

Measuring Organizational R&D Capability and External Knowledge Integration: A Longitudinal Scientometric Study of Aviation Higher Education Institutions

Arthur Dela Peña

artair248@gmail.com

National Aviation Academy of the Philippines

Jefferson Clariza

National Aviation Academy of the Philippines

Noel Florencondia

Nueva Ecija University of Science and Technology

Research Article

Keywords: absorptive capacity, aviation higher education, bibliometrics, external knowledge integration, longitudinal panel, OpenAlex

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11137586/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Measuring Organizational R&D Capability and External Knowledge Integration: A Longitudinal Scientometric Study of Aviation Higher Education Institutions

Arthur Dela Peña^{1*}, Jefferson Clariza¹, Noel Florencondia²

¹National Aviation Academy of the Philippines, Pampanga, Philippines

²Nueva Ecija University of Science and Technology, Nueva Ecija, Philippines

Corresponding author email: atair248@gmail.com

Abstract

Aviation-focused higher education institutions connect specialized research with geographically and organizationally distributed knowledge, yet longitudinal evidence concerning this boundary-spanning function remains limited. This study examined whether prior revealed organizational R&D capability (RORC) was associated with subsequent external knowledge integration (EKII) in an international purposive panel of 40 aviation-focused institutions from 2015–2025. OpenAlex metadata, Research Organization Registry lineages, and the frozen AVI-TAX-v1.0 protocol yielded 49,274 strict institution–work records, 62,624 broad records, 35,010 strict collaboration edges, and a 440-row institution-year skeleton; 38 institutions met the primary eligibility criteria. A final denominator audit corrected AVI038 total-output counts using approved OpenAlex lineage I107428990, enforced $\text{Total_Institutional_Works} \geq \text{Aviation_Works}$, rebuilt RORC, and reran all analyses. Two-way fixed-effects models estimated EKII in year s from RORC and institutional output in $s-1$, contemporaneous national R&D conditions, and institution and year effects. The strict primary model included 320 observations from 38 institutions. RORC and EKII were positively correlated in pooled observations ($r = .327$), but the within-institution estimate was positive and imprecise in the strict corpus ($\beta = .098$, 95% CI $[-.074, .270]$, wild-cluster $p = .230$) and smaller in the broad corpus ($\beta = .041$, 95% CI $[-.090, .173]$, $p = .510$). Exploratory cross-sector linkage estimates were also imprecise. Patent, institutional-report, and independent validation evidence were excluded. The study contributes a lineage-aware longitudinal scientometric dataset and measurement framework while demonstrating that pooled institutional relationships among bibliometric indicators should not be interpreted automatically as within-institution change.

Keywords: absorptive capacity; aviation higher education; bibliometrics; external knowledge integration; longitudinal panel; OpenAlex

1. Introduction

Aviation higher education institutions (HEIs) occupy an unusually demanding boundary between academic research and operational practice. Their research connects universities with manufacturers, airlines, airports, maintenance organizations, public laboratories, regulators, and regional innovation systems. The knowledge produced in this environment is multidisciplinary, infrastructure-intensive, and safety-critical. Technical promise alone is rarely sufficient: evidence must often withstand system integration, operational testing, explainability, certification, and extended development cycles before it becomes usable. Data-driven aerospace engineering illustrates this challenge because apparently high-performing methods must also be generalizable, interpretable, and compatible with stringent engineering constraints (Brunton et al., 2021). External relationships can provide complementary expertise and implementation settings, but their value depends on communication, trust, governance, and relational fit (Rybnicek & Königsgruber, 2019).

This setting creates a central organizational puzzle. Some aviation HEIs sustain large and diverse research portfolios yet remain weakly connected to external research systems, whereas others become deeply embedded in international, interinstitutional, or cross-sector networks. Publication counts alone cannot distinguish accumulated research capability from the processes through which organizations recognize and integrate external knowledge. This distinction also presents a scientometric measurement problem. Institutional research performance is commonly represented through publication volume, collaboration intensity, internationalization, or network indicators, yet these measures capture different dimensions of scientific activity and may behave differently across institutions and over time. University collaboration has heterogeneous associations with research quality and technology transfer (Sun et al., 2020), while societal influence typically develops through sequences of interaction rather than from scholarly output alone (Muhonen et al., 2020). Cross-sectional associations among such indicators may therefore reflect persistent institutional heterogeneity rather than longitudinal organizational change. Distinguishing productive research capability from external knowledge integration provides both a substantive organizational test and a methodological test of how institutional bibliometric indicators should be

interpreted. A longitudinal design is consequently needed to separate stable differences between institutions from changes occurring within the same institution over time.

Absorptive-capacity theory provides a useful explanation because prior related knowledge should help organizations recognize, assimilate, transform, and exploit external knowledge (Cohen & Levinthal, 1990; Zahra & George, 2002). Yet three limitations constrain its application to aviation HEIs. First, much evidence comes from firms or cross-sectional surveys; institution-level, historically reconstructed measures remain uncommon. Second, studies frequently collapse research inputs, collaboration, routines, and innovation outcomes into a single score, despite warnings against reifying absorptive capacity (Lane et al., 2006). Third, bibliometric studies often compare organizations cross-sectionally even when the theoretical mechanism concerns change. These limitations are material because potential and realized dimensions of absorptive capacity may have different antecedents and consequences (Algarni et al., 2023; Sancho-Zamora et al., 2021).

Existing higher-education evidence also leaves a design gap. Survey research has linked absorptive capacity to innovation in HEIs through internal coordination, but perceptual, single-period measures cannot reconstruct institutional trajectories or separate respondent assessment from observable behavior (Asiedu et al., 2023). A systematic review of absorptive capacity in technology transfer likewise found persistent contextual and methodological fragmentation (da Silva Florêncio & de Oliveira Júnior, 2022). The problem is not that surveys are intrinsically weak; they reveal routines and perceptions that bibliometric records cannot. Rather, the field lacks complementary archival designs that can test whether changes in an institution's revealed knowledge base precede changes in its external knowledge relations. Aviation HEIs provide a focused setting for that complement because their organizational boundaries and research applications are comparatively identifiable.

The present study addressed this gap through a transparent longitudinal computational analysis. Revealed Organizational R&D Capability (RORC) represented the observable scale, continuity, specialization, and topical breadth of an institution's aviation research base. The External Knowledge Integration Index (EKII) represented the breadth and intensity of research relationships with external higher-education and research organizations. The primary question was whether RORC in year $s-1$ was positively associated with EKII in year s after accounting for institution and year fixed effects, prior institutional output, and national R&D conditions. Exploratory questions examined annual construct trajectories and associations with a deliberately limited cross-sector research-linkage index based on company and government coauthorship. This index was not interpreted as research translation because patent-family, licensing, implementation, and institutional-report evidence were not included in the completed evidence base.

The study makes four bounded contributions. Theoretically, it separates a revealed prior knowledge base from observable external integration rather than treating both as a single measure of absorptive capacity. Empirically, it provides an international purposive 2015–2025 aviation-HEI panel with work-level provenance and collaboration edges. Methodologically, the study contributes to longitudinal institutional scientometrics by separating capability and integration indicators, preserving organizational lineage and work-level provenance, and testing whether pooled bibliometric relationships remain evident under within-institution fixed-effects estimation. The workflow further incorporates strict and broad aviation classifications, explicit blank-versus-zero rules, construction buffers, deduplication, denominator-invariant checks, and reconciliation audits. Substantively, it shows that the positive pooled capability–integration relationship was not reproduced as a precisely estimated within-institution association. This contrast is consistent with an important role for stable institutional differences, although it does not by itself estimate or identify a formal between-institution effect.

2. Theoretical Background and Hypothesis

2.1 Aviation HEIs as Boundary Organizations

Aviation HEIs were defined as legally recognized universities, academies, and institutes with sustained teaching, research, or institutional activity in aviation, aeronautics, aerospace systems, air transportation, flight operations, airports, safety, unmanned aircraft, human factors, or aviation regulation. Their analytical importance derives from their boundary position. Universities can be knowledge producers, conveners, and institutional catalysts

within innovation ecosystems, but these functions depend on connections to other actors and extend beyond publication output alone (Cai et al., 2020; Compagnucci & Spigarelli, 2020). Research on university embeddedness similarly emphasizes spatial, organizational, and relational ties as conditions for useful knowledge flow (Taxt et al., 2022).

The aerospace context intensifies this dependence. Comparative ecosystem research identifies public R&D support, specialized skills, infrastructure, supporting organizations, and strategic alliances as mutually reinforcing enablers (Luna-Andrade et al., 2021). Civil aerospace innovation systems may remain incomplete even when technically strong organizations are present because missing capabilities or unstable relationships prevent knowledge from being combined (Hickie & Hickie, 2025). An aviation HEI can therefore maintain high scientific productivity without achieving broad external integration. Conversely, an institution with a smaller internal base may become highly connected through deliberate boundary-spanning arrangements. Treating scale and integration as separate constructs is essential.

2.2 Absorptive Capacity and Organizational Change

Absorptive capacity originally described the ability to recognize the value of external knowledge, assimilate it, and apply it (Cohen & Levinthal, 1990). Zahra and George (2002) later distinguished acquisition and assimilation—potential absorptive capacity—from transformation and exploitation—realized absorptive capacity. The distinction explains why access does not guarantee use. More recent evidence links absorptive capacity to strategic flexibility and business-model innovation (Miroshnychenko et al., 2021), while multidimensional studies show that its components have non-equivalent relationships with innovation behavior (Algarni et al., 2023; Sancho-Zamora et al., 2021). A large meta-analysis further suggests that context conditions these relationships and that absorptive capacity may be more strongly related to commercialization than invention (Stettler et al., 2025).

For archival research, the theoretical challenge is to avoid treating any observable proxy as the construct itself. R&D output, collaboration, and downstream outcomes are traces of different stages in a process. Lane et al. (2006) argued that indiscriminate operationalization weakens cumulative explanation by reifying absorptive capacity. The present study therefore treated RORC as revealed prior related knowledge and productive capacity, and EKII as observable boundary-spanning integration. Neither index was equated with the entire absorptive-capacity process. The theory would be supported if within-institution increases in prior RORC systematically preceded greater external knowledge integration, but the design could not directly observe internal recognition, assimilation, or transformation routines.

The distinction also changes what constitutes a theoretically useful negative result. If high-RORC institutions are more integrated only because both features reflect stable historical advantages, a positive pooled association offers limited guidance to organizations seeking change. Conversely, a positive fixed-effects coefficient would indicate that deviations from an institution's own capability level were followed by deviations in integration, a closer match to the proposed process. Failure to estimate that coefficient precisely narrows the claim: it does not falsify all absorptive-capacity mechanisms, but it does challenge a simple, near-term, monotonic translation from bibliometrically revealed capability to external integration. The relevant enabling conditions may reside in governance, incentives, boundary-spanning roles, or alliance experience rather than in research scale alone.

This emphasis on change also matters statistically. Organizations with historically large laboratories, mature doctoral programs, or favorable national systems may display high values on both capability and integration. A pooled correlation can therefore arise from stable institutional sorting, even when changes within an institution are weakly related. Longitudinal absorptive-capacity research has demonstrated the value of tracing capabilities over time rather than inferring dynamics from a single cross-section (Awwad et al., 2025). Two-way fixed effects were accordingly used to focus inference on deviations from each institution's own average while accounting for common year shocks.

2.3 Revealed Organizational R&D Capability

RORC was defined as an institution's sustained and observable capacity to produce aviation-related research through an active researcher base, continuity, strategic specialization, and topical breadth. The adjective revealed signaled an inferential boundary: internal routines, budgets, laboratory quality, and managerial processes were not directly observed. Knowledge-based dynamic-capability research remains conceptually fragmented, making explicit construct boundaries particularly important (Kaur, 2023). Cross-country research nevertheless distinguishes knowledge creation, diffusion, absorption, and impact as related but separable capabilities within innovation ecosystems (Robertson et al., 2023).

Five indicators represented complementary foundations. Aviation output captured accumulated production; unique aviation researchers represented the active human knowledge base; prior-year author continuity approximated organizational memory; aviation research intensity measured specialization relative to total institutional output; and normalized topic breadth represented recombination opportunities across technical and operational domains. Breadth was not assumed to be unconditionally beneficial. Persistent R&D may build learning advantages, whereas excessive diversity can dilute attention or coordination (Ipinnaie et al., 2026). Equal weighting therefore served as a transparent baseline and was challenged through leave-one-indicator-out analyses rather than defended as a uniquely correct latent measurement model.

2.4 External Knowledge Integration

EKII represented the breadth and intensity with which an aviation HEI connected its research to external higher-education and research organizations. Its components captured partner diversity, country diversity, the share of works with an external academic or research partner, and the share involving international academic or research collaboration. These traces were interpreted as observable opportunities for knowledge acquisition and combination, not proof that substantive assimilation occurred. University–industry collaboration encompasses partner identification, governance, exchange, and output realization, while academic engagement also occurs through channels beyond coauthorship (Passos et al., 2023; Perkmann et al., 2021).

Network research supports both the opportunity and constraint sides of integration. International ties and brokerage can improve access to nonredundant knowledge, while dense networks may eventually impose redundancy or coordination costs (Iino et al., 2021). Patent-network evidence similarly shows that the structure and evolution of interorganizational collaboration shape innovation opportunities (Du et al., 2024). Institutions with stronger prior RORC should be better able to signal competence, evaluate potential partners, support complex projects, and bear the fixed costs of international collaboration. Yet these benefits may materialize slowly, and administrative or strategic choices can interrupt the expected pathway.

2.5 Hypothesis and Exploratory Analytical Framework

Absorptive capacity theory posits that prior related knowledge enables organizations to recognize, assimilate, and exploit external knowledge effectively (Cohen & Levinthal, 1990; Zahra & George, 2002). Applied to aviation-focused higher education institutions (HEIs), a broader, more continuous, and specialized research base should strengthen the capacity to establish and sustain relationships with external knowledge sources. The longitudinal design therefore focused on within-institution change rather than cross-sectional differences in average capability. H1 states that, within an aviation-focused HEI, higher revealed organizational R&D capability in year $s-1$ is positively associated with external knowledge integration in year s , conditional on prior institutional output, contemporaneous national R&D conditions, and institution and year fixed effects.

To formalize this primary relationship, the core econometric specification is defined as

$$EKII_{i,s} = \beta_1 RORC_{i,s-1} + \gamma' X_{i,s-1} + \alpha_i + \lambda_s + \varepsilon_{i,s} \quad (1)$$

where $EKII_{i,s}$ represents external knowledge integration for institution i in year s ; $RORC_{i,s-1}$ is revealed organizational R&D capability one year earlier; X contains the time-varying controls; α_i and λ_s denote institution and year fixed effects; and $\varepsilon_{i,s}$ is the idiosyncratic error term.

In addition to testing H1, the study examines two exploratory linkage questions. First, it assesses whether prior RORC is associated with subsequent cross-sector research linkage (CSLI). Second, it evaluates whether lagged EKII is associated with subsequent CSLI after conditioning on prior RORC. These specifications are descriptive diagnostics and are not interpreted as causal mediation or research translation.

$$CSLI_{i,s+1} = \delta_1 RORC_{i,s-1} + \delta_2 EKII_{i,s} + \theta' X_{i,s-1} + \alpha_i + \lambda_s + u_{i,s+1} \quad (2)$$

As illustrated in Figure 1, the solid arrow represents the primary one-year-lagged RORC–EKII hypothesis, whereas the dashed arrows depict these exploratory cross-sector linkage diagnostics.

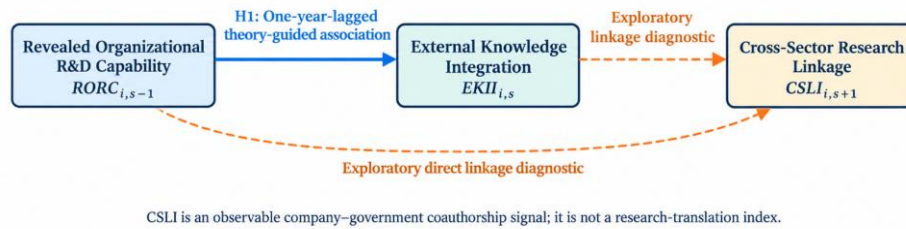


Figure 1. Analytical framework detailing primary and exploratory paths

Note. The solid arrow indicates the primary one-year-lagged RORC–EKII hypothesis (H1). Dashed arrows represent exploratory cross-sector research linkage (CSLI) diagnostics. CSLI measures observable company and government coauthorship signals and does not constitute a direct measure of downstream research translation or societal impact.

These exploratory specifications serve a descriptive, theory-guided diagnostic purpose. They are deliberately not interpreted as formal tests of research translation, causal mediation, technology adoption, or downstream economic impact, given that *CSLI* captures observable coauthorship affiliations with industry and government partners rather than commercial implementation or downstream societal adoption.

3. Methodology

3.1 Research Design and Analytical Framework

This study employed a quantitative, longitudinal computational design with the institution-year as the unit of analysis, drawing on public bibliographic and institutional metadata. The nonexperimental, explanatory framework interprets coefficients as temporally ordered associations rather than causal impacts. Institution fixed effects absorbed time-invariant differences—such as longstanding mission, location, and baseline infrastructure—while year fixed effects controlled for macro-level temporal shocks (Bell et al., 2019). Dataset assembly, dated extractions, and reproducible transformations adhered to established bibliometric standards (Donthu et al., 2021; Weston et al., 2019).

In the archived analysis plan, the RORC–EKII relationship appeared as an intermediate path within a proposed Research Translation Index (RTI) sequence and was labeled H2. Once incomplete RTI, patent, and institutional-report paths were removed from the empirical claims, this capability–integration relationship became the sole theory-guided primary hypothesis and was renumbered H1 in the revised manuscript. This renumbering did not convert the analysis into a preregistered confirmatory test; legacy labels were retained only in the protocol-deviation register and explicitly mapped to the revised terminology in the supplementary reproducibility note. H1 remains the primary theory-driven test supported by the completed OpenAlex panel, whereas *CSLI* specifications

are strictly exploratory. Transparently documenting these scope revisions aligns with best practices for secondary-data research disclosure (van den Akker et al., 2021).

3.2 Sampling Frame and Observation Period

The study used a purposive international frame of 40 legally recognized higher education institutions with identifiable aviation, aeronautical, aerospace, air-transport, or closely related teaching and research activity. The frame was assembled before outcome analysis using documented institutional-program and research evidence. It was intended to provide theoretically relevant institutional variation rather than a probability sample of all aviation-focused HEIs. Supplementary Table S2 reports the discovery source, screening date, country, qualifying aviation activity, institutional identifier, lineage decision, and inclusion rationale for every institution. Accordingly, results were interpreted within the sampled frame and were not treated as population-level estimates for aviation higher education globally.

Thirty-eight institutions were connected to canonical OpenAlex identifiers and Research Organization Registry lineages. National Aviation Academy of the Philippines/former PhilSCA and PATTS College of Aeronautics required documented raw-affiliation fallback searches because available canonical identifiers did not adequately represent the target institutions. The panel skeleton covered 2015–2025, producing 440 possible institution-year rows. Records from 2013–2014 were used only as construction buffers for lagged and author-continuity measures.

A research-active year contained at least one record satisfying the frozen aviation classification. Primary eligibility required at least seven active aviation-research years and at least 20 strict institution–work records over 2015–2025. These rules were frozen in the study control files before outcome analysis and model estimation and were intended to ensure sufficient longitudinal observability for within-institution analysis rather than to classify institutional quality, research performance, or substantive importance. Thirty-eight institutions met the thresholds. The complete 40-institution frame was retained for descriptive reporting; however, the two below-threshold fallback institutions did not provide sufficient like-for-like construct coverage for a complete-frame fixed-effects re-estimation. Alternative eligibility-threshold models were not completed and are therefore not presented as sensitivity evidence. Verified zero-output years were coded zero, whereas unresolved evidence remained missing.

3.3 Data Sources, Extraction, and Deduplication

OpenAlex supplied work identifiers, publication years and types, author identifiers, affiliation strings, institutional identifiers, topics, abstracts when available, citation counts, countries, and collaborating organizations. A frozen Research Organization Registry–linked lineage file harmonized canonical names, aliases, campuses, predecessor–successor relationships, and approved organizational variants. Retrieval used 38 canonical institutional lineages and two documented raw-affiliation fallbacks. Because OpenAlex coverage can change over time and vary geographically, all retrievals were date-stamped, cached responses were reused, query provenance was preserved, and canonical and fallback searches were logged separately (Culbert et al., 2025; Haunschild & Bornmann, 2024; Zheng et al., 2025). Figure 2 summarizes the complete workflow from sampling and lineage-aware retrieval through classification, panel construction, computational auditing, and longitudinal estimation.

The retrieval stream contained 80,405 institution–work rows. Deduplication by the composite key of institution identifier and OpenAlex Work ID removed 6,199 repeated stream rows, leaving 74,206 unique candidate institution–work records for AVI-TAX classification. OpenAlex Work ID, rather than DOI, served as the record-identity key because distinct OpenAlex records can share a DOI. In the strict corpus, 34 institution–DOI groups, representing 31 unique DOI values, contained more than one distinct Work ID; all such Work IDs were retained. A work was also retained once for each sampled institution when its authors reported affiliations with multiple institutions in the study frame. Table 1 summarizes the resulting sample formation, corpus sizes, eligibility counts, and data-quality controls.

Application of AVI-TAX-v1.0 produced 49,274 strict institution–work records representing 46,382 distinct Work IDs and 62,624 broad records representing 59,114 distinct Work IDs. The strict collaboration dataset contained 35,010 unique institution–year–partner edges. During final denominator quality control, AVI038 (Istituto

Tecnológico de Aeronáutica) was re-linked to the approved OpenAlex lineage I107428990 and its 2015–2025 all-work counts were re-retrieved under the same annual scope, yielding 467–747 total works per year. An absent or failed denominator was treated as missing rather than as a verified zero, and the pipeline enforced $\text{Total_Institutional_Works} \geq \text{Aviation_Works}$ in both corpora before aviation intensity and RORC were rebuilt. Automated checks then confirmed that the final institution–work, collaboration-edge, and institution-year keys contained no duplicates and that work-level aggregations reconciled with the 440-row panel skeleton. The lineage-complete extraction also yielded 2,606 TU Delft institution–work records, superseding the earlier 2,482-record pilot count.

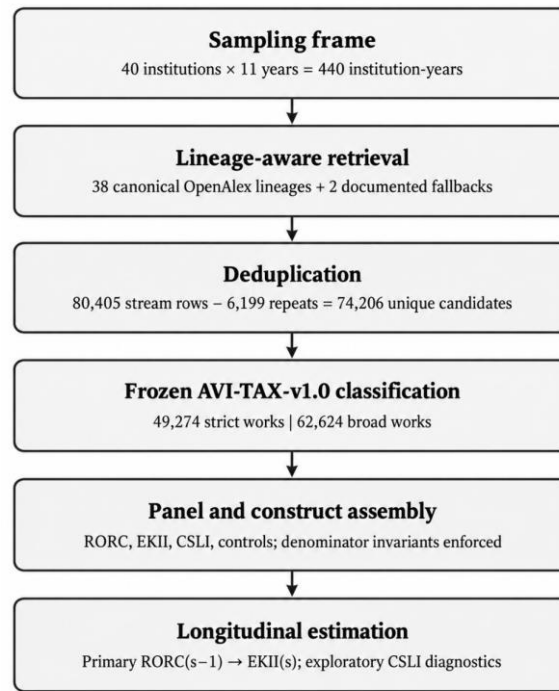


Figure 2. Reproducible data-construction and analysis workflow

Note. Counts refer to the corrected computational materials. The 74,206 records are unique institution–Work ID candidates after stream deduplication. No classification-concordance statistic is reported because neither an independent human validation file nor a provenance-complete archived AI-coded audit file was available for the completed analysis.

Table 1. Sample formation and data-quality controls

Component	Observed result	Interpretation
Panel skeleton	40 institutions × 11 years = 440 rows	Construct availability varied
Retrieved stream rows	80,405	Before institution–Work ID deduplication
Repeated stream rows removed	6,199	Deduplicated by institution × Work ID
Unique candidate records	74,206	Entered AVI-TAX classification
Strict corpus	49,274 institution–work records; 46,382 distinct Work IDs	Primary specification
Broad corpus	62,624 institution–work records; 59,114 distinct Work IDs	Boundary sensitivity
Eligible institutions	38	418 possible descriptive rows
Primary strict model	320 observations; 38 institutions	Complete aligned primary sample

3.4 Aviation Classification and Validation Status

AVI-TAX-v1.0 classified candidate works using a frozen set of direct aviation topics, conditional adjacent topics, keywords, and exclusion rules. Inclusion required aviation or aerospace to constitute a substantive research object, technology, operational context, or application; lexical matches alone were insufficient when the available title, abstract, or topic evidence indicated a non-aviation meaning. Each classification retained its matched rule, evidence field, publication type, taxonomy version, exclusion reason, and source provenance. The strict corpus served as the primary specification, while the broad corpus incorporated defensible adjacent aerospace, transport, and operational applications to assess classification-boundary sensitivity.

A stratified 600-record audit sample was prepared across automated decision, confidence, period, publication type, and institution. However, no independent human coding was completed, and no provenance-complete archived AI-coded audit file was available. Because defensible reliability assessment requires independent coding decisions and transparent procedures for evaluating disagreement (O'Connor & Joffe, 2020), no intercoder-reliability, criterion-validity, or human–AI concordance statistics were reported as completed results. The blinded sample and frozen codebook were retained for future subject-matter validation.

3.5 Constructs and Measures

Table 2 defines the completed constructs. Count variables were transformed as $\log(1 + x)$, whereas proportions and normalized diversity remained on their natural 0–1 scales. Components were standardized as pooled z-scores among eligible observations, and each index was calculated as the arithmetic mean of its complete components. Equal weighting was selected for transparency and reproducibility, not because all components were assumed to be psychometrically interchangeable. The indices were designed as transparent formative scientometric indicators for longitudinal institutional analysis rather than as latent scales or composite rankings of institutional quality. Leave-one-indicator-out analyses assessed dependence on index composition because weighting and aggregation choices can materially affect composite indices (Greco et al., 2019).

Table 2. Operationalization of completed constructs

Construct	Components	Interpretation	Availability
RORC	Aviation works; unique aviation researchers; prior-year author continuity; aviation R&D intensity; normalized topic breadth	Revealed prior knowledge base and productive capability; not a direct measure of internal routines	394/440 strict rows; 422/440 broad rows
EKII	External academic/research partner count; partner-country count; external academic/research collaboration rate; international academic/research collaboration rate	Observable breadth and intensity of research-sector integration	440/440 rows
CSLI	Company coauthorship rate; government collaboration rate	Exploratory cross-sector research-linkage signal; not research translation	440/440 rows
Controls	$\log(1 + \text{total institutional output})$ at RORC exposure year; national GERD (% GDP) at EKII observation year	Time-varying institutional scale and national R&D environment	Complete-case model basis

Note. RORC = Revealed Organizational R&D Capability; EKII = External Knowledge Integration Index; CSLI = Cross-Sector Research Linkage Index; GERD = gross domestic expenditure on R&D. Missing RORC values reflect construction requirements, including lagged continuity requirements, rather than verified zero values. Availability counts in this table refer to the complete 440-row panel; Section 4.1 reports corresponding availability within the 418-row primary-eligible frame.

RORC's output and researcher indicators were logarithmically transformed before standardization. Continuity was the share of current-year aviation authors also observed in the prior year. Aviation intensity was aviation works divided by all institutional works; a missing denominator remained missing, and every observed denominator was required to satisfy $\text{Total_Institutional_Works} \geq \text{Aviation_Works}$. Topic breadth was normalized Shannon diversity across frozen aviation domains. EKII partner counts included external higher-education and research-institute organizations, while its rate components used aviation works as the denominator. CSLI averaged standardized company-coauthorship and government-collaboration rates. It was deliberately named *linkage* rather than *translation* because it contained no patent, licensing, implementation,

standards, curriculum, or verified operational evidence. Detailed formulas and the machine-readable variable dictionary appear in Supplementary Tables S3 and S4.

National R&D conditions were measured using gross domestic expenditure on R&D (GERD) as a percentage of GDP from the World Bank World Development Indicators (indicator GB.XPD.RSDV.GD.ZS; underlying source: UNESCO Institute for Statistics). Observed country-year values were retained directly. Internal gaps of no more than two consecutive years bounded by observed values were filled by linear interpolation, whereas terminal gaps of no more than two years were filled by carrying forward the most recent observed value; longer unresolved gaps remained missing. Each country-year was provenance-flagged as observed, interpolated, carried forward, or missing. Within the 418 institution-years in the primary-eligible frame, 318 (76.1%) used observed GERD values, 12 (2.9%) used interpolated values, 56 (13.4%) used values carried forward for no more than two years, and 32 (7.7%) remained missing. In the final strict primary model, 64 of 320 observations (20.0%) used a filled GERD value—8 interpolated and 56 carried forward—while the remaining 256 observations used directly observed GERD values.

3.6 Statistical Analysis

Analysis began with sample-flow verification, missingness, distributions, annual trends, pooled correlations, and decomposition of within- and between-institution variation. Let s denote the EKII observation year. H1 was estimated using a two-way fixed-effects specification in which RORC and log total institutional output were measured in $s-1$, national GERD was aligned with year s , institution fixed effects absorbed time-invariant institutional characteristics, and EKII-year fixed effects controlled for common annual shocks:

$$\text{EKII}_{i,s} = \beta \text{RORC}_{i,s-1} + \gamma_1 \ln(1 + \text{Output}_{i,s-1}) + \gamma_2 \text{GERD}_{c,s} + \alpha_i + \lambda_s + \epsilon_{i,s} \quad (3)$$

The one-year alignment established temporal ordering but did not identify a causal effect. The individual RORC and EKII components were first standardized during index construction as described in Section 3.5. For coefficient comparability, the completed composite scores and continuous controls were then re-standardized within each model-specific aligned complete-case sample before estimation. Accordingly, reported β coefficients represent standardized within-sample associations and can be compared across the strict and broad specifications. Confidence intervals were calculated using institution-clustered CR1 standard errors. For H1, the p -value was obtained from a null-imposed wild-cluster bootstrap using Rademacher weights, 999 replications, and random seed 20260917; exploratory coefficient rows used the same procedure with deterministic analysis-specific seed offsets. Because the confidence intervals and p -values were derived using different small-cluster procedures, both are identified explicitly in the results tables.

For the exploratory sequential diagnostics, CSLI in year t was aligned with EKII in $t-1$ and RORC and institutional output in $t-2$; national GERD was aligned with the EKII year. Models estimated the RORC–CSLI association and the EKII–CSLI association conditional on prior RORC. For the exploratory sequential diagnostics, CSLI in year t was aligned with EKII in $t-1$ and RORC and institutional output in $t-2$; national GERD was aligned with the EKII year. Models estimated the RORC–CSLI association and the EKII–CSLI association conditional on prior RORC. The descriptive coefficient-product interval was obtained from 2,000 institution-history bootstrap replications in which contributing institutions were sampled with replacement, the complete aligned longitudinal history of each sampled institution was retained, and both constituent regressions—RORC→EKII and EKII→CSLI conditional on prior RORC—were re-estimated in each replication before their coefficients were multiplied (random seed = 20261017). The coefficient product was treated only as a summary of sequential associations and was not interpreted as an indirect causal effect or mediation estimate.

3.7 Transparency, Protocol Deviations, and Ethics

The sampling frame, eligibility rules, taxonomy, construct definitions, and lag logic were frozen in dated control files during data assembly, but the study had no public preregistration record. The revised manuscript therefore

distinguishes a theory-guided primary analysis from exploratory analyses and does not use the term confirmatory for the unavailable RTI sequence. Table 3 documents departures from the earlier analytical plan and their treatment in the revised study. The expanded protocol-deviation register is provided in Supplementary Table S5, and the contents and completion status of the reproducibility archive are summarized in Supplementary Table S1.

Table 3. Protocol-deviation and claim-alignment register

Earlier plan or analytical component	Completed status	Treatment in revised study
Two independent human coders and adjudication	Not completed	No independent human validation or intercoder-reliability claim was made; the 600-record sample and codebook were retained only for future validation.
Authenticated aviation patent-family coverage	Not completed; 0 of 440 institution-year searches were authenticated	Patent evidence was excluded from all constructs, models, and substantive claims; unresolved cells remained missing.
Institutional-report event coding	Report ledger prepared, but events were not systematically coded	Institutional-report evidence was excluded from the constructs and results.
Research Translation Index and mediation sequence	Not constructible from the completed evidence	RTI and mediation claims were removed from the title, hypotheses, models, results, and conclusions.
Capability–integration path labeled H2 within the planned RTI sequence	RTI sequence removed; the completed RORC–EKII path was retained	The capability–integration path was renumbered H1; its legacy H2 mapping was retained in the supplementary reproducibility crosswalk.
Legacy confirmatory analysis script and Holm-adjustment fields	Legacy script retained as historical protocol artifact; manuscript-aligned analysis created for this revision	A manuscript-aligned longitudinal-analysis script was created and rerun after denominator correction; reported estimates and checksums were regenerated from that workflow.
Formal ethics determination	No institutional determination was archived	No approval, exemption, or non-human-participant determination was claimed; the public, aggregate-data scope was disclosed.
OpenAlex panel and RORC, EKII, and CSLI constructs	Completed for bibliometric extraction and corrected construct/model workflow	Retained as the evidence base for the primary capability–integration analysis and exploratory linkage diagnostics.

Note. RORC = Revealed Organizational R&D Capability; EKII = External Knowledge Integration Index; CSLI = Cross-Sector Research Linkage Index; RTI = Research Translation Index.

The reproducibility workflow retained extraction and lineage provenance, strict and broad classified work corpora, collaboration edges, deduplication and reconciliation audits, corrected institution-year and construct panels, a manuscript-aligned longitudinal-analysis script, model and sensitivity outputs, the AVI038 denominator-correction log, and SHA-256 checksums. These materials support reproduction of the reported analyses from the classified-corpus stage onward; exact re-execution of the original OpenAlex retrieval requires the source API and archived query provenance and is not represented as fully self-contained in the submission archive.

4. Results

4.1 Sample Formation, Descriptive Coverage, and Annual Trends

The panel-construction procedure generated a complete skeleton of 440 institution-year observations for 40 institutions over 2015–2025. A skeleton row represented an intended institution-year observation and did not imply that every construct was observed. Thirty-eight institutions met the primary eligibility criteria, yielding 418 possible eligible descriptive observations. After correction of the AVI038 denominator, strict RORC was available for 390 eligible observations and broad RORC for all 418 eligible observations. Temporal alignment, GERD availability, and complete-case requirements yielded 320 observations from 38 institutions in the primary strict model and 348 observations from 38 institutions in the corresponding broad model.

OpenAlex retrieval was completed for 38 canonical institutional lineages and two documented affiliation-based fallback searches. Deduplication and reconciliation identified no duplicate final keys or discrepancies between work-level records and institution-year counts. The strict corpus contained 49,274 institution–work records representing 46,382 distinct OpenAlex Work IDs; the broad corpus contained 62,624 institution–work records representing 59,114 distinct Work IDs. Final denominator checks also confirmed `Total_Institutional_Works ≥`

Aviation_Works for every observed institution-year in both corpora. These checks established computational completeness for the specified bibliometric extraction and panel-construction scope, not for patent, institutional-report, research-translation, or independent validation evidence. The complete extraction, duplication, denominator, and reconciliation checklist is reported in Supplementary Table S9.

Figure 3 presents the annual descriptive patterns. Mean EKII rose from substantially below the panel-wide reference level in the early years to its highest observed value in 2025. Mean RORC followed a more moderate and uneven trajectory, remaining near or below the reference level through 2021, increasing in 2022–2023, declining in 2024, and rising again in 2025. The two composites therefore did not evolve in parallel throughout the observation period. These annual means were descriptive and were not interpreted as evidence of an intervention effect or structural break. Year-specific descriptive values are reported in Supplementary Table S6.

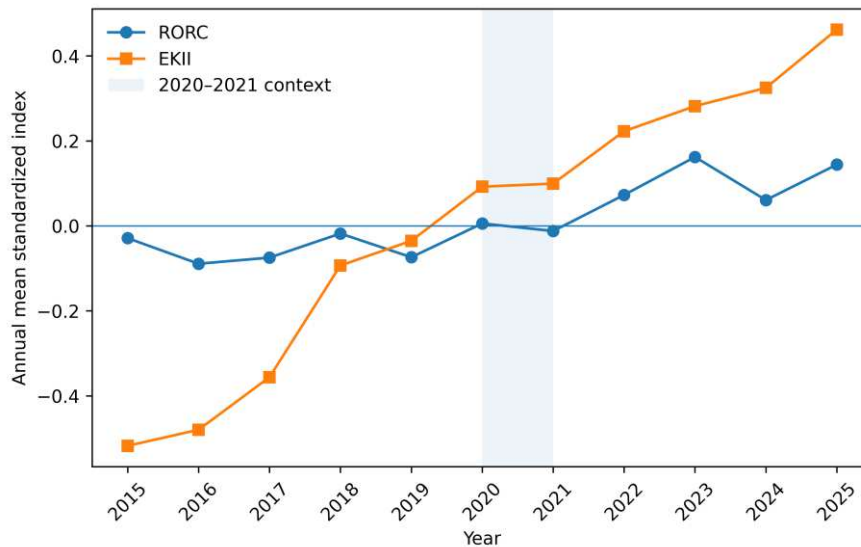


Figure 3. Annual mean RORC and EKII in the strict eligible panel, 2015–2025

Note. Means are based on available complete construct rows among the 38 eligible institutions. The shaded 2020–2021 interval is contextual pandemic-period shading only; it is not a causal intervention specification.

4.2 Distributions and Within–Between Structure

The strict eligible panel exhibited substantial heterogeneity across institution-years (Table 4). Aviation output averaged 117.7 works and 186.5 unique researchers, whereas the corresponding medians were 77 and 111, indicating right-skewed distributions; maxima reached 1,127 works and 2,234 researchers. Mean aviation R&D intensity was .041. Mean external academic/research and international collaboration rates were .541 and .342, respectively, while company coauthorship and government collaboration rates averaged .092 and .075.

Between-institution variation exceeded within-institution variation for several component measures, including aviation works (SD = 122.36 versus 74.98), unique aviation researchers (SD = 252.98 versus 143.14), and aviation R&D intensity (SD = .060 versus .012). The pooled RORC–EKII correlation was .327; RORC was weakly correlated with CSLI ($r = .012$), and the EKII–CSLI correlation was .076. This descriptive contrast was consistent with stable institutional heterogeneity contributing to the pooled RORC–EKII relationship, but it did not constitute a formal estimate of the between-institution association. A correlated-random-effects or within–between model would be required to estimate the two components jointly.

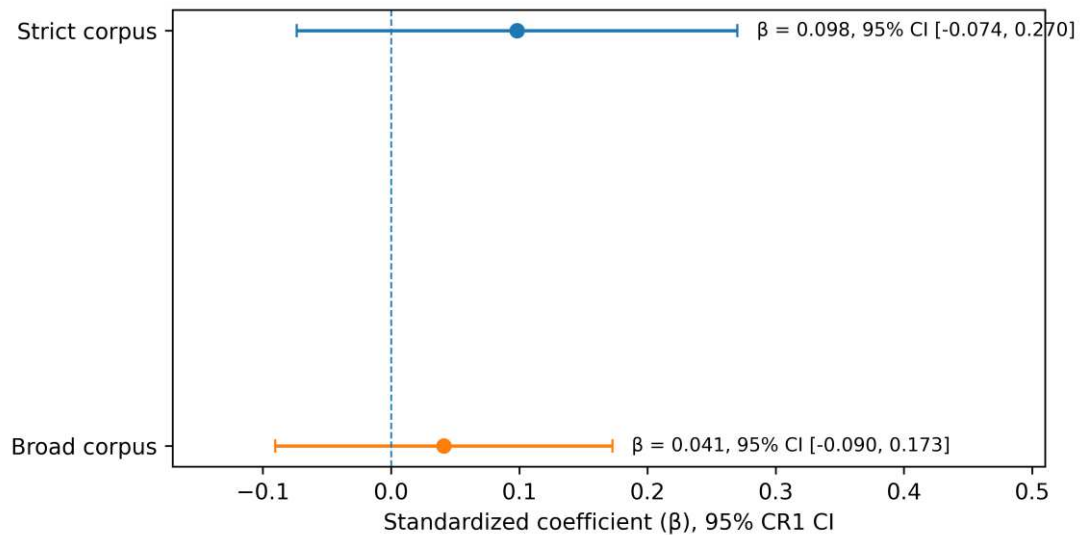
Table 4. Descriptive statistics for the strictly eligible panel

Variable	N	Mean	SD	Median	Minimum	Maximum
Aviation works	418	117.739	142.251	77	7	1,127
Unique aviation researchers	418	186.536	288.011	111	11	2,234
Researcher continuity	390	.335	.087	.339	.037	.636
Aviation R&D intensity	418	.041	.061	.014	.000	.350
Normalized topic breadth	418	.754	.098	.763	.212	.922
Academic/research partners	418	66.627	50.871	54	2	349
Partner countries	418	16.328	7.968	15	1	40
External academic/research rate	418	.541	.157	.517	.138	.932
International academic/research rate	418	.342	.179	.322	.000	.863
Company coauthorship rate	418	.092	.061	.081	.000	.296
Government collaboration rate	418	.075	.067	.055	.000	.385

4.3 Primary Longitudinal Association

The strict primary model included 320 aligned observations from 38 institutions. A one-standard-deviation increase in RORC in year $s-1$ was associated with a .098-standard-deviation increase in EKII in year s ($\beta = .098$, 95% CR1 CI $[-.074, .270]$, wild-cluster $p = .230$; within $R^2 = .363$). The broad-corpus estimate was smaller ($\beta = .041$, 95% CR1 CI $[-.090, .173]$, wild-cluster $p = .510$; within $R^2 = .335$; $N = 348$; 38 institutions). Exact estimates are reported in Table 5 and visually compared in Figure 4.

The estimates did not provide clear support for H1. Both point estimates were positive, but their confidence intervals crossed zero and the broad-corpus magnitude was smaller. The strict interval remained compatible with effects ranging from a small negative association to a moderately positive association. The evidence was therefore interpreted as inconclusive rather than as demonstrating either a positive association or the absence of a relationship between prior RORC and subsequent EKII.

**Figure 4. Primary one-year-lagged RORC–EKII association**

4.4 Exploratory Cross-Sector Linkage Analyses

The exploratory sequential-association analysis did not identify a stable relationship between prior RORC and subsequent CSLI. In the strict corpus, the RORC–CSLI association estimated without lagged EKII was negative

and imprecise ($\beta = -.053$, 95% CR1 CI $[-.291, .186]$, wild-cluster $p = .647$). In the joint model containing both lagged EKII and prior RORC, EKII was positively but imprecisely associated with subsequent CSLI ($\beta = .078$, 95% CR1 CI $[-.183, .340]$, wild-cluster $p = .500$), while the conditional RORC coefficient remained negative and imprecise ($\beta = -.061$, 95% CR1 CI $[-.298, .177]$, wild-cluster $p = .589$).

The broad-corpus estimates yielded the same substantive pattern. The RORC–CSLI association estimated without lagged EKII was $\beta = -.103$ (95% CR1 CI $[-.338, .131]$, wild-cluster $p = .369$). In the joint model, the EKII–CSLI coefficient was $\beta = .081$ (95% CR1 CI $[-.146, .307]$, wild-cluster $p = .450$), whereas the conditional RORC coefficient was $\beta = -.106$ (95% CR1 CI $[-.339, .127]$, wild-cluster $p = .327$). Thus, neither corpus provided sufficiently precise evidence that changes in organizational R&D capability or external knowledge integration were followed by changes in cross-sector research linkage. Complete estimates are reported in Table 5 and Supplementary Table S7 and are displayed in Figure 5.

Table 5. Primary and exploratory longitudinal estimates

Analysis	Strict β (SE)	Strict 95% CI	$p^\dagger$	Broad β (SE)	Broad 95% CI	$p^\dagger$
Primary H1: RORC $\rightarrow$ EKII	.098 (.085)	$[-.074, .270]$	0.230	.041 (.065)	$[-.090, .173]$	0.510
Exploratory: RORC $\rightarrow$ CSLI	-.053 (.118)	$[-.291, .186]$	0.647	-.103 (.116)	$[-.338, .131]$	0.369
Exploratory: EKII $\rightarrow$ CSLI RORC	.078 (.129)	$[-.183, .340]$	0.500	.081 (.112)	$[-.146, .307]$	0.450
Conditional RORC $\rightarrow$ CSLI EKII	-.061 (.117)	$[-.298, .177]$	0.589	-.106 (.115)	$[-.339, .127]$	0.327
Descriptive coefficient product	0.008	$[-.026, .034]$	0.676	0.003	$[-.014, .023]$	0.813

Note. Primary H1: strict $N = 320$ and broad $N = 348$; exploratory CSLI models: strict $N = 290$ and broad $N = 318$; all models include 38 institutions. The EKII $\rightarrow$ CSLI and conditional RORC $\rightarrow$ CSLI coefficients are estimated jointly in the same exploratory model. RORC = Revealed Organizational R&D Capability; EKII = External Knowledge Integration Index; CSLI = Cross-Sector Research Linkage Index. † Coefficient-row p -values use 999 null-imposed wild-cluster replications with Rademacher weights; confidence intervals use institution-clustered CR1 standard errors. The coefficient-product row uses 2,000 institution-history bootstrap replications and must not be interpreted as a causal pathway.

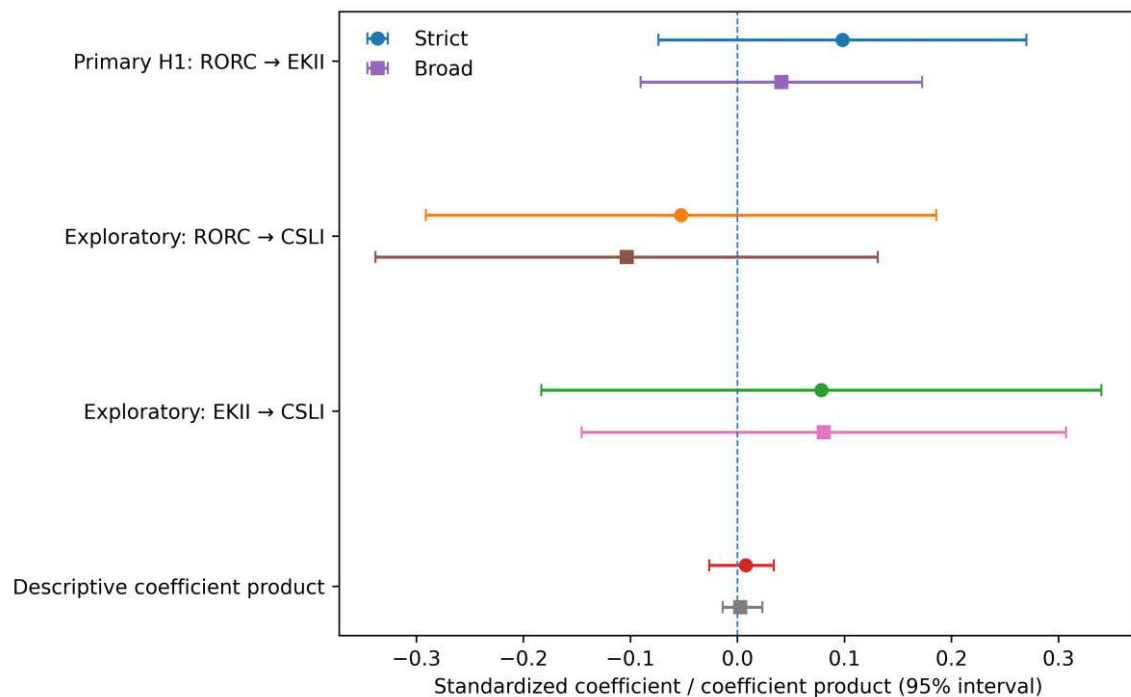


Figure 5. Strict- and broad-corpus longitudinal estimates

Note. Intervals for coefficient estimates are institution-clustered 95% CR1 confidence intervals. Intervals for the descriptive coefficient product are institution-history bootstrap intervals. The coefficient product summarizes sequential associations and must not be interpreted as a causal indirect effect.

4.5 Robustness and Sensitivity

Alternative temporal alignments did not establish a stable RORC–EKII relationship. Across the exploratory contemporaneous, one-year, and two-year RORC–EKII gap specifications, strict estimates ranged from $\beta = .043$ to $.099$, whereas broad estimates ranged from $\beta = -.003$ to $.142$. The broad contemporaneous specification produced the largest estimate ($\beta = .142$, 95% CR1 CI $[.004, .279]$, wild-cluster $p = .043$; $N = 386$; 38 institutions). This contemporaneous model was an exploratory temporal sensitivity specification and was not the theory-guided H1 test. Moreover, the association did not persist under the primary one-year gap ($\beta = .041$, 95% CR1 CI $[-.090, .173]$, wild-cluster $p = .510$) or the two-year gap ($\beta = -.003$, 95% CR1 CI $[-.137, .132]$). The temporal sensitivity results therefore indicated that the estimated RORC–EKII association was dependent on temporal alignment rather than stable across specifications.

As summarized in Table 6, leave-one-indicator-out RORC–EKII estimates ranged from $.049$ to $.195$ in the strict corpus and from $-.013$ to $.099$ in the broad corpus, showing sensitivity to index composition. In leave-one-institution-out exploratory analyses, the EKII–CSLI coefficient remained positive but imprecise, ranging from $.027$ to $.145$ in the strict corpus and from $.046$ to $.135$ in the broad corpus. Collectively, the sensitivity analyses did not identify a specification-independent relationship; effect magnitude and precision depended on temporal alignment, aviation-classification scope, and index composition. Complete coefficient-level robustness results are reported in Supplementary Table S7.

Table 6. Summary of robustness and sensitivity evidence

Specification	Strict result	Broad result	Interpretation
RORC $\rightarrow$ EKII, 0–2-year RORC–EKII gaps	$\beta = .043$ to $.099$	$\beta = -.003$ to $.142$	Broad contemporaneous estimate was positive and more precise, but not stable across temporal alignments.
RORC $\rightarrow$ EKII, leave-one-indicator-out	$\beta = .049$ to $.195$	$\beta = -.013$ to $.099$	Sensitive to index composition
Primary H1, classification boundary	$\beta = .098$	$\beta = .041$	Positive but imprecise in both corpora
EKII $\rightarrow$ CSLI, leave-one-institution-out	$\beta = .027$ to $.145$	$\beta = .046$ to $.135$	Direction remained positive; estimates were imprecise
AVI038 denominator invariant	467–747 total works; total $\geq$ strict aviation works	467–747 total works; total $\geq$ broad aviation works	All 11 corrected institution-years passed
Data reconciliation	0 duplicate or mismatched keys	0 duplicate or mismatched keys	Computational integrity checks passed

Note. RORC = Revealed Organizational R&D Capability; EKII = External Knowledge Integration Index; CSLI = Cross-Sector Research Linkage Index. A zero-year gap denotes contemporaneity.

5. Discussion

5.1 Interpretation of the Main Finding

The principal result was a contrast between the positive pooled relationship and the imprecisely estimated within-institution association. The pooled RORC–EKII correlation was $.327$, indicating that institutions observed at persistently higher levels of revealed R&D capability also tended to exhibit greater external knowledge integration. However, annual increases in RORC were not followed consistently by higher EKII one year later. The strict fixed-effects coefficient was positive but uncertain ($\beta = .098$), while the broad-corpus estimate was closer to zero ($\beta = .041$). This contrast is consistent with durable institutional characteristics—including mission, infrastructure, reputation, national-system position, and accumulated partnerships—contributing to the pooled relationship. It does not establish that between-institution differences dominated, because the between- and within-institution components were not estimated jointly.

The finding therefore refines rather than rejects an absorptive-capacity interpretation. Prior related knowledge may support an institution’s ability to recognize relevant external knowledge without being sufficient to establish or expand organizational relationships. Partnership development may also depend on search routines,

administrative support, incentives, funding continuity, intellectual-property arrangements, boundary-spanning personnel, and trust. Strategic flexibility may help organizations convert knowledge resources into innovation outcomes (Miroshnychenko et al., 2021), but such mechanisms were not directly captured by the bibliometric indicators. The observed pattern may therefore reflect a gap between the productive research resources represented by RORC and the organizational mechanisms required to expand external integration.

Temporal scale provides an additional explanation. A one-year interval may be short relative to research-center formation, consortium development, doctoral training, infrastructure investment, and aviation certification cycles. Sensitivity analyses did not identify a stable RORC–EKII relationship across contemporaneous, one-year, and two-year gaps (Table 6). The broad contemporaneous estimate was positive, but it did not persist under the one- or two-year alignments. Future research should therefore examine longer periods and identifiable organizational events rather than assume that annual changes in research capability translate immediately into changes in external knowledge networks.

5.2 Theoretical Contribution

The study's primary theoretical contribution is the separation of revealed organizational capability, external knowledge integration, and downstream research translation. RORC represented an observable aviation-research knowledge base, whereas EKII represented relationships with external academic and research organizations. Neither construct alone was treated as a complete measure of absorptive capacity. This separation responds to concerns about reifying absorptive capacity as a single observable attribute and recognizes that its potential and realized dimensions may operate differently (Algarni et al., 2023; Lane et al., 2006; Sancho-Zamora et al., 2021).

The pooled and fixed-effects results further demonstrate why cross-sectional and longitudinal implications should not be conflated. A pooled relationship may reflect enduring configurations of resources, reputation, institutional mission, and network position, whereas a fixed-effects coefficient addresses whether deviations from an institution's usual capability level are followed by changes in integration. The coexistence of a pooled correlation of .327 and an imprecisely estimated one-year longitudinal coefficient is consistent with absorptive capacity being embedded in relatively persistent institutional configurations. Nevertheless, a correlated-random-effects or within-between model is required before the relative contributions of stable institutional differences and within-institution change can be estimated directly.

The study also establishes an explicit conceptual boundary between knowledge integration and downstream translation. CSLI captured observable company and government coauthorship but could not demonstrate licensing, implementation, regulatory use, operational improvement, or broader societal impact. Research-impact scholarship emphasizes multiple pathways and contribution processes that cannot be reduced to publication relationships alone (Dotti & Walczyk, 2022; Reed et al., 2021; Viana-Lora & Nel-lo-Andreu, 2021). By retaining CSLI as a cross-sector linkage signal rather than relabeling it as research translation, the study clarifies the additional evidence required for a defensible translation analysis.

5.3 Implications for Longitudinal Institutional Scientometrics

Methodologically, the study contributes a transparent international purposive panel covering 40 aviation-focused HEIs from 2015 to 2025. The reproducibility architecture retained work-level provenance, institutional lineages, affiliation fallbacks, construction-buffer records, classification evidence, collaboration edges, and auditable deduplication and reconciliation rules. Deduplication was performed by institution and OpenAlex Work ID while preserving distinct records that shared a DOI. Explicit blank-versus-verified-zero rules also prevented unresolved external evidence from being converted into artificial non-events. These safeguards are important because OpenAlex coverage and metadata quality can differ across sources, countries, institutions, and publication types (Culbert et al., 2025; Zheng et al., 2025).

Classification uncertainty was retained analytically through strict and broad specifications rather than concealed as a preprocessing decision. After correction of the AVIO38 denominator and complete re-estimation, the primary RORC–EKII coefficient decreased from .098 in the strict corpus to .041 in the broad corpus, and leave-one-

indicator-out estimates varied materially. The denominator audit also demonstrated the importance of treating a failed denominator as missing rather than as a verified zero and of enforcing $\text{Total_Institutional_Works} \geq \text{Aviation_Works}$ before computing intensity. Because no independent human validation or provenance-complete AI-coded audit file was available, no classification-concordance statistic was retained as a completed result.

Empirically, the panel documents substantial growth in aviation-related research and external integration across the sampled institutions. Nevertheless, the primary longitudinal estimate remained imprecise and sensitive to classification scope and composite construction. The result should therefore be interpreted as an informative constraint on the available evidence rather than as proof that no relationship exists. It also challenges the assumption that expanding publication activity alone will automatically broaden an institution's external knowledge network. University–industry and research–impact pathways are multidimensional and may unfold over different organizational and temporal scales (Cohen et al., 2025; Cunningham et al., 2025). The present estimates provide a transparent benchmark for future studies using direct organizational measures, expert-validated classifications, longer observation periods, or quasi-experimental designs.

5.4 Practical Implications

The findings do not establish that any specific managerial intervention will increase external knowledge integration. They nevertheless indicate that publication growth should not be assumed to produce broader knowledge relationships automatically. Investments in boundary-spanning personnel, consortium governance, shared facilities, data-sharing arrangements, research administration, and regulatory engagement are theoretically plausible complements to research scale and specialization, but their effectiveness should be evaluated prospectively rather than inferred from the present associations.

Aviation HEIs and funding agencies should therefore monitor longitudinal trajectories rather than rely solely on annual publication totals or institutional rankings. Relevant indicators may include partner diversity, international reach, relationship persistence, acquisition of new partners, and the continuity of cross-sector engagement. Claims of downstream translation should additionally require independently verified evidence of patent families, standards participation, licensing, curriculum adoption, demonstrations, policy use, certification support, safety interventions, or operational implementation. Such evidence should remain traceable to its source and should not be incorporated into an opaque composite.

5.5 Limitations and Future Research

Several limitations bound the findings. First, the 40 institutions formed a purposive international frame rather than a probability sample, limiting generalizability beyond the sampled aviation-focused HEIs. Second, primary eligibility required at least seven active aviation-research years and 20 strict institution–work records. Although these criteria were frozen before outcome analysis to support longitudinal observability rather than institutional ranking, complete-frame and alternative-threshold fixed-effects models were not completed; future work should test reasonable alternative eligibility specifications where data coverage permits. Third, AVI-TAX-v1.0 was frozen and reproducible but lacked independent human criterion validation. Strict and broad corpora provided classification-boundary sensitivity only; future validation should use blinded coding by qualified aviation subject-matter reviewers and documented adjudication, rather than AI-assisted agreement as a substitute for expert validation.

Fourth, OpenAlex affiliation, topic, abstract, author, and institutional metadata may vary across countries, years, institutions, and publication types, as illustrated by the AVI038 denominator correction. Fifth, RORC and EKII were formative bibliometric composites rather than direct measures of organizational recognition, assimilation, transformation, or exploitation. Sixth, some GERD country-years required documented interpolation or limited carry-forward, introducing additional measurement uncertainty despite provenance flagging. Seventh, fixed effects and temporal ordering did not eliminate measurement error, time-varying confounding, reciprocal relationships, or selection into collaboration. Finally, CSLI captured publication-based company and government linkages only and should not be interpreted as evidence of research translation or broader impact.

Future research should independently validate the aviation taxonomy using qualified subject-matter coders and estimate within- and between-institution effects jointly through correlated-random-effects or within-between models. Alternative eligibility thresholds should be examined, and network models should distinguish the acquisition of new partners from the persistence or deepening of existing relationships. Event-study or quasi-experimental designs could evaluate identifiable changes such as research-center openings, major funding awards, consortium membership, or policy reforms. Authenticated patent-family, standards, licensing, curriculum, policy, demonstration, certification, and operational-use evidence should be collected prospectively under a separately registered translation protocol rather than retrospectively inferred from coauthorship. These extensions would preserve the present study's distinction between research capability, external knowledge integration, cross-sector research linkage, and independently verified downstream use.

6. Conclusion

This study assembled a lineage-aware 2015–2025 computational panel of 40 aviation-focused higher education institutions and developed distinct measures of revealed organizational R&D capability and external knowledge integration. Following correction of the AVI038 institutional-output denominator and complete re-estimation, RORC and EKII remained positively correlated in pooled observations, whereas the one-year-lagged within-institution estimates were positive but imprecise in both the strict ($\beta = .098$) and broad ($\beta = .041$) corpora. The evidence therefore did not provide clear support for a stable short-run RORC–EKII association, while remaining compatible with a substantively meaningful positive relationship.

The study contributes by distinguishing research capability, external integration, cross-sector publication linkage, and downstream translation; preserving work-level provenance and classification uncertainty; and enforcing explicit denominator and deduplication invariants. More broadly, the findings demonstrate the value of longitudinal institutional scientometrics that separates research-production capability from external knowledge integration, preserves uncertainty in indicator construction, and distinguishes persistent cross-sectional differences from within-institution change. By excluding incomplete patent, institutional-report, and validation evidence and by relying on a manuscript-aligned longitudinal workflow rather than the legacy RTI framework, the study provides a transparent and reproducible foundation for future research on institutional research capability, knowledge integration, and absorptive capacity in aviation higher education.

Data and Code Availability

The data and reproducibility materials used in this study are publicly available in Mendeley Data: Dela Peña, Arthur (2026), “Aviation HEI RORC–EKII Dataset, 2015–2025,” Version 2, DOI: 10.17632/48mv7gd4sx.2. The repository includes the processed institution-year and construct datasets, analysis-ready files, statistical outputs, quality-control records, and supporting analysis code required to reproduce the reported analyses. Raw OpenAlex API-response caches and unprocessed affiliation/contact strings are not included; underlying OpenAlex and ROR metadata remain subject to their respective terms.

Ethics and Responsible Research Statement

The study analyzed publicly available bibliographic and organizational metadata and reported institution-level results. It involved no participant recruitment, intervention, private records, or individual performance evaluation. Author identifiers were used only for record linkage and aggregated continuity measures. No formal institutional ethics approval, exemption, or non-human-participant determination was obtained or is claimed in this manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement

The authors declare no conflict of interest relevant to this study.

References

- Algarni, M. A., Ali, M., Leal-Rodríguez, A. L., & Albort-Morant, G. (2023). The differential effects of potential and realized absorptive capacity on imitation and innovation strategies, and its impact on sustained competitive advantage. *Journal of Business Research*, 158, Article 113674. <https://doi.org/10.1016/j.jbusres.2023.113674>
- Asiedu, M. A., Anyigba, H., & Doe, J. K. (2023). Absorptive capacity and innovation generation in higher education institutions: The mediating role of inter-functional coordination. *The Learning Organization*, 30(4), 385–405. <https://doi.org/10.1108/TLO-11-2022-0128>
- Awwad, M. S., Abuhommous, A. A., & Adaileh, A. M. (2025). Unravelling the relationship between absorptive capacity, innovation, and financial performance: A longitudinal study. *International Journal of Innovation Studies*, 9(2), 116–128. <https://doi.org/10.1016/j.ijis.2025.02.001>
- Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity*, 53, 1051–1074. <https://doi.org/10.1007/s11135-018-0802-x>
- Brunton, S. L., Kutz, J. N., Manohar, K., Aravkin, A. Y., Morgansen, K., Klemisch, J., Goebel, N., Buttrick, J., Poskin, J., Blom-Schieber, A., Hogan, T., & McDonald, D. (2021). Data-driven aerospace engineering: Reframing the industry with machine learning. *AIAA Journal*, 59(8), 2820–2847. <https://doi.org/10.2514/1.J060131>
- Cai, X., Coffman, D. L., Piper, M. E., & Li, R. (2022). Estimation and inference for the mediation effect in a time-varying mediation model. *BMC Medical Research Methodology*, 22, Article 113. <https://doi.org/10.1186/s12874-022-01585-x>
- Cai, Y., Ma, J., & Chen, Q. (2020). Higher education in innovation ecosystems. *Sustainability*, 12(11), Article 4376. <https://doi.org/10.3390/su12114376>
- Cohen, M., Fernandes, G., & Godinho, P. (2025). Measuring the impacts of university–industry R&D collaborations: A systematic literature review. *The Journal of Technology Transfer*, 50, 345–374. <https://doi.org/10.1007/s10961-024-10114-5>
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128–152. <https://doi.org/10.2307/2393553>
- Compagnucci, L., & Spigarelli, F. (2020). The third mission of the university: A systematic literature review on potentials and constraints. *Technological Forecasting and Social Change*, 161, Article 120284. <https://doi.org/10.1016/j.techfore.2020.120284>
- Culbert, J. H., Hobert, A., Jahn, N., Haupka, N., Schmidt, M., Donner, P., & Mayr, P. (2025). Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. *Scientometrics*, 130, 2475–2492. <https://doi.org/10.1007/s11192-025-05293-3>
- Cunningham, J. A., Menter, M., & Starke, F. (2025). The evolution of university technology transfer research: A text mining approach. *The Journal of Technology Transfer*, 50, 1231–1268. <https://doi.org/10.1007/s10961-024-10133-2>
- da Silva Florêncio, M. N., & de Oliveira Júnior, A. M. (2022). The importance of absorptive capacity in technology transfer and organisational performance: A systematic review. *International Journal of Innovation Management*, 26(2), Article 2230001. <https://doi.org/10.1142/S136391962230001X>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Dotti, N. F., & Walczyk, J. (2022). What is the societal impact of university research? A policy-oriented review to map approaches, identify monitoring methods and success factors. *Evaluation and Program Planning*, 95, Article 102157. <https://doi.org/10.1016/j.evalprogplan.2022.102157>
- Du, Y., Wang, M., & Yin, Y. (2024). Research on innovation cooperation network of Chinese universities based on patent data. *Journal of Engineering and Technology Management*, 71, Article 101784. <https://doi.org/10.1016/j.jengtecman.2023.101784>
- Greco, S., Ishizaka, A., Tasiou, M., & Torrì, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, 141, 61–94. <https://doi.org/10.1007/s11205-017-1832-9>
- Haunschild, R., & Bornmann, L. (2024). The use of OpenAlex to produce meaningful bibliometric global overlay maps of science on the individual, institutional, and national levels. *PLOS ONE*, 19(12), Article e0308041. <https://doi.org/10.1371/journal.pone.0308041>
- Hickie, J., & Hickie, D. (2025). Complete and incomplete innovation systems in civil aerospace. *European Planning Studies*, 33(8), 1336–1350. <https://doi.org/10.1080/09654313.2025.2541246>
- Iino, T., Inoue, H., Saito, Y. U., & Todo, Y. (2021). How does the global network of research collaboration affect the quality of innovation? *The Japanese Economic Review*, 72(1), 5–48. <https://doi.org/10.1007/s42973-020-00042-y>

- Ipinnaiye, O., Lenihan, H., Mulligan, K., Doran, J., & Roper, S. (2026). Innovation success in small and larger firms: Does persistence and diversity in R&D matter when building a unique knowledge base? *The Journal of Technology Transfer*, 51(3), 1564–1600. <https://doi.org/10.1007/s10961-025-10209-7>
- Kaur, V. (2023). Knowledge-based dynamic capabilities: A scientometric analysis of marriage between knowledge management and dynamic capabilities. *Journal of Knowledge Management*, 27(4), 919–952. <https://doi.org/10.1108/JKM-02-2022-0112>
- Lane, P. J., Koka, B. R., & Pathak, S. (2006). The reification of absorptive capacity: A critical review and rejuvenation of the construct. *Academy of Management Review*, 31(4), 833–863. <https://doi.org/10.5465/amr.2006.22527456>
- Luna-Andrade, J. J., Salonitis, K., & Brintrup, A. (2021). Key enablers for the evolution of aerospace ecosystems. *Journal of Aerospace Technology and Management*, 13, Article e3321. <https://doi.org/10.1590/jatm.v13.1225>
- MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2023). Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, 232(2), 272–299. <https://doi.org/10.1016/j.jeconom.2022.04.001>
- Miroshnychenko, I., Strobl, A., Matzler, K., & De Massis, A. (2021). Absorptive capacity, strategic flexibility, and business model innovation: Empirical evidence from Italian SMEs. *Journal of Business Research*, 130, 670–682. <https://doi.org/10.1016/j.jbusres.2020.02.015>
- Muhonen, R., Benneworth, P., & Olmos-Peñuela, J. (2020). From productive interactions to impact pathways: Understanding the key dimensions in developing SSH research societal impact. *Research Evaluation*, 29(1), 34–47. <https://doi.org/10.1093/reseval/rvz003>
- O'Connor, C., & Joffe, H. (2020). Inter-coder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1–13. <https://doi.org/10.1177/1609406919899220>
- Passos, J. B., Enrique, D. V., Dutra, C. C., & ten Caten, C. S. (2023). University–industry collaboration process: A systematic review of literature. *International Journal of Innovation Science*, 15(3), 479–506. <https://doi.org/10.1108/IJIS-11-2021-0216>
- Perkmann, M., Salandra, R., Tartari, V., McKelvey, M., & Hughes, A. (2021). Academic engagement: A review of the literature 2011–2019. *Research Policy*, 50(1), Article 104114. <https://doi.org/10.1016/j.respol.2020.104114>
- Reed, M. S., Ferré, M., Martin-Ortega, J., Blanche, R., Lawford-Rolfe, R., Dallimer, M., & Holden, J. (2021). Evaluating impact from research: A methodological framework. *Research Policy*, 50(4), Article 104147. <https://doi.org/10.1016/j.respol.2020.104147>
- Robertson, J., Caruana, A., & Ferreira, C. (2023). Innovation performance: The effect of knowledge-based dynamic capabilities in cross-country innovation ecosystems. *International Business Review*, 32(2), Article 101866. <https://doi.org/10.1016/j.ibusrev.2021.101866>
- Rybníček, R., & Königsgruber, R. (2019). What makes industry–university collaboration succeed? A systematic review of the literature. *Journal of Business Economics*, 89, 221–250. <https://doi.org/10.1007/s11573-018-0916-6>
- Sancho-Zamora, R., Gutiérrez-Broncano, S., Hernández-Perlines, F., & Peña-García, I. (2021). A multidimensional study of absorptive capacity and innovation capacity and their impact on business performance. *Frontiers in Psychology*, 12, Article 751997. <https://doi.org/10.3389/fpsyg.2021.751997>
- Stettler, T. R., Moosauer, E. J., Schweiger, S. A., Baldauf, A., & Audretsch, D. (2025). Absorptive capacity in a more (or less) absorptive environment: A meta-analysis of contextual effects on firm innovation. *Journal of Product Innovation Management*, 42(1), 18–47. <https://doi.org/10.1111/jpim.12758>
- Sun, Y., Zhang, C., & Kok, R. A. W. (2020). The role of research outcome quality in the relationship between university research collaboration and technology transfer: Empirical results from China. *Scientometrics*, 122(2), 1003–1026. <https://doi.org/10.1007/s11192-019-03330-6>
- Taxt, R. E., Robinson, D. K. R., Schoen, A., & Fløysand, A. (2022). The embedding of universities in innovation ecosystems: The case of marine research at the University of Bergen. *Norsk Geografisk Tidsskrift–Norwegian Journal of Geography*, 76(1), 42–60. <https://doi.org/10.1080/00291951.2022.2041718>
- van den Akker, O. R., Weston, S. J., Campbell, L., Chopik, W. J., Damian, R. I., Davis-Kean, P. E., Hall, A. N., Kosie, J. E., Kruse, E., Olsen, J., Ritchie, S. J., Valentine, K. D., van 't Veer, A. E., & Bakker, M. (2021). Preregistration of secondary data analysis: A template and tutorial. *Meta-Psychology*, 5, Article 2625. <https://doi.org/10.15626/MP.2020.2625>
- Viana-Lora, A., & Nel-lo-Andreu, M. G. (2021). Approaching the social impact of research through a literature review. *International Journal of Qualitative Methods*, 20, Article 16094069211052189. <https://doi.org/10.1177/16094069211052189>

- Weston, S. J., Ritchie, S. J., Rohrer, J. M., & Przybylski, A. K. (2019). Recommendations for increasing the transparency of analyses of preexisting data sets. *Advances in Methods and Practices in Psychological Science*, 2(3), 214–227. <https://doi.org/10.1177/2515245919848684>
- Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review*, 27(2), 185–203. <https://doi.org/10.5465/amr.2002.6587995>
- Zheng, M., Miao, L., Bu, Y., & Larivière, V. (2025). Understanding discrepancies in the coverage of OpenAlex: The case of China. *Journal of the Association for Information Science and Technology*, 76(11), 1591–1601. <https://doi.org/10.1002/asi.70013>

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)