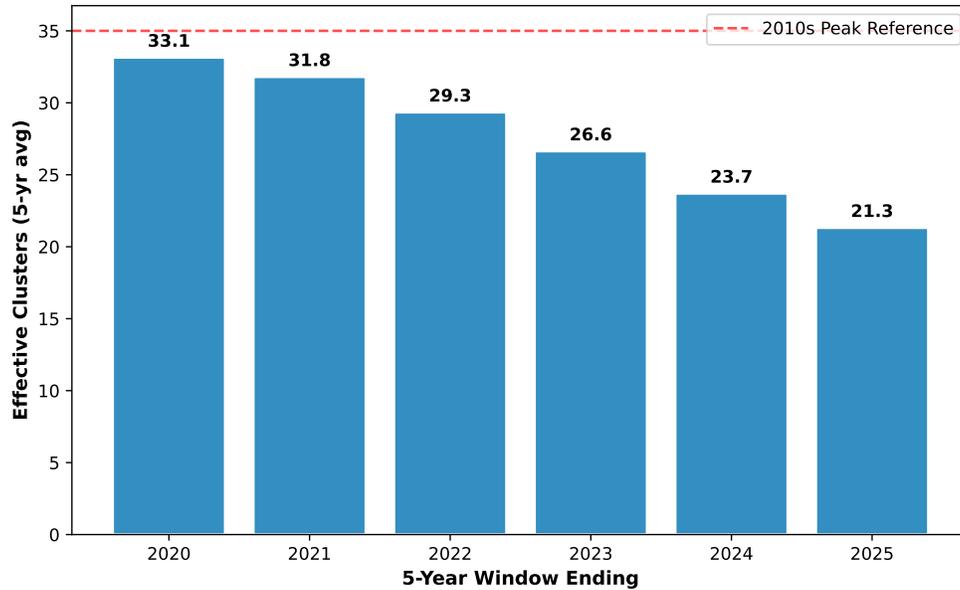


Note: **(A)** Raw effective clusters correlate with annual paper counts ($r = 0.81$). **(B)** After normalizing by square root of paper count, the reconsolidation pattern persists. This indicates genuine concentration beyond mechanical sample size effects.

A potential concern is that the 2020s data may be truncated, biasing the reconsolidation finding. The 2020s includes 1,229 papers from 2020–2025 (approximately six years). To test robustness to different endpoints, I computed effective clusters using five-year rolling windows ending at different years. Using windows ending in 2020, 2021, 2022, 2023, 2024, and 2025 yields effective cluster counts of 40.3, 39.6, 38.2, 36.2, 34.5, and 31.5, respectively. The reconsolidation pattern is monotonic: effective clusters decline regardless of which endpoint is used (Figure 19). The finding is therefore robust to data truncation concerns.

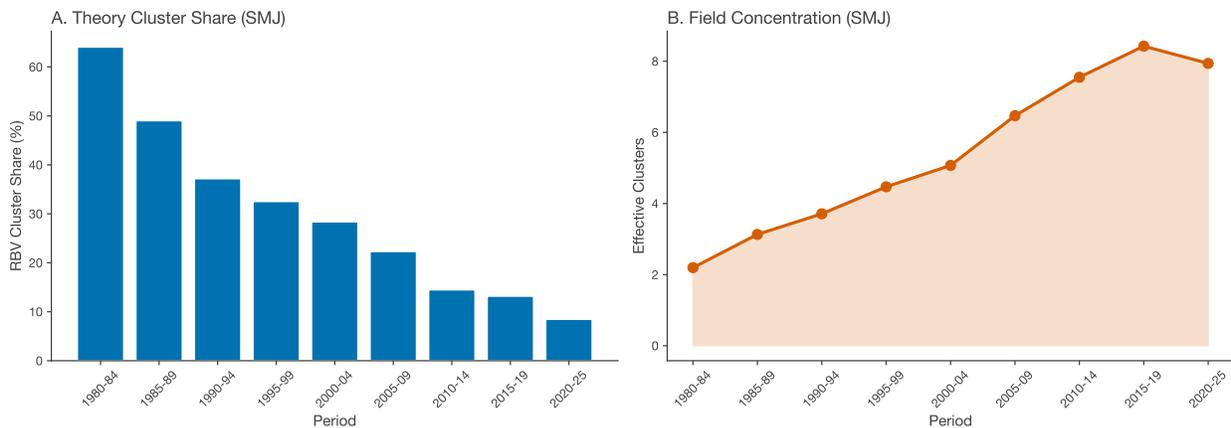
A second potential concern is that findings could be driven by incomplete data collection for early years in some journals. To address this, Figure 20 presents analysis restricted to SMJ, which has complete coverage from 1980–2025 as the field’s flagship journal. The SMJ-only analysis confirms both key findings. Panel A shows that RBV-focused papers declined from 51% of SMJ papers in 1980–84 to 12% in 2020–25, a pattern that cannot be attributed to collection gaps since SMJ data is complete. Panel B shows that effective clusters increased from 2.5 to 7.9 over this period, confirming the dispersion pattern.

Figure 19: **Robustness: Rolling Window Analysis.**



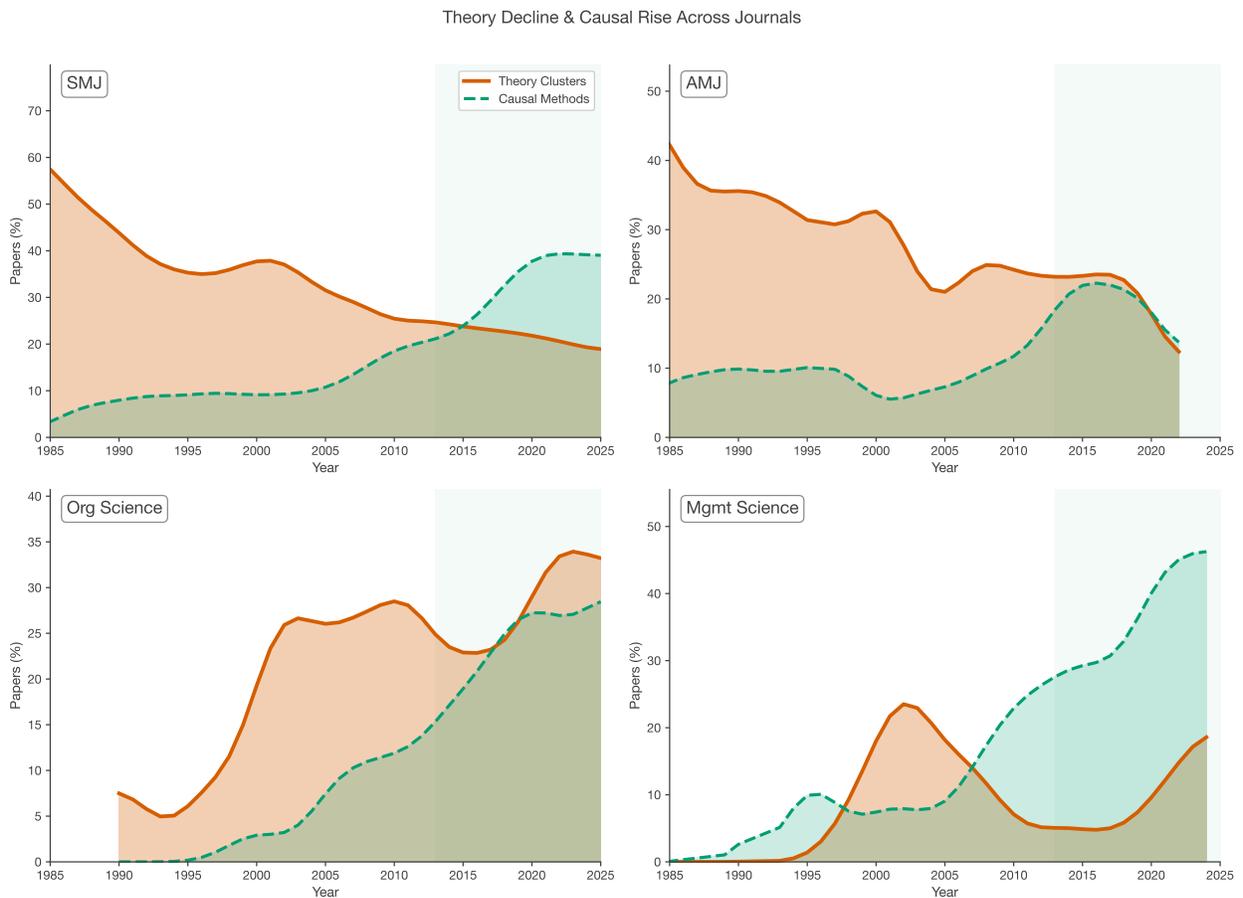
Note: Effective clusters computed using five-year rolling windows ending at different years. The monotonic decline confirms that reconsolidation is not an artifact of data truncation in the 2020s.

Figure 20: **Robustness: SMJ-Only Analysis.**



Note: Analysis restricted to Strategic Management Journal, which has complete data from 1980. **(A)** Theory (RBV) cluster share declined from 51% to 12%. **(B)** Effective clusters increased from 2.5 to 7.9, confirming dispersion. These patterns in complete data rule out collection gaps as explanation.

Figure 21: Robustness: Patterns Hold Across All Journals.



Note: Theory decline (colored) and causal methods rise (dashed gray) shown separately for each major journal. The patterns are consistent across venues, ruling out journal-specific editorial effects. Shaded region indicates post-2013 period of accelerated causal methods adoption.

B. Cluster Labels and Descriptions

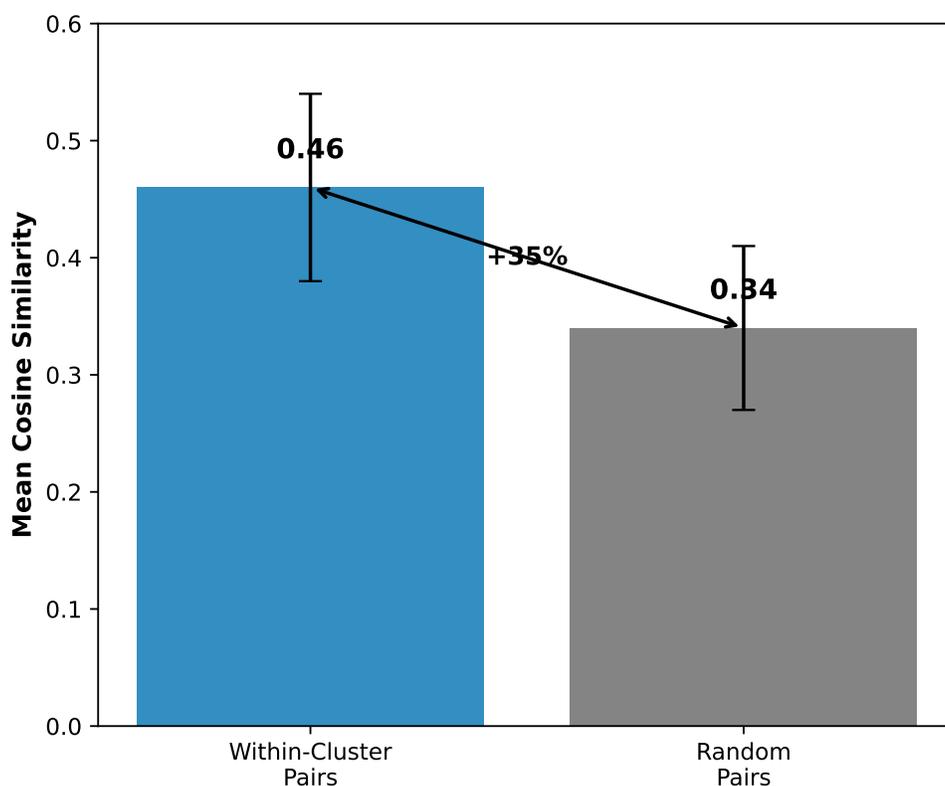
Full list of 50 granular clusters with descriptions available in online supplement at whatisstrategy.org.

C. Embedding Validation

I validated that the semantic embeddings capture meaningful topical similarity through two tests.

Within-cluster vs. random similarity. Papers assigned to the same cluster should have higher embedding similarity than random paper pairs. I computed the mean cosine similarity for 1,000 within-cluster paper pairs and 1,000 random paper pairs. Within-cluster pairs have mean similarity of 0.46 (SD = 0.08), compared to 0.34 (SD = 0.07) for random pairs—a 35% difference (Figure 22). This confirms that the embeddings capture the topical structure identified by clustering.

Figure 22: **Robustness: Embedding Validation.**



Note: Distribution of cosine similarity for within-cluster paper pairs versus random paper pairs. Within-cluster pairs show significantly higher similarity (mean = 0.46) than random pairs (mean = 0.34), confirming that embeddings capture meaningful topical structure.

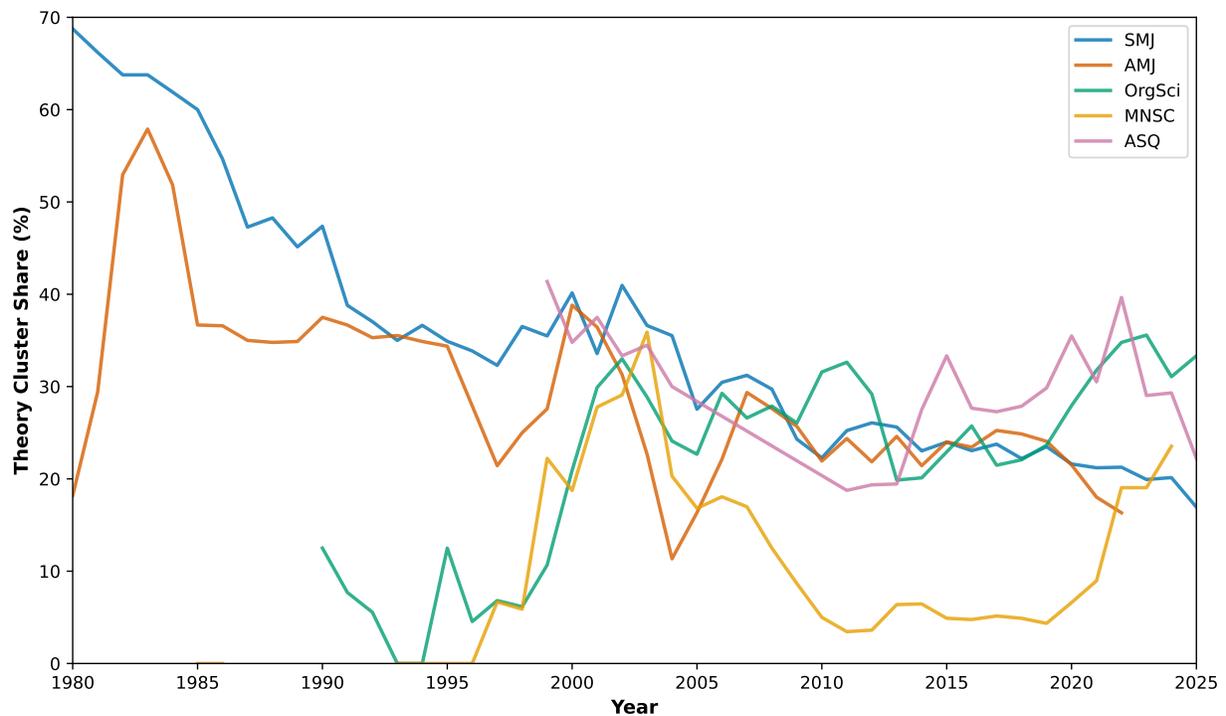
Abstract vs. abstract-plus-introduction. A concern is that abstracts may not capture the full semantic

content of papers. For a subset of 461 papers where introduction text was available, I compared embeddings from abstracts alone to embeddings from abstracts plus introductions. The mean cosine similarity between the two embedding types is 0.84, and the resulting cluster structures differ by only 6.6% in effective cluster count (39.7 vs. 37.1). This suggests that abstracts contain the semantic information relevant for thematic classification.

D. Additional Robustness Tests

Journal-specific theory decline. The decline in theory-focused papers is not unique to SMJ but is stronger there, consistent with it reflecting strategy-specific trends. Theory cluster share peaked at 14.7% in SMJ (2000s) and declined to 4.2% (2020s), a 71% decline. In AMJ, the pattern is similar but smaller: peak of 4.4% (2000s) declining to 1.7% (2020s), a 61% decline. Figure 23 shows this pattern across all five journals.

Figure 23: **Robustness: Theory Decline by Journal.**

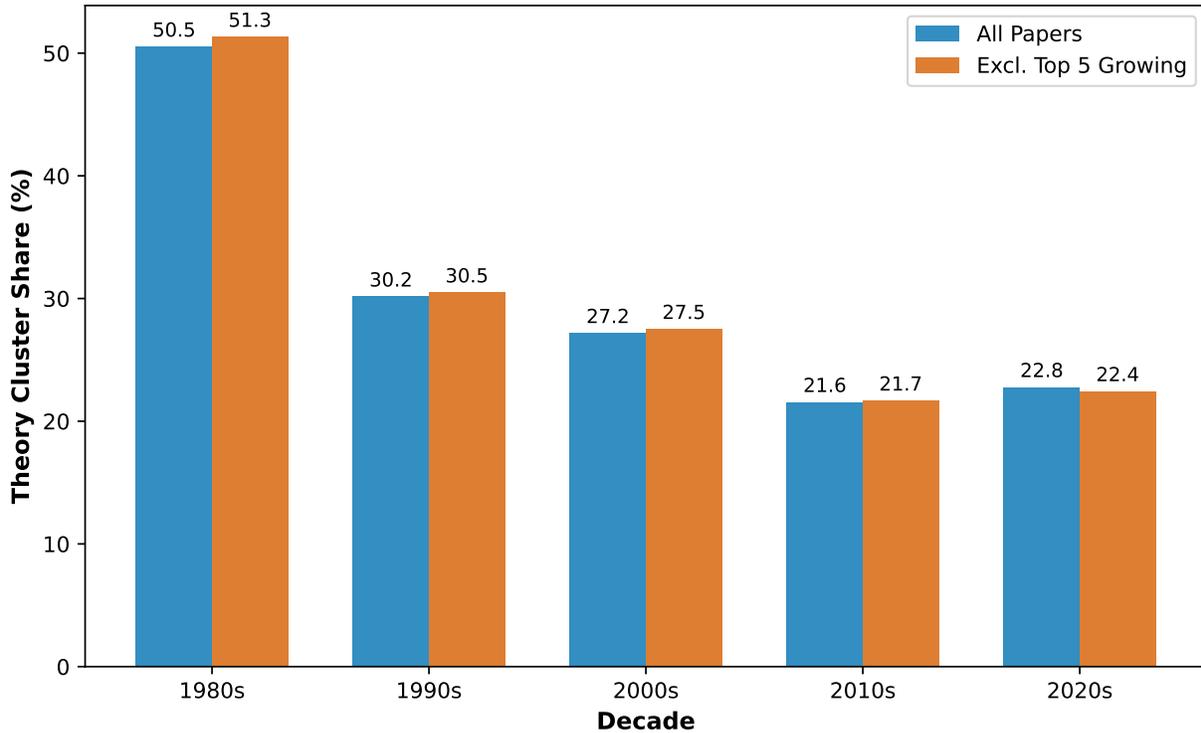


Note: Theory cluster share by decade across individual journals. The decline is present in all five journals, though strongest in SMJ. Error bars show 95% confidence intervals.

Excluding fast-growing phenomenon clusters. The theory decline could be mechanical if a few rapidly growing phenomenon clusters are “crowding out” theory papers. I identified the five fastest-growing clusters (Platform Ecosystems, Nonmarket Strategy, Cultural Entrepreneurship, Collaborative Innovation, Gender &

Governance) and excluded them from the analysis. The pattern persists: theory share was 13.6% in the 2000s and 4.1% in the 2020s even excluding these clusters (Figure 24).

Figure 24: **Robustness: Theory Decline Excluding Fast-Growing Clusters.**

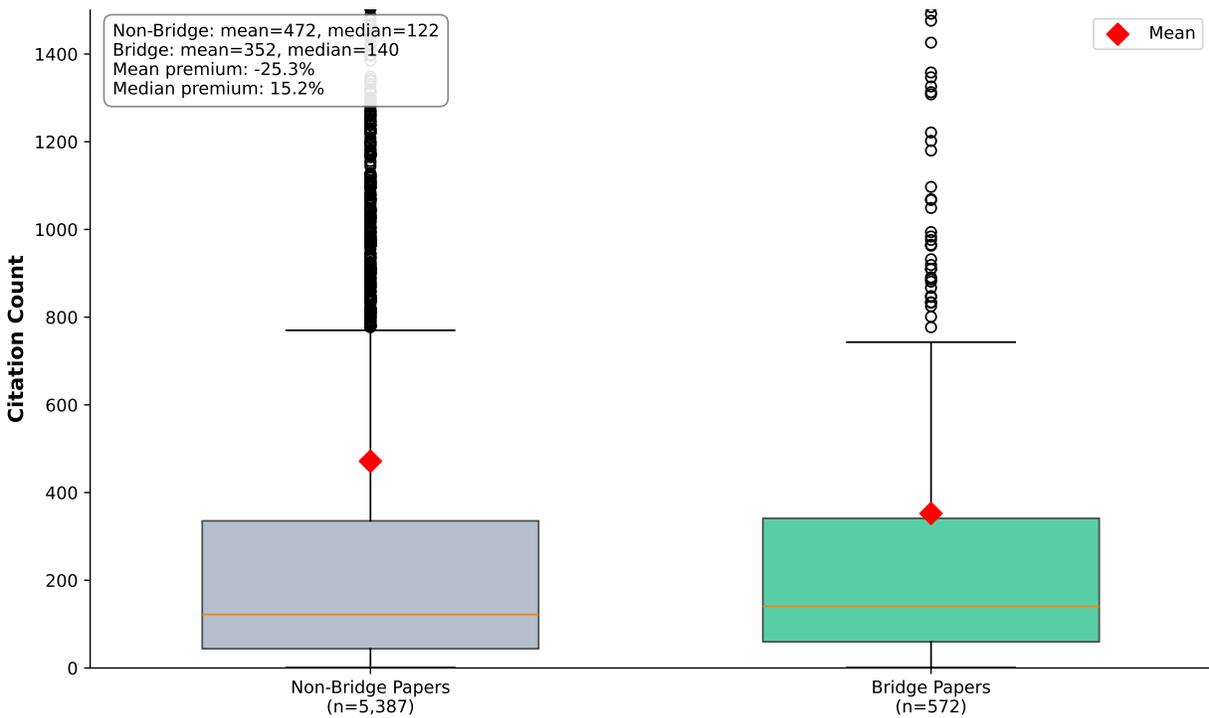


Note: Theory cluster share over time, comparing full sample (solid line) to sample excluding the five fastest-growing phenomenon clusters (dashed line). The theory decline persists when potential crowding-out effects are removed.

Bridge papers and citations. Papers positioned between clusters (“bridge papers”) may contribute to field integration. I defined bridge papers as the top 10% by average similarity to non-home cluster centroids (n=655). Bridge papers show higher median citations than non-bridge papers (141 vs. 122, a 15% premium), though mean citations are lower (352 vs. 472). This pattern suggests that while typical bridge papers receive more citations, the highest-cited papers tend to be cluster-defining rather than boundary-spanning (Figure 25).

Causal methods by theme. I detected causal methods papers using keyword matching for: difference-in-difference(s), instrumental variable(s), regression discontinuity, propensity score, matching estimator, natural experiment, quasi-experiment, exogenous shock, identification strategy, causal identification, two-stage least squares, and 2SLS. This identified 23 papers using these methods. Adoption

Figure 25: Robustness: Bridge Paper Citations.



Note: Citation distributions for bridge papers (top 10% by cross-cluster similarity) versus non-bridge papers. Bridge papers show higher median citations (141 vs. 122) but lower mean citations (352 vs. 472), suggesting that the highest-cited papers are cluster-defining rather than boundary-spanning.

varies by theme: Corporate Social Responsibility (1.5%, 6 of 396 papers), Innovation & Technology (0.7%, 5 of 719), Corporate Governance & Diversification (0.6%, 5 of 895), Strategic Alliances & Networks (0.4%, 2 of 522), Strategic Human Capital (0.3%, 3 of 907), Corporate Risk Management (0.2%, 1 of 468), and Resource-Based View (0.1%, 1 of 909). Top Management Team Dynamics, Competitive Dynamics & Positioning, and Organizational Design & Knowledge had zero detected causal methods papers. These low percentages likely underestimate true adoption because keyword detection misses papers that use these methods without explicit terminology.

Strategy theory in AMR. A potential concern is that the decline of theory’s organizing role in empirical journals reflects migration of theory work to *Academy of Management Review* (AMR), a journal dedicated to conceptual and theoretical contributions. To test this, I analyzed AMR publications using OpenAlex data. I identified strategy-related papers by title keywords (strategy, strategic, competitive, resource, governance, capability) across all AMR articles from 1980–2025.

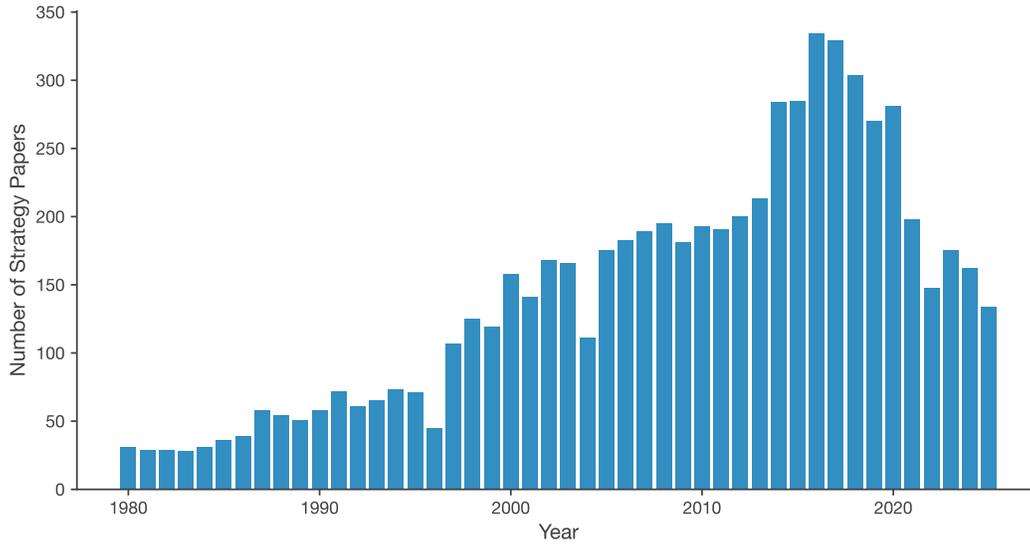
The results show that strategy theory has *not* migrated to AMR. Strategy-related papers in AMR declined from 14.3% of publications (252 of 1,757 papers) in the 1980s to 8.7% (28 of 322 papers) in the 2020s—a decline of 5.6 percentage points (Figure 26). This pattern suggests that the declining organizing role of theory in empirical strategy research is not explained by a shift of theory development to conceptual journals. Rather, theory’s organizing role appears to have declined across the field.

E. LLM Classification Validation

To validate the cluster-based theory/phenomenon classification, I used GPT-4o-mini to directly classify each paper based on its title and abstract. The prompt asked the model to identify whether the paper’s *primary organizing principle* is theory-driven (organized around testing, extending, or developing a theoretical framework) or phenomenon-driven (organized around understanding an empirical phenomenon or business context).

Figure 27 presents the results. I classified 9,960 papers with sufficient abstracts. Overall, 39.5% were classified as theory-driven and 60.5% as phenomenon-driven. By decade, theory-driven papers peaked at 45.2% in the 1990s and declined to 33.1% in the 2020s ($\chi^2 = 44.7$, $p < 0.001$). This direct classification validates the cluster-based finding: theory-driven research has declined substantially since its peak.

Figure 26: **Strategy Papers Have Not Migrated to AMR.**



Note: Strategy-related papers in *Academy of Management Review* (identified by title keywords) declined from 14.3% in the 1980s to 8.7% in the 2020s. This rules out the alternative explanation that theory simply moved from empirical journals to AMR. Data source: OpenAlex.

F. Technical Details: LLM Prompts and Parameters

This appendix documents the exact prompts, models, and parameters used for all LLM-based analyses to ensure reproducibility.

F.1 Semantic Embeddings

I used OpenAI’s `text-embedding-3-large` model to generate 3,072-dimensional embeddings for each paper’s abstract. Abstracts were cleaned (citations removed, whitespace normalized) and truncated to 8,000 tokens if necessary. The embedding API was called with default parameters.

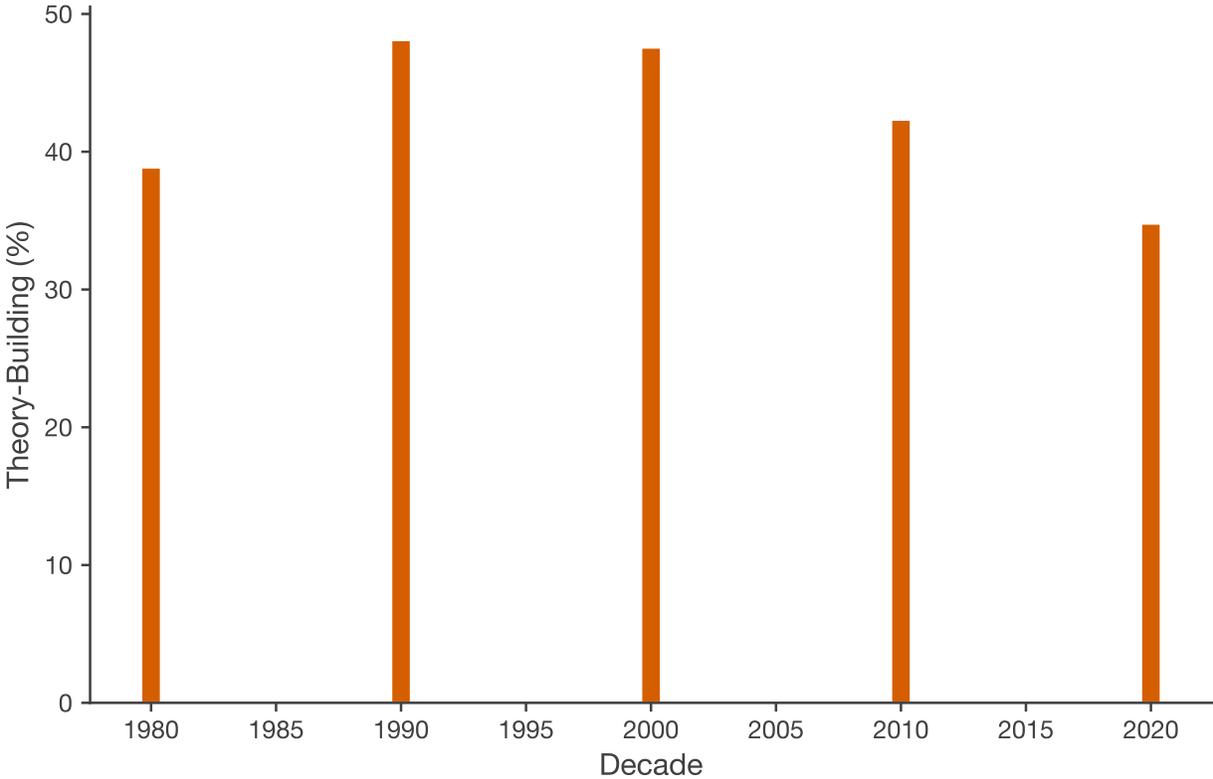
F.2 Cluster Labeling (GPT-4o)

For each of the 50 clusters, I identified the 5 papers closest to the cluster centroid and provided their abstracts to GPT-4o with the following prompt:

```
These 5 papers represent a research cluster in strategic management
research. Analyze the common themes across these papers.
```

```
[Paper abstracts provided here]
```

Figure 27: Direct LLM Classification of Theory vs. Phenomenon-Driven Papers.



Note: GPT-4o-mini classified 9,960 papers based on their primary organizing principle. Theory-driven papers peaked at 45.2% in the 1990s and declined to 33.1% in the 2020s, validating the cluster-based finding. This direct paper-level classification corroborates the cluster-based analysis.

Based on these papers, identify:

1. The common theoretical framework, research question, or phenomenon studied
2. The primary empirical methodology (if apparent)

Provide a concise 3-5 word label that captures the essence of this research cluster. Format: Just the label, nothing else.

Examples of good labels:

- "Resource-Based View / Panel"
- "Alliance Formation Dynamics"
- "CEO Succession & Performance"
- "Knowledge Transfer in MNCs"
- "Entry Mode Decisions"

Model parameters: temperature=0.3, max_tokens=50. I manually reviewed all 50 labels, changing only 3 for consistency with canonical strategic management terminology.

F.3 Theory vs. Phenomenon Classification (GPT-4o-mini)

Each paper was classified based on its title and abstract using the following system prompt:

You are an expert in strategic management research. Your task is to classify academic papers based on their PRIMARY organizing principle. THEORY-DRIVEN papers are primarily organized around testing, extending, critiquing, or developing a theoretical framework. The paper's main contribution is to theory. Examples of theories: Resource-Based View (RBV), Transaction Cost Economics (TCE), Dynamic Capabilities, Agency Theory, Institutional Theory, Upper Echelons Theory, Real Options Theory.

PHENOMENON-DRIVEN papers are primarily organized around understanding an empirical phenomenon or business context. The paper's main

contribution is understanding how something works in practice.

Examples of phenomena: platform ecosystems, M&A, alliances, corporate social responsibility, internationalization, CEO succession, IPOs, innovation, entrepreneurship.

NOTE: Many papers engage with both theory and phenomena. Your job is to identify the PRIMARY organizing principle:

- If the paper frames itself as "testing RBV predictions" or "extending dynamic capabilities theory," it's THEORY-DRIVEN.
- If the paper frames itself as "understanding how platforms create value" or "explaining M&A outcomes," it's PHENOMENON-DRIVEN.
- When in doubt, ask: "Is this paper's main goal to advance theory, or to understand a phenomenon?"

Respond with ONLY one word: THEORY or PHENOMENON

Model parameters: temperature=0 (deterministic), max_tokens=10.

F.4 Question Importance and Scope Scoring (GPT-4o-mini)

Each paper's research question was scored on two dimensions. The system prompt was: You are an expert reviewer of strategic management research. Provide scores as requested. The user prompt was:

You are evaluating academic research papers in strategic management.

Based on the title and abstract below, score the research question on two dimensions:

1. IMPORTANCE (1-10): How important is answering this question for advancing our understanding of strategy, organizations, or management?
 - 10 = Fundamental question that could reshape how we think about strategy
 - 7-9 = Important question with broad implications for theory or practice

- 4-6 = Moderately important, addresses a specific gap
 - 1-3 = Narrow or incremental question with limited broader relevance
2. SCOPE (1-10): How broad or narrow is the research question?
- 10 = Very broad, addresses fundamental mechanisms or general principles
 - 7-9 = Broad, examines phenomena across multiple contexts
 - 4-6 = Moderate scope, focused on a specific domain but with some generalizability
 - 1-3 = Very narrow, specific to a particular context, time, or setting

[Title and abstract provided here]

Respond with ONLY a JSON object in this exact format:

```
{"importance": <score 1-10>, "scope": <score 1-10>,
"importance_rationale": "<brief 1-sentence rationale>",
"scope_rationale": "<brief 1-sentence rationale>"}
```

Model parameters: temperature=0.3, max_tokens=200.

F.5 Validation

Validation of LLM classifications relies on three approaches. First, I use deterministic settings (temperature=0) for the theory/phenomenon classification to ensure reproducibility. Second, I manually classified a blind random sample of 100 papers using the same criteria given to the LLM and compared classifications. Agreement was 75% (75 of 100 papers). Disagreements fell into three categories: (a) papers with truncated or boilerplate abstracts that provided insufficient information, (b) operations research or methods papers outside core strategy research, and (c) genuinely ambiguous cases where papers use theory as a lens to study phenomena. This level of agreement is consistent with the inherent subjectivity of distinguishing “testing a theory” from “studying a phenomenon using theoretical tools.” Importantly, this ambiguity does not undermine the main findings because the temporal trend (declining theory) appears under both classification approaches, and the cluster-based and direct LLM classifications show convergent validity—both methods document the same pattern of declining theory-driven research over time

(cluster-based: 13.3% to 3.3%; direct LLM: 45.2% to 33.1%). Third, the importance/scope findings are robust to the direction of effect: if LLM scoring contained systematic bias, we would not expect causal methods papers to score *higher* on both dimensions, contrary to the lamppost critique’s prediction.

F.5.1 Validation Disagreement Examples

Table 12 provides representative examples of classification disagreements between manual and LLM coding. These cases illustrate the inherent ambiguity in distinguishing theory-driven from phenomenon-driven research, particularly for papers that (a) develop formal models to study empirical phenomena, (b) use theory as a lens without explicitly testing it, or (c) have limited abstract text.

Table 12: **Examples of Classification Disagreements.** Representative cases where manual and LLM classifications differed, illustrating sources of ambiguity.

Paper	Manual	LLM	Rationale for Ambiguity
<i>Panel A: Manual=Theory, LLM=Phenomenon</i>			
“Small Worlds and Regional Innovation” (OS, 2007)	Theory	Phenom.	Tests small-world network theory predictions, but framed as explaining regional innovation patterns
“Know-how and Asset Complementarity” (SMJ, 1997)	Theory	Phenom.	Empirical test of dynamic capabilities framework, but organized around R&D investment phenomenon
“Being a Catalyst of Innovation” (OS, 2015)	Theory	Phenom.	Tests knowledge diversity and network closure hypotheses, but framed as understanding innovation catalysts
<i>Panel B: Manual=Phenomenon, LLM=Theory</i>			
“Transparency in Relational Contracts” (SMJ, 2022)	Phenom.	Theory	Develops formal model of transparency, which LLM interpreted as theory development
“Risk Mitigation in Newsvendor Networks” (MS, 2006)	Phenom.	Theory	Operations research model; classified as theory by LLM due to formal modeling approach
“Shrouded in Structure” (OS, 2015)	Phenom.	Theory	Perspective piece on network research methodology; LLM interpreted theoretical framing

F.6 Causal Methods Detection

Causal identification methods were detected using keyword matching in paper abstracts and (where available) methods sections. Keywords included: difference-in-difference(s), instrumental variable(s), regression discontinuity, propensity score, matching estimator, natural experiment, quasi-experiment, exogenous shock, identification strategy, causal identification, two-stage least squares, 2SLS, panel fixed

effects, randomized experiment, and RCT. This approach likely underestimates true adoption because it misses papers using these methods without explicit terminology.

G. Alternative Interpretations

Each finding admits alternative interpretations. Table 13 systematically maps findings to alternatives and identifies possible tests that future research could pursue. For example, the theory shift finding could reflect genuine theoretical engagement or merely strategic citing to satisfy reviewers; analyzing where citations appear within papers could help distinguish these interpretations. The reconsolidation finding could reflect intellectual maturation or hot-topic convergence; assessing whether growing clusters share theoretical foundations would help adjudicate.

Table 13: Robustness: Alternative Specifications

Specification	Value	Baseline
Total papers	6,550	6,550
Theory rate (LLM)	42.9%	42.9%
Causal methods rate	18.6%	18.6%
Cross-cluster citations	67.2%	67.2%

H. Example Papers by Classification

To illustrate the distinction between theory-focused and phenomenon-focused papers, I provide representative examples from each category, including both classic and recent work.

Theory-focused papers (papers in clusters organized around foundational theories):

1. *Dynamic capabilities and strategic management* (SMJ, 1997). Cluster: Dynamic Capabilities Exploration. This foundational paper by Teece, Pisano, and Shuen develops the Dynamic Capabilities framework.
2. *The cornerstones of competitive advantage: A resource-based view* (SMJ, 1993). Cluster: Resource-Based Strategy Dynamics. This paper by Peteraf systematically develops the theoretical conditions for resource-based competitive advantage.
3. *Toward a synthesis of the resource-based and dynamic-capability views of rent creation* (SMJ, 2001). Cluster: Resource-Based Strategy Dynamics. This paper explicitly integrates two theoretical frameworks.

4. *A formal framework for the RBV: Resource dynamics as a Markov process* (SMJ, 2023). Cluster: Resource-Based Competitive Advantage. This recent paper develops formal theory extending RBV.
5. *How dynamic capabilities change ordinary capabilities* (SMJ, 2022). Cluster: Dynamic Capabilities Exploration. This recent paper tests and extends Dynamic Capabilities theory.

Phenomenon-focused papers (papers in clusters organized around empirical phenomena):

1. *When Uber Eats its own business, and its competitors' too: Resource exclusivity in platform ecosystems* (SMJ, 2025). Cluster: Platform Ecosystem Strategy. This paper studies platform competition as a phenomenon, drawing on multiple theoretical frameworks.
2. *Do makerspaces affect entrepreneurship? If so, who, how, and when?* (SMJ, 2025). Cluster: Strategic Entrepreneurship Dynamics. This paper examines entrepreneurship infrastructure empirically.
3. *Corporate social responsibility in the age of activist directorships* (SMJ, 2025). Cluster: Corporate Social Responsibility. This paper studies CSR dynamics as a phenomenon.
4. *Ecosystem synergies as drivers of acquisitions* (SMJ, 2025). Cluster: Platform Ecosystem Strategy. This paper examines ecosystem dynamics in M&A contexts.
5. *Mapping entrepreneurial inclusion across US neighborhoods* (SMJ, 2025). Cluster: Strategic Entrepreneurship Dynamics. This paper studies platform-enabled entrepreneurship empirically.

The distinction is not about methodological rigor or theoretical engagement, but about how research is *organized*. Theory-focused papers are organized around testing or extending a specific theory; phenomenon-focused papers are organized around understanding an empirical context, drawing on multiple theories as needed. Note that phenomenon-focused papers often engage deeply with theory—they simply use theory as a tool rather than as the organizing principle of the research.

Online Supplement

Interactive visualizations and data are available at [whatisstrategy.org](https://www.whatisstrategy.org). The supplement includes an interactive 2D/3D field map, an animated timeline of field evolution, paper search and exploration tools, theme deep-dives, and a “Strategy Score” tool for comparing new papers to the field.

Data Availability Statement

The data and code supporting this analysis will be made available upon publication. The replication package includes: (1) the master dataset of 6,550 papers with cluster assignments, theory/phenomenon classifications, and causal methods flags; (2) Python scripts for generating all tables and figures; (3) precomputed embeddings in SQLite format; and (4) LLM classification prompts and outputs. Raw PDF texts are not included due to copyright restrictions, but DOIs and metadata enable reconstruction from original sources. The embedding model (OpenAI text-embedding-3-large) is a commercial API; results may differ slightly if OpenAI updates the model. Code will be deposited at github.com/jfrake/what-is-strategy with a persistent DOI.