Advances in Management & Applied Economics, Vol. 17, No. 1, 2027, 1-21
ISSN: 1792-7544 (print version), 1792-7552(online)
https://doi.org/10.47260/amae/1711
Scientific Press International Limited

Data-related Disclosure and Interregional Collaborative Innovation: Partner Search and External Finance

Feng Hong[1], Haoran Guo[1*], Jiayuan Wu[1], and Zhengyang Li[1]

Abstract

Interregional collaborative innovation requires firms to identify compatible partners and finance projects with uncertain returns. Grounded in an information-frictions perspective, this study examines how corporate data-related disclosure facilitates interregional collaborative innovation through two parallel channels: partner-search capability and access to external finance. Using an unbalanced panel of 3,105 Chinese A-share listed firms (24,250 firm-year observations) from 2016 to 2024, our baseline firm fixed-effects Poisson estimates show that a one-standard-deviation increase in data-related disclosure is associated with an 18.95% increase in the conditional expected count of interregional joint patent applications. This positive effect remains robust across alternative disclosure and outcome measures, nonmanufacturing subsamples, lagged specifications, high-dimensional industry-by-year and province-by-year fixed effects, and instrumental-variable control-function estimation. Mechanism analyses confirm that both partner-search capability and external financing access serve as significant transmission pathways. Furthermore, cross-sectional heterogeneity analyses reveal that the positive effect of disclosure is significantly more pronounced among firms with higher initial R&D intensity, smaller firms, and non-state-owned enterprises (non-SOEs). Overall, our findings demonstrate that transparent data-related disclosure mitigates spatial information asymmetries by enhancing mutual partner evaluation and external financing access.

JEL classification numbers: D83, O31, O32.
Keywords: Data-related disclosure, Interregional collaborative innovation, Partner search, External finance, Information frictions.

[1] School of Economics and Management, North China University of Technology, Beijing 100144, China. [*]Corresponding author.

Article Info: *Received:* October 24, 2026. *Revised*: November 17, 2026.
Published online: November 24, 2026.

1. Introduction

Firms increasingly depend on external knowledge to develop new products and technologies. While interregional collaboration broadens access to geographically dispersed knowledge, it also requires managers to evaluate unfamiliar counterparts and commit resources before joint ventures yield observable returns. Geographic distance inherently hinders firms from learning about prospective partners through repeated local interactions. Consequently, even firms endowed with valuable data resources and technical competencies may remain obscure to potential collaborators across regions. The core managerial challenge is therefore how to render these underlying capabilities sufficiently visible and credible, enabling firms to identify suitable partners and structure financially viable collaborations.

Data-related disclosure can mitigate this information friction. Corporate reports convey critical information regarding data assets, application scenarios, governance mechanisms, and related commercial activities. Such disclosures provide external stakeholders with a verifiable basis for assessing a firm's knowledge base and potential contributions to collaborative endeavors. This informational role is distinct from both the formal accounting recognition of data assets and the physical deployment of digital technologies. Corporate disclosure research links transparent reporting to external evaluation and efficient resource allocation (Healy and Palepu, 2001), while innovation narratives offer forward-looking signals of future firm performance (Chu et al., 2025). Together, these insights motivate examining corporate data-related disclosure as a strategic informational input into interregional collaborative innovation.

Prior studies have explored closely related phenomena from multiple angles. For instance, Huang et al. (2026) investigate data asset disclosure and corporate innovation, whereas Li and Cao (2026) focus on government data openness and interregional collaborative innovation, delineating firm-level communication from the macro-level public information environment. In the domain of supply chain management, Qiao et al. (2026) examine digital finance and supply chain resilience across information, financing, and innovation channels. Furthermore, Hu (2026) analyzes how downstream digital transformation drives upstream green innovation, illustrating knowledge spillovers across organizational boundaries. While these studies underscore the dual importance of information flow and capital access, they primarily differentiate corporate communication from digital financial services and technology adoption. An unaddressed question remains: how does firm-level disclosure simultaneously resolve the twin challenges of locating external cross-regional partners and securing project-level financing?

To address this gap, we develop an information-frictions framework centered on two parallel mechanisms: partner-search capability and external financing access. Partner search is inherently a mutual evaluation process: a focal firm must screen compatible collaborators, while prospective partners require verifiable signals to assess the firm in return. Transparent disclosure facilitates this matching process by equipping managers with credible signals during outreach and guiding inbound

inquiries from external partners. Concurrently, project financing addresses a distinct constraint: capital providers must rigorously evaluate project feasibility, resource commitments, and return profiles. We therefore model partner search and capital acquisition as parallel, non-sequential pathways of collaboration. Finally, we examine how these mechanisms vary across three critical organizational contingencies: baseline R&D capacity, firm size, and ownership structure.

We empirically test this framework using an unbalanced panel of 3,105 nonfinancial A-share listed firms (24,250 firm-year observations) on the Shanghai and Shenzhen main boards from 2016 to 2024, combining corporate annual report disclosures, interregional joint patent application records, and firm-level financial and governance data. Interregional collaborative innovation is measured by the count of joint patent applications co-filed with external organizations located in other provincial-level regions over the 12-month window following annual report release. Controlling for firm-level characteristics as well as firm and fiscal-year fixed effects, our baseline fixed-effects Poisson estimates indicate that a one-standard-deviation increase in data-related disclosure is associated with an 18.95% increase in the conditional expected collaboration count. This positive relationship remains robust across alternative variable definitions, sample restrictions, lagged explanatory variables, stricter multi-way fixed effects, and a Poisson control-function specification utilizing an excluded peer-reporting instrument. Mechanism and heterogeneity analyses further substantiate the parallel mediating roles of partner-search capability and external financing access, and demonstrate that the effect of disclosure is significantly stronger among high-R&D, smaller, and non-state-owned enterprises.

This study makes two primary contributions to the literature on corporate disclosure and collaborative innovation. First, it advances the disclosure literature by addressing a concrete interorganizational coordination challenge—the screening, evaluation, and financing of cross-provincial collaborative ventures—thereby linking corporate communication directly to boundary-spanning joint innovation rather than solely to internal R&D output. Second, our framework disaggregates collaboration constraints into partner search and project financing as parallel operational pathways, and empirically demonstrates how organizational endowments in absorptive capacity, firm scale, and ownership structure moderate the returns to corporate transparency. Section 2 reviews the literature and develops the hypotheses; Section 3 describes the research design and data; Section 4 presents the empirical results and robustness checks. Section 5 concludes with the main findings, theoretical and managerial implications, and directions for future research.

2. Literature Review and Hypothesis Development

2.1 Data-related disclosure and interregional collaborative innovation

Data-related disclosure captures what firms communicate regarding their data resources, analytics capabilities, and organizational applications. Conceptually, it is distinct from the formal balance-sheet recognition of data assets, general financial reporting quality, and the physical deployment of digital technologies. By conveying verifiable operational details, such disclosure enables external stakeholders to evaluate intangible competencies that are difficult to observe directly. Prior disclosure research demonstrates that reducing informational uncertainty facilitates contracting efficiency and capital allocation (Healy and Palepu, 2001). These informational benefits become particularly salient when interorganizational collaboration requires prospective partners to evaluate each other's technical capabilities before committing specialized resources.

Geographic separation amplifies this evaluation burden. While spatial proximity facilitates informal knowledge exchange and trust building, interregional alliances allow firms to tap into complementary knowledge pools beyond local networks (Boschma, 2005; Rosenkopf and Almeida, 2003). Public descriptions of data resources, technical architectures, and application scenarios help geographically distant counterparts assess whether their technological expertise and R&D objectives are mutually compatible. The economic value of such disclosure stems from facilitating an informed assessment of technical and organizational fit rather than from reporting volume alone. Consistently, scholarship on corporate disclosure, patenting, and trade secrecy highlights how strategic information sharing shapes incentives for interorganizational technological cooperation (Boot and Vladimirov, 2025).

Nevertheless, corporate transparency entails strategic trade-offs. Granular disclosure can expose commercially sensitive knowledge to rivals, and managerial reporting incentives may vary across institutional contexts (Chen et al., 2024; Kao, 2025). However, when corporate communications convey credible information regarding technical capabilities and project readiness without divulging core proprietary secrets, the reduction in partner-screening and contracting frictions should outweigh potential proprietary costs. Accordingly, we expect data-related disclosure to facilitate the formation and execution of collaborative innovation across regional boundaries.

H1: Data-related disclosure is positively associated with interregional collaborative innovation.

2.2 Partner-search capability

Partner-search capability refers to a firm's capacity to identify, evaluate, and engage prospective collaborators possessing complementary knowledge and technical assets. Locating a candidate organization is only the initial step; managers must also evaluate technological compatibility, cooperative alignment, and coordination demands. Crucially, partner selection is a two-sided matching process, as a prospective counterpart must simultaneously evaluate the focal firm as a viable collaborator. Recent scholarship on expert partner selection demonstrates that observable organizational signals play a pivotal role in capability assessments when prior direct interaction is absent (Borchhardt et al., 2025). In interregional settings, the primary barrier is therefore not merely identifying distant organizations, but establishing sufficient mutual information to evaluate a potential match. Corporate disclosure enriches this external information environment and strengthens the firm's boundary-spanning search routines.

Specifically, data-related disclosure enhances partner-search capability through both outward and inward matching channels. In outward search, clear documentation of data resources, technical applications, and R&D priorities provides managers with a verifiable reference point when articulating what the firm can contribute and what complementary capabilities it seeks. In inward search, public disclosure reduces screening costs for external organizations, attracting better-informed and technologically aligned inquiries that expand the focal firm's pool of viable collaboration candidates. Combined with internal absorptive capacity to evaluate and assimilate external knowledge (Cohen and Levinthal, 1990), transparent disclosure lowers mutual assessment frictions and streamlines partner selection. More effective partner identification and evaluation, in turn, directly facilitate interregional joint research, leading to our first mechanism hypothesis.

H2a: Partner-search capability forms a positive pathway linking data-related disclosure to interregional collaborative innovation.

2.3 Access to external finance

Access to external finance reflects a firm's ability to secure capital from outside investors and creditors on feasible terms. Collaborative innovation requires substantial upfront expenditures—spanning specialized R&D personnel, cross-regional coordination, prototyping, and knowledge integration—long before technical or commercial returns materialize. Sustaining these resource commitments is challenging when internal cash flows are constrained. As the innovation financing literature emphasizes, high outcome uncertainty and limited collateralizability frequently ration innovative projects out of external capital markets (Hall and Lerner, 2010). Consequently, identifying a technically compatible partner does not guarantee project execution; securing adequate external financing represents a distinct, parallel prerequisite for interregional collaboration.

Data-related disclosure can alleviate external financing constraints by providing

investors and creditors with a transparent account of the data assets, technical competencies, and governance structures underpinning proposed collaborative projects. Substantive reporting on data applications and commercialization pathways enables capital providers to better evaluate project feasibility and monitor managerial resource commitments. Building on the foundational disclosure–financing framework of Diamond and Verrecchia (1991) and recent evidence on how informational and financial channels shape corporate resilience (Qiao et al., 2026), we argue that enhanced data-related transparency reduces information asymmetry in capital markets. Improved access to external finance enables firms to fund boundary-spanning R&D commitments without crowding out routine operations, supporting our second mechanism hypothesis.

H2b: Access to external finance forms a positive pathway linking data-related disclosure to interregional collaborative innovation.

2.4 R&D capacity, firm size, and ownership

The extent to which data-related disclosure translates into collaborative innovation should depend on key organizational contingencies, beginning with baseline R&D capacity. Firms endowed with stronger prior research capabilities are better positioned to evaluate technical compatibility, contribute specialized knowledge, and assimilate external expertise uncovered during partner search. Grounded in the absorptive-capacity perspective (Cohen and Levinthal, 1990), internal R&D competence complements external visibility by enabling firms to effectively screen inbound inquiries and execute complex joint projects. We therefore expect the positive impact of data-related disclosure on interregional collaborative innovation to be more pronounced among firms with higher initial R&D capacity.

H3a: The positive relationship between data-related disclosure and interregional collaborative innovation is stronger among firms with higher initial R&D capacity.

Firm size represents a second contingency shaping the marginal value of public disclosure. Smaller firms typically face greater informational obscurity and possess fewer established interorganizational networks or institutional channels through which distant partners and capital providers can evaluate them. By contrast, larger enterprises benefit from broader market recognition, extensive alliance histories, and diversified financing channels. Drawing on the principle that public disclosure yields greater incremental value where baseline information asymmetry is more severe (Healy and Palepu, 2001), transparent data-related reporting should play a larger marginal role in enhancing the visibility and credibility of smaller firms.

H3b: The positive relationship between data-related disclosure and interregional collaborative innovation is stronger among smaller firms.

Corporate ownership structure similarly conditions firms' reliance on public market signals. In China's institutional environment, state-owned enterprises (SOEs) often enjoy established administrative networks, policy support, and privileged access to state-backed credit, providing alternative channels for interregional coordination and project funding. Conversely, non-state-owned enterprises (non-SOEs) face heightened market screening frictions and rely more heavily on transparent public disclosure to establish technical and financial credibility with unfamiliar cross-regional partners and external financiers. We therefore expect data-related disclosure to exhibit a stronger positive association with interregional collaborative innovation among non-SOEs.

H3c: The positive relationship between data-related disclosure and interregional collaborative innovation is stronger among non-state-owned firms.

3. Research Design

3.1 Sample and timing

Our empirical analysis employs an unbalanced panel of nonfinancial A-share listed firms on the Shanghai and Shenzhen main boards from 2016 to 2024, cross-checked against the China Stock Market & Accounting Research (CSMAR) database. Following standard sample-selection procedures, we exclude financial sector firms (CSRC industry code J), firm-year observations flagged for special treatment (ST or *ST), and observations with missing industry, province, listing date, or ownership information. Firms enter the sample after their initial public offering (IPO), and we require at least two valid firm-year observations per firm to accommodate panel fixed-effects estimation.

The final analytical panel comprises 24,250 firm-year observations across 3,105 unique listed firms. Basic corporate characteristics—including firm identifiers, IPO dates, CSRC industry classifications, provincial headquarters locations, and ultimate ownership status—are obtained from the CSMAR database. Listed age is defined as ln(fiscal year − IPO year + 1). Corporate data-related disclosure metrics are constructed via automated textual analysis of corporate annual reports, interregional joint patent application data are compiled from official patent records of the China National Intellectual Property Administration (CNIPA) matched with CSMAR subsidiary files, and firm-level financial and governance variables are drawn from audited financial statements in CSMAR.

To ensure a clear temporal ordering between corporate disclosure and subsequent collaborative innovation outcomes, explanatory and control variables are measured in fiscal year t, whereas collaborative patent applications are tracked over the 12-month window immediately following the release of the fiscal year t annual report (released in April of year t + 1). Specifically, for each firm-year, the outcome observation window commences on the day following the annual report release date and spans the subsequent 12 months. Accordingly, fiscal years 2016–2024 correspond to annual report releases in 2017–2025, with the final 12-month patent

application window concluding in April 2026.

The dependent variable, Collaboration, is the nonnegative count of interregional collaborative patent applications filed within the 12-month post-release window. Specifically, it captures joint invention and utility model patent applications co-filed by the focal firm (including its majority-owned consolidated subsidiaries) and at least one external organization—such as another enterprise, university, or research institute—located in a different provincial-level administrative region. To capture genuine boundary-spanning interorganizational collaboration, co-filings solely among entities within the same corporate group (e.g., between a parent company and its subsidiaries or between sibling affiliates) are excluded.

In our empirical estimation, conditional fixed-effects Poisson models exclude firms with zero collaborative patent filings across all sample years, yielding an effective baseline estimation sample of 19,353 observations across 2,443 firms. To ensure comparability across specifications, we maintain this common sample across all columns of our baseline regressions, while our linear fixed-effects robustness specification retains the full panel of 24,250 observations.

3.2 Variables

The primary explanatory variable, Disclosure, measures the intensity of corporate data-related disclosure in annual reports. Following established textual-analysis methodologies in accounting and innovation research (Chu et al., 2025; Huang et al., 2026), we conduct automated keyword frequency analysis of annual reports to quantify narratives concerning data resources, data analytics, application scenarios, and data governance. Raw disclosure density is calculated as the frequency of data-related terms per 10,000 effective report words. To mitigate right-skewness and facilitate economic interpretation, our primary measure (Disclosure) takes the natural logarithm of positive reporting density and standardizes it across the full panel (mean = 0, standard deviation = 1), so that a one-unit increase represents a one-standard-deviation increase in log disclosure density. For robustness, we also construct Alternative disclosure, defined as the full-panel standardized raw disclosure density without logarithmic transformation.

We operationalize our two theoretical mechanisms—Partner-search capability and Access to external finance—as full-panel standardized composite indices (mean = 0, standard deviation = 1). Partner-search capability captures a firm's capacity to identify, screen, and engage external R&D collaborators, integrating indicators of boundary-spanning R&D outreach, external technical cooperation breadth, and collaborative search efficiency. Access to external finance measures the availability of external debt and equity capital on feasible terms, constructed as a reverse-coded composite index of external financing constraints and capital-raising capacity so that higher values indicate greater access to external funds. Both indices enter the joint outcome specification simultaneously to evaluate the contribution of each pathway conditional on the other.

Following prior literature (Huang et al., 2026; Li and Cao, 2026), we control for seven time-varying firm-level financial and governance characteristics: Firm size (natural logarithm of total assets), Leverage (total liabilities divided by total assets), Return on assets (net profit divided by total assets), Operating cash flow (net operating cash flow divided by total assets), Revenue growth (annual rate of change in operating revenue), Largest shareholder stake (proportion of shares held by the largest shareholder), and Listed age (ln(fiscal year − IPO year + 1)). Table 1 provides detailed definitions for all variables.

To examine cross-sectional heterogeneity, we classify firms into time-invariant subgroups based on their initial sample observation. Specifically, firms are partitioned at the firm-level sample median by initial R&D intensity (R&D expenditure scaled by operating revenue) and initial firm size (natural logarithm of total assets), as well as by initial ownership structure (non-state-owned versus state-owned enterprises). Table 2 reports summary statistics for the full analytical panel.

Table 1: Variable definitions

Variable	Definition
Collaboration	Count of interregional joint patent applications co-filed with external organizations located in other provincial-level regions over the 12-month window following annual report release.
Disclosure	Standardized log data-related reporting density; density is positive and expressed per 10,000 effective report words.
Alternative disclosure	Standardized raw disclosure density, without a logarithmic transformation.
Firm size	Natural logarithm of total assets.
Leverage	Total liabilities divided by total assets.
Return on assets	Net profit divided by total assets.
Operating cash flow	Operating cash flow divided by total assets.
Revenue growth	Annual change in revenue divided by previous-year revenue.
Largest shareholder stake	Fraction of shares held by the largest shareholder.
Listed age	ln(fiscal year − IPO year + 1).
Partner-search capability	Standardized composite index measuring a firm's capacity to identify, evaluate, and engage external collaborative R&D partners.
Access to external finance	Standardized composite index capturing a firm's ability to obtain external debt and equity financing on feasible terms.
High initial R&D	Initial R&D intensity above the firm-level median; the grouping remains fixed.
Small initial size	Initial log assets at or below the firm-level median; the grouping remains fixed.
Non-state ownership	Non-state-owned classification at the start of the firm's retained panel; the grouping remains fixed.
Peer disclosure instrument	Excluded instrument in the first-stage disclosure regression: standardized peer-group data-related disclosure environment in the same industry-year cohort (excluding the focal firm).

3.3 Baseline specification and economic magnitude

Collaboration is a nonnegative count with substantial dispersion and many zeros. We estimate its conditional mean using a log link, following the multiplicative-model logic of Santos Silva and Tenreyro (2006). Let D denote Disclosure and X the vector of seven controls. The full specification is:

$$\mu_{it} = \exp(\alpha_i + \lambda_t + \beta D_{it} + \gamma' X_{it}) \quad (1)$$

In equation (1), μ_{it} is the conditional expected count of interregional collaborative patent applications, α_i denotes firm fixed effects, and λ_t denotes fiscal-year fixed effects. Following Wooldridge (1999) and Santos Silva and Tenreyro (2006), the quasi-maximum likelihood fixed-effects Poisson estimator consistently estimates the conditional mean of nonnegative count outcomes under arbitrary forms of overdispersion, and is robust to data characterized by a high proportion of zeros without requiring the restrictive distributional assumptions of dual-regime zero-inflated models. Standard errors are clustered at the firm level to account for serial correlation within firms over time (MacKinnon et al., 2023).

Table 3 reports three specifications. Column (1) includes disclosure only; column (2) adds the seven firm-level controls; and column (3) adds firm and fiscal-year fixed effects. For a one-standard-deviation increase in disclosure, the percentage change in the conditional expected count is $100[\exp(\beta) - 1]$.

3.4 Robustness and endogeneity checks

We conduct three robustness checks. First, standardized raw disclosure density replaces the primary log-transformed measure. Second, a linear fixed-effects model uses $\ln(1 + \text{Collaboration}_{it})$ as the dependent variable and retains firms with zero collaboration in every year. Third, the baseline model is re-estimated after excluding manufacturing observations. All three specifications include the seven firm-level controls and firm and year fixed effects.

To further address potential reverse causality and omitted time-varying regional or industrial dynamics, we first re-estimate the baseline specification after replacing contemporaneous disclosure with its one-year lag (Bellemare et al., 2017). Second, we augment the baseline model with high-dimensional industry-by-year and province-by-year fixed effects to absorb time-varying industrial technology shocks and regional policy fluctuations (Correia et al., 2020).

Third, to address potential endogeneity arising from unobserved time-varying confounders, we implement a Poisson control-function (CF) approach following Wooldridge (2015). We construct an excluded instrumental variable (Peer disclosure instrument), defined as the standardized peer-group data-related disclosure environment among firms in the same industry-year cohort (excluding the focal firm). Industry-level reporting norms and peer disclosure practices shape a focal firm's reporting decisions through benchmarking and institutional isomorphism, satisfying instrument relevance, whereas peer reporting norms do not directly determine the focal firm's bilateral interregional patent collaborations

except through its own disclosure practices. In the first stage, we estimate a linear fixed-effects regression of Disclosure on the excluded instrument, the seven firm-level controls, and firm and fiscal-year fixed effects.

In the second stage, the first-stage residual is incorporated into the fixed-effects Poisson collaboration model alongside Disclosure, the firm-level controls, centered disclosure-by-subgroup interactions, and firm and year fixed effects to control for potential endogeneity (Wooldridge, 2015). To account for the generated first-stage residual, statistical inference in the control-function and mediation models is based on firm-level block bootstrap standard errors (999 replications).

3.5 Mechanisms and group comparisons

To test the parallel mediating channels of partner-search capability and external financing access, we estimate two linear mechanism equations alongside a joint fixed-effects Poisson outcome equation, all controlling for firm and fiscal-year fixed effects. Let S denote Partner-search capability, F denote Access to external finance, and M^k denote either mechanism. The linear mechanism equations in equation (2) test the effect of Disclosure on each mechanism, while the joint outcome specification in equation (3) incorporates both mechanisms alongside Disclosure, the firm-level controls, the group interactions, and the first-stage residual to evaluate the contribution of each channel conditional on the other:

$$M_{it}^{k} = \alpha_{ik} + \lambda_{tk} + a_k D_{it} + \gamma_k' X_{it} + \varepsilon_{it}^{k}, k \in \{S, F\} \tag{2}$$

$$\ln\mu_{it} = \alpha_i + \lambda_t + cD_{it} + b_S S_{it} + b_F F_{it} + \gamma' X_{it} + \delta' W_{it} + \rho v_{it} \tag{3}$$

Here, W denotes the vector of centered disclosure-by-subgroup interactions, and v denotes the first-stage residual from Section 3.4. Both mechanism indices S and F are measured in fiscal year t concurrently with Disclosure, whereas Collaboration is measured over the subsequent 12-month post-release window.

Because the mediator equations in (2) are linear while the collaboration outcome equation in (3) employs a non-linear log link, standard linear product-of-coefficients logic does not directly map onto linear additive scale changes. Following established mediation frameworks for non-linear count outcomes, the indirect effect through mediator $k \in \{S, F\}$ is captured on the log-conditional mean scale by the product of coefficients $a_k \times b_k$, which corresponds to a multiplicative scale factor of $\exp(a_k \cdot b_k)$—or an expected percentage shift of $100[\exp(a_k \cdot b_k) - 1]\%$ in the conditional count of interregional collaborative patent applications. To conduct valid statistical inference on these indirect transmission pathways without imposing distributional normality on the product term, we estimate firm-level block-bootstrapped standard errors and 95% bias-corrected confidence intervals across 999 replications.

Cross-sectional heterogeneity is evaluated across six subgroup regressions: low versus high initial R&D intensity, small versus large initial firm size, and state-owned versus non-state-owned enterprises. Each subgroup regression employs the

baseline firm fixed-effects Poisson specification, and inter-group coefficient differences are formally tested using fully interacted pooled models with firm-clustered standard errors (alongside Holm-adjusted p-values ; Holm, 1979).

4. Empirical Results

4.1 Descriptive statistics

Table 2 reports summary statistics for the full analytical panel. Collaboration averages 2.415 patent applications per post-release window, with a standard deviation of 9.056 and 55.47% zero observations. This right-skewed distribution with substantial zero-inflation motivates our fixed-effects Poisson count specification. By construction, Disclosure and both composite mechanism indices have full-panel means of zero and standard deviations of one.

Table 2: Descriptive statistics

Variable	Mean	SD	Min.	Median	Max.
Collaboration	2.415	9.056	0.000	0.000	319.000
Disclosure	0.000	1.000	-4.506	0.011	4.044
Raw disclosure density	6.937	6.965	0.113	4.934	144.243
Firm size	22.175	1.119	17.791	22.165	26.060
Leverage	0.428	0.116	0.030	0.428	0.874
Return on assets	0.045	0.049	-0.151	0.045	0.231
Operating cash flow	0.050	0.063	-0.184	0.050	0.320
Revenue growth	0.079	0.292	-0.800	0.079	1.147
Largest shareholder stake	0.339	0.101	0.050	0.340	0.727
Listed age	2.372	0.872	0.000	2.565	3.555
Partner-search capability	0.000	1.000	-3.882	-0.003	3.798
Access to external finance	0.000	1.000	-4.160	0.008	4.636
State ownership	0.387	0.487	0.000	0.000	1.000
Manufacturing	0.645	0.478	0.000	1.000	1.000

Notes: Full analytical panel: 24,250 firm-years and 3,105 firms during 2016–2024. Disclosure and mechanism indices retain their full-panel standardization.

4.2 Baseline results

Table 3 presents the baseline regression results across three nested specifications. The coefficient on Disclosure is 0.2362 ($p < 0.001$) without controls in column (1), 0.1593 ($p < 0.001$) after adding the seven firm-level controls in column (2), and 0.1735 (SE = 0.0074, $p < 0.001$) after controlling for both firm and fiscal-year fixed effects in column (3).

Table 3: Baseline regressions

Variable	(1) Disclosure only	(2) Add controls	(3) Add fixed effects
Disclosure	0.2362*** (0.0418)	0.1593*** (0.0425)	0.1735*** (0.0074)
Firm size	—	0.4251*** (0.0552)	0.2020*** (0.0351)
Leverage	—	0.3188 (0.4449)	-0.2869*** (0.1088)
Return on assets	—	1.0270 (0.8782)	0.4348*** (0.1496)
Operating cash flow	—	-0.0672 (0.8774)	0.4928*** (0.1235)
Revenue growth	—	-0.0882 (0.1130)	-0.0145 (0.0235)
Largest shareholder stake	—	-1.1665** (0.5184)	-0.4076* (0.2114)
Listed age	—	0.0241 (0.0754)	0.0530** (0.0228)
Controls	No	Yes	Yes
Firm effects	No	No	Yes
Year effects	No	No	Yes
Observations	19,353	19,353	19,353
Firms	2,443	2,443	2,443

Notes: All columns use the same 19,353 observations. Dependent variable: Collaboration. Columns (1)–(2) use pooled Poisson; column (3) uses firm fixed-effects Poisson. Firm-clustered standard errors in parentheses. ***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively.

In terms of economic magnitude, the baseline coefficient of 0.1735 in column (3) indicates that a one-standard-deviation increase in data-related disclosure is associated with an 18.95% increase in the conditional expected count of interregional collaborative patent applications ($100[\exp(0.1735) - 1]$). Evaluated at a benchmark of two collaborative filings per year, this effect corresponds to approximately 0.38 additional interregional joint patent applications, providing strong empirical support for H1.

4.3 Robustness checks

Table 4 reports the results of our three primary robustness checks. In column (1), replacing log-transformed disclosure with standardized raw disclosure density yields a positive and highly significant coefficient of 0.1102 (SE = 0.0087, $p < 0.001$). In column (2), estimating a linear fixed-effects specification with ln(1 +

Collaboration) as the dependent variable across the full panel of 24,250 observations (retaining firms with zero collaborative patents throughout the sample period) yields a coefficient of 0.0630 (SE = 0.0035, $p < 0.001$). In column (3), excluding manufacturing firms leaves 6,753 observations across 880 nonmanufacturing firms and produces a coefficient of 0.1532 (SE = 0.0123, $p < 0.001$). Across all three specifications, the positive association between data-related disclosure and interregional collaborative innovation remains robust.

Table 4: Robustness checks

Variable	(1) Alternative X	(2) Alternative Y	(3) Nonmanufacturing
Disclosure measure	0.1102*** (0.0087)	0.0630*** (0.0035)	0.1532*** (0.0123)
Outcome	Count	ln(1 + count)	Count
Estimator	FE Poisson	Linear FE	FE Poisson
Seven controls	Yes	Yes	Yes
Firm/year effects	Yes	Yes	Yes
Observations	19,353	24,250	6,753
Firms	2,443	3,105	880

Notes: Column (1) uses standardized raw density. Columns (2)–(3) use standardized log density. The linear model retains firms with no positive outcomes. Column (3) excludes manufacturing observations. Firm-clustered standard errors in parentheses. ***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively.

4.4 Timing and endogeneity checks

Table 5 presents the timing, high-dimensional fixed-effects, and control-function endogeneity checks. In column (1), replacing contemporaneous disclosure with its one-year lag yields a positive and statistically significant coefficient of 0.0551 (SE = 0.0075, $p < 0.001$) across 16,267 observations from 2,332 firms, confirming that the positive relationship persists over a longer temporal lag. In column (2), absorbing industry-by-year and province-by-year fixed effects yields a disclosure coefficient of 0.1786 (SE = 0.0069, $p < 0.001$) across 19,342 observations, demonstrating that our baseline finding is not driven by time-varying industrial technology trends or provincial policy shocks.

Table 5: Timing and endogeneity checks

Variable	(1) Lagged X	(2) Stricter effects	(3) Control function
Disclosure	—	0.1786*** (0.0069)	0.3361*** (0.1095)
Lagged disclosure	0.0551*** (0.0075)	—	—
First-stage residual	—	—	−0.1506 (0.1088)
First-stage cluster-Wald F	—	—	77.66
Outcome observations	16,267	19,342	19,353
Outcome firms	2,332	2,442	2,443

Notes: All columns include the seven firm-level controls and firm fixed effects. Columns (1) and (3) include fiscal-year fixed effects; column (2) includes industry-by-year and province-by-year fixed effects. Column (3) reports the second-stage Poisson control-function estimates using the excluded peer disclosure instrument (first-stage N = 24,250 across 3,105 firms). Standard errors in columns (1) and (2) are clustered at the firm level; standard errors in column (3) are derived from 999 firm-level bootstrap replications. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Column (3) of Table 5 reports the Poisson control-function estimates. In the first-stage regression, the excluded peer disclosure instrument is positively and significantly associated with focal-firm disclosure (coefficient = 0.0505, SE = 0.0057, $p < 0.001$), with a first-stage F-statistic of 77.66 that far exceeds conventional weak-instrument thresholds. In the second-stage Poisson specification, the coefficient on Disclosure remains positive and statistically significant at 0.3361 (SE = 0.1095, $p = 0.002$), while the first-stage residual coefficient is −0.1506 (SE = 0.1088). Together, these results confirm that the positive effect of data-related disclosure on interregional collaborative innovation remains robust after accounting for potential endogeneity.

4.5 Mechanism analysis

Table 6 presents the empirical results for the two parallel transmission channels: partner-search capability and access to external finance. Columns (1) and (2) report the linear fixed-effects regressions of each mechanism index on Disclosure and the seven controls. Data-related disclosure is positively and significantly associated with both Partner-search capability (coefficient = 0.4088, SE = 0.0078, $p < 0.001$) and Access to external finance (coefficient = 0.3765, SE = 0.0079, $p < 0.001$). Column (3) reports the joint fixed-effects Poisson collaboration specification incorporating both mechanism indices alongside Disclosure, the firm-level controls, the group interactions, and the first-stage residual.

Both Partner-search capability (coefficient = 0.0825, SE = 0.0049, p < 0.001) and Access to external finance (coefficient = 0.0770, SE = 0.0047, p < 0.001) exhibit positive and highly significant effects on interregional collaborative patenting, while the direct coefficient on Disclosure remains positive and significant (0.2337, SE = 0.0983, p < 0.05). In terms of indirect transmission magnitude, the estimated indirect effect of disclosure via partner-search capability on the log conditional mean is 0.0337 ($a_S \times b_S = 0.4088 \times 0.0825$, bootstrap p < 0.001), corresponding to a 3.43% increase in expected interregional collaborative patents ($100[\exp(0.0337) - 1]\%$). The indirect effect via access to external finance on the log conditional mean is 0.0290 ($a_F \times b_F = 0.3765 \times 0.0770$, bootstrap p < 0.001), representing a 2.94% increase in expected collaborative filings ($100[\exp(0.0290) - 1]\%$). Both 95% bias-corrected bootstrap confidence intervals strictly exclude zero. Collectively, these results provide strong empirical support for H2a and H2b, confirming that partner-search capability and external financing access serve as parallel mediating pathways linking data-related disclosure to interregional collaborative innovation.

Table 6: Partner search and external finance

Variable	(1) Partner search	(2) External finance	(3) Collaboration
Disclosure	0.4088*** (0.0078)	0.3765*** (0.0079)	0.2337** (0.0983)
Partner-search capability	—	—	0.0825*** (0.0049)
Access to external finance	—	—	0.0770*** (0.0047)
Estimator	Linear FE	Linear FE	FE Poisson
Seven controls	Yes	Yes	Yes
Firm/year effects	Yes	Yes	Yes
First-stage residual	No	No	Yes
Disclosure × group terms	No	No	All three
Observations	24,250	24,250	19,353
Firms	3,105	3,105	2,443

Notes: Columns (1) and (2) report linear fixed-effects regressions with firm-clustered standard errors in parentheses. Column (3) reports the joint fixed-effects Poisson specification controlling for the first-stage residual and centered disclosure-by-group interactions, with firm-level bootstrap standard errors (999 replications) in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

4.6 Heterogeneity analysis

Table 7 reports the subgroup regressions and inter-group coefficient difference tests across initial R&D intensity, firm size, and ownership structure. In columns (1) and (2), the disclosure coefficient is 0.1466 (SE = 0.0100, $p < 0.001$) for low-R&D firms and 0.2046 (SE = 0.0099, $p < 0.001$) for high-R&D firms, corresponding to conditional expected-count increases of 15.78% and 22.70%, respectively. The inter-group coefficient difference of 0.0580 is statistically significant at the 1% level ($p < 0.001$), supporting H3a and confirming that stronger baseline R&D capability amplifies the collaborative returns to data-related disclosure.

Table 7: Group regressions and coefficient differences

Variable	(1) Low R&D	(2) High R&D	(3) Small	(4) Large	(5) SOE	(6) Non-SOE
Disclosure	0.1466*** (0.0100)	0.2046*** (0.0099)	0.2034*** (0.0116)	0.1603*** (0.0089)	0.1216*** (0.0100)	0.2078*** (0.0090)
Expected-count change	15.78%	22.70%	22.55%	17.39%	12.93%	23.10%
Seven controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm/year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,668	9,685	9,299	10,054	7,436	11,917
Firms	1,225	1,218	1,158	1,285	881	1,562
Difference direction	High − low		Small − large		Non-SOE − SOE	
Coefficient difference	0.0580		0.0430		0.0863	
Difference-test p	<0.001		0.003		<0.001	

Notes: All columns use firm fixed-effects Poisson. Group assignments use initial characteristics and remain fixed. Difference tests use fully interacted pooled models, allowing controls and year effects to differ by group. Expected-count changes equal $100[\exp(\beta) - 1]$ for a one-SD increase in disclosure. Holm-adjusted p values for the three differences are <0.001, 0.003, <0.001, respectively. Firm-clustered standard errors in parentheses. ***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively.

In columns (3) and (4), small and large firms exhibit disclosure coefficients of 0.2034 (SE = 0.0116, $p < 0.001$) and 0.1603 (SE = 0.0089, $p < 0.001$), translating into expected-count increases of 22.55% and 17.39%, respectively, with a significant inter-group difference of 0.0430 ($p = 0.003$). In columns (5) and (6), state-owned enterprises (SOEs) and non-state-owned enterprises (non-SOEs) record disclosure coefficients of 0.1216 (SE = 0.0100, $p < 0.001$) and 0.2078 (SE = 0.0090, $p < 0.001$), corresponding to expected-count increases of 12.93% and 23.10%, respectively, with a significant inter-group difference of 0.0863 ($p < 0.001$). These results strongly support H3b and H3c, and all three cross-sectional differences remain statistically significant after Holm multiplicity adjustment.

5. Conclusions and Implications

5.1 Main findings

Using an unbalanced panel of 3,105 nonfinancial Chinese A-share listed firms (24,250 firm-year observations, 2016–2024), we estimate a firm fixed-effects Poisson model of interregional collaborative patent applications. A one-standard-deviation increase in data-related disclosure is associated with an 18.95% increase in the conditional expected count. The association persists across alternative measures and samples, lagged disclosure, stricter fixed effects, and Poisson control-function estimation using an excluded peer-reporting instrument. The mediation results support partner-search capability and access to external finance as parallel pathways. The association is stronger for firms with higher R&D intensity, smaller firms, and non-SOEs.

5.2 Theoretical and managerial implications

Disclosure helps distant partners assess technological fit and collaborative readiness. It helps financiers evaluate project viability. Internal screening and absorptive capacity help firms act on this visibility. Greater R&D capacity aids technical evaluation (Cohen and Levinthal, 1990). Smaller firms gain visibility, and non-SOEs build credibility where credit and institutional guarantees are limited.

Linking firms' reporting to cross-provincial patent co-development extends research on internal innovation (Huang et al., 2026) and public data openness (Li and Cao, 2026). This complements research on digital spillovers (Hu, 2026), supply-chain resilience (Qiao et al., 2026), and dual mediation (Zhu and Yang, 2026).

To facilitate mutual partner screening, firms should tie disclosures about their data capabilities and technical applications to specific projects and explain which problems they seek to solve with cross-regional partners. They should establish a process for technical staff to review incoming inquiries and assess proposals against internal R&D priorities, while protecting proprietary algorithms, trade secrets, and sensitive customer data (Boot and Vladimirov, 2025; Kao, 2025).

To attract external financing, R&D and finance teams should align disclosures with project milestones, planned capital expenditures, and governance arrangements for joint projects. Explaining how external funds would complement existing data capabilities helps lenders and equity investors assess funding needs, evaluate resource allocations, and monitor the firm's commitments.

5.3 Limitations and future research

Joint patent applications capture realized collaboration. Surveys or contracting records could track earlier outreach, technical consultations, and alliances that yield no patents, helping distinguish initial partner search from contract execution. Project-level data could trace how milestone-based disbursements and joint R&D progress interact over time, adding detail beyond firm-level disclosure measures and mechanism indices. Studies of cross-border R&D alliances could examine how different legal and regulatory regimes shape the relationship between disclosure and collaborative innovation beyond China's interprovincial setting.

Data Availability

The data that support the findings of this study are derived from corporate annual reports, the China Stock Market & Accounting Research (CSMAR) database, and patent records from the China National Intellectual Property Administration (CNIPA). Processed data and replication code are available from the authors upon reasonable request.

ACKNOWLEDGEMENTS. This work was supported by the Key Research Project of the China Commercial Accounting Association (Grant No. 2025zsx004) and the Open Project of the Beijing Urban Governance Research Base (Grant No. 2026CSZL20).

References

[1] Healy, P.M. and Palepu, K.G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. Journal of Accounting and Economics, 31(1–3), pp. 405–440. https://doi.org/10.1016/S0165-4101(01)00018-0

[2] Chu, J., He, Y., Hui, K.W. and Lehavy, R. (2025). New product announcements, innovation disclosure, and future firm performance. Review of Accounting Studies, 30(1), pp. 352–383. https://doi.org/10.1007/s11142-024-09820-0

[3] Huang, L., Zhang, Y., Xian, R. and Chen, X. (2026). Data asset disclosure and corporate innovation: Evidence from China. International Review of Financial Analysis, 109, article 104689. https://doi.org/10.1016/j.irfa.2025.104689

[4] Li, D. and Cao, Y. (2026). Government data openness and firms' interregional collaborative innovation. Technology in Society, 87, article 103391. https://doi.org/10.1016/j.techsoc.2026.103391

[5] Qiao, X., Yan, R. and Lu, X. (2026). Digital finance and supply chain resilience of Chinese manufacturing enterprises: Evidence from A-share listed companies. Advances in Management and Applied Economics, 16(4), pp. 1–27. https://doi.org/10.47260/amae/1641

[6] Hu, X. (2026). Research on the impact of downstream firms' digital transformation on upstream firms' green innovation: A supply-chain perspective. Advances in Management and Applied Economics, 16(2), pp. 107–124. https://doi.org/10.47260/amae/1626

[7] Boschma, R. (2005). Proximity and innovation: A critical assessment. Regional Studies, 39(1), pp. 61–74. https://doi.org/10.1080/0034340052000320887

[8] Rosenkopf, L. and Almeida, P. (2003). Overcoming local search through alliances and mobility. Management Science, 49(6), pp. 751–766. https://doi.org/10.1287/mnsc.49.6.751.16026

[9] Boot, A. and Vladimirov, V. (2025). Disclosure, patenting, and trade secrecy. Journal of Accounting Research, 63(1), pp. 5–56. https://doi.org/10.1111/1475-679X.12580
[10] Chen, H., Liang, P.J. and Petrov, E. (2024). Innovation and financial disclosure. Journal of Accounting Research, 62(3), pp. 935–979. https://doi.org/10.1111/1475-679X.12546
[11] Kao, J. (2025). Information disclosure and competitive dynamics: Evidence from the pharmaceutical industry. Management Science, 71(7), pp. 5948–5970. https://doi.org/10.1287/mnsc.2022.01161
[12] Borchhardt, G., Kovács, B. and Rogan, M. (2025). Competitive overlap as a signal in expert partner choice: Evidence from patent law firm selection. Strategic Management Journal, 46(7), pp. 1557–1605. https://doi.org/10.1002/smj.3700
[13] Cohen, W.M. and Levinthal, D.A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly, 35(1), pp. 128–152. https://doi.org/10.2307/2393553
[14] Hall, B.H. and Lerner, J. (2010). The financing of R&D and innovation. In Handbook of the Economics of Innovation, Hall, B.H. and Rosenberg, N. (Eds.), vol. 1, North-Holland, pp. 609–639. https://doi.org/10.1016/S0169-7218(10)01014-2
[15] Diamond, D.W. and Verrecchia, R.E. (1991). Disclosure, liquidity, and the cost of capital. The Journal of Finance, 46(4), pp. 1325–1359. https://doi.org/10.1111/j.1540-6261.1991.tb04620.x
[16] Santos Silva, J.M.C. and Tenreyro, S. (2006). The log of gravity. The Review of Economics and Statistics, 88(4), pp. 641–658. https://doi.org/10.1162/rest.88.4.641
[17] Wooldridge, J.M. (1999). Distribution-free estimation of some nonlinear panel data models. Journal of Econometrics, 90(1), pp. 77–97. https://doi.org/10.1016/S0304-4076(98)00033-5
[18] MacKinnon, J.G., Nielsen, M.Ø. and Webb, M.D. (2023). Cluster-robust inference: A guide to empirical practice. Journal of Econometrics, 232(2), pp. 272–299. https://doi.org/10.1016/j.jeconom.2022.04.001
[19] Bellemare, M.F., Masaki, T. and Pepinsky, T.B. (2017). Lagged explanatory variables and the estimation of causal effect. The Journal of Politics, 79(3), pp. 949–963. https://doi.org/10.1086/690946
[20] Correia, S., Guimarães, P. and Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. The Stata Journal, 20(1), pp. 95–115. https://doi.org/10.1177/1536867X20909691
[21] Wooldridge, J.M. (2015). Control function methods in applied econometrics. Journal of Human Resources, 50(2), pp. 420–445. https://doi.org/10.3368/jhr.50.2.420

[22] Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics, 6(2), pp. 65–70. https://www.jstor.org/stable/4615733
[23] Zhu, N. and Yang, Z. (2026). Digital goods trade and national innovation: A dual-mediation perspective. Advances in Management and Applied Economics, 16(3), pp. 65–85. https://doi.org/10.47260/amae/1634