Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

When Silence Is Not Exit: What Publication-Affiliation Traces Can and Cannot Reveal About Researcher Mobility in Oman, Qatar, and the United Arab Emirates

AHMED Shehata

a.shehata@squ.edu.om

Sultan Qaboos University

Research Article

Keywords: researcher mobility, incomplete observation, publication disappearance, affiliation metadata, competing risks, research evaluation

Posted Date: October 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10904356/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Bibliometric mobility indicators often treat disappearance of a national publication affiliation as international departure. We develop a framework that separates first foreign-only publication transitions, first completed three-year publication gaps, later return or publication resumption, and administrative censoring. The analysis covers 49,587 country-specific spells for 49,270 Scopus Author IDs entering Oman, Qatar, or the United Arab Emirates during 2010–2020 and followed through 2025. At ten years, prospective first-foreign-only cumulative incidence was 14.3% (95% bootstrap CI [13.6%, 15.1%]) in Oman, 23.4% [22.6%, 24.1%] in Qatar, and 18.0% [17.5%, 18.5%] in the UAE; first-completed-gap incidence was 77.4%, 62.8%, and 70.6%. Publication later resumed after 22.3%, 20.0%, and 17.3% of first completed gaps, while target affiliation returned after 26.2%, 21.1%, and 21.1% of first foreign-only events. Terminal status at 2025 and ORCID current employment are retained as secondary sensitivity analyses. Adjusted country-by-time models show heterogeneous publication-affiliation contrasts rather than stable retention rankings. Publication traces reveal observable research participation, not complete migration histories.

1. Introduction

Internationally mobile researchers transmit skills, tacit knowledge, and professional networks across institutions and national research systems. Mobility may expand collaboration and impact, but its consequences differ across career stages, fields, destinations, and institutional conditions (Franzoni et al., 2012; Netz et al., 2020; Sugimoto et al., 2017). Governments and universities therefore seek scalable indicators of researcher inflow, retention, circulation, and departure.

Publication databases offer broad longitudinal coverage, but they record authorship and publication affiliations—not residence, employment contracts, citizenship, or border crossings. Bibliometric studies have nevertheless reconstructed mobility by comparing the countries attached to an author's publications over time (Appelt et al., 2015; Moed et al., 2013; Moed & Halevi, 2014). This approach is useful when interpreted as an indicator of publication-affiliation change. It becomes problematic when the last appearance of a national affiliation is automatically called emigration. If a researcher ceases to publish, moves to an outlet outside database coverage, enters a long publication gap, or acquires a split Author ID, the observed trace ends without positive evidence of movement.

This paper treats the problem as incomplete observation rather than formal identification of a physical-migration parameter. The primary outcomes are prospective in the annual record: the first foreign-only publication state and the first point at which three consecutive publication-free years have elapsed. Neither event is revoked by later return or publication resumption. Fixed-window terminal classifications and ORCID current status are secondary sensitivity analyses, not population bounds.

The empirical setting comprises Oman, Qatar, and the UAE: three small, internationally connected Gulf research systems that expanded capacity through international recruitment, collaboration, new universities, and targeted investment (Al-Marzouqi & Arabi, 2022). They were selected because the same country-seed and complete-author-history workflow could be executed and verified across systems of different publication scale, while keeping the first comparative application computationally and conceptually bounded. Saudi Arabia, Bahrain, and Kuwait were outside this prespecified three-country scope; the sample is not presented as representative of the entire Gulf. The comparison tests portability of a common measurement framework rather than ranking national policy performance (El-Ouahi et al., 2021; Paul et al., 2026).

The study makes four contributions. First, it separates first observed transitions from later return and resumption. Second, it implements an author-specific alignment rule that refuses to assign every publication country to every

coauthor. Third, it combines prospective competing-event description with nonabsorbing annual multistate analysis and adjusted discrete-time models. Fourth, it treats terminal 2025 status and current ORCID employment as explicitly secondary, selection-sensitive checks.

The research questions are:

1. What lower and upper bibliographic classifications of departure are supported by the observed affiliation histories?
2. How do terminal foreign-affiliation transition, publication disappearance, and event-free participation evolve over time, with uncertainty and risk sets reported explicitly?
3. Do country differences persist when analysis is conditioned on continued publication visibility and when field and academic age are controlled?
4. How common are co-affiliation, foreign-only states, and observed returns?
5. How much can external ORCID employment evidence narrow the unresolved disappearance category?

2. Literature review and conceptual framework

2.1 Inferring mobility from publication affiliations

Large-scale bibliometric mobility research developed from the capacity to link disambiguated author profiles to publication affiliations. Moed et al. (2013) demonstrated the feasibility of reconstructing international mobility using Scopus Author IDs, while documenting the risks created by split or merged profiles. Moed and Halevi (2014) subsequently distinguished synchronous and asynchronous indicators and emphasized that bibliometric results require careful interpretation and complementary data. OECD experimental indicators adopted similar affiliation-change logic to describe bilateral flows, while explicitly warning that multiple affiliations, missing authors, indexing gaps, and the absence of academic status complicate physical-mobility interpretations (Appelt et al., 2015). The issue remains operationally current: the OECD's production dataset reports "implied bilateral mobility flows" from authors' main countries in their first and last indexed publications during 2010–2024 (OECD, 2026). The present study examines the uncertainty hidden by that shortcut.

The globalization of scientific careers makes such indicators policy-relevant beyond any one region (Czaika & Orazbayev, 2018). Later work refined the categories encoded by affiliation trajectories. Robinson-Garcia et al. (2019) distinguished migrants, whose earlier country link is interrupted, from travelers who add countries while retaining prior affiliations. This distinction matters because co-affiliation can represent a transition, a visiting appointment, continuing collaboration, honorary affiliation, or strategic multiple affiliation. Treating every new country as a move overstates mobility and erases circulation. Global studies have nevertheless shown that affiliation-defined mobility is associated with collaboration and citation impact, reinforcing its policy relevance while underscoring its selected nature (Sugimoto et al., 2017).

Bibliometric mobility is also career- and field-dependent. Momeni et al. (2022) documented differences by discipline, gender, career stage, nationality, and collaboration position. Netz et al.'s (2020) systematic review found the most consistent benefit of international mobility in broader networks, with heterogeneous evidence for productivity, impact, employment, and other outcomes. Academic age therefore cannot be treated as optional in comparative models, and country coefficients cannot be interpreted without accounting for disciplinary composition.

2.2 Identity, affiliation, and coverage errors

A mobility observation requires three correct links: publications must be assigned to the correct author, affiliations must be attached to that author rather than another coauthor, and the affiliation must be mapped to the correct country. Scopus Author IDs have sufficient precision for large-scale work, but split and merged profiles remain consequential for longitudinal careers. In a comparison with curricula vitae, Aman (2018) found that most mobility episodes were recoverable while identifying inconsistencies caused by multiple Author IDs, missing coauthor affiliations, and incomplete CVs.

Affiliation quality is a separate problem from author disambiguation. Purnell (2022) demonstrated substantial disagreement in university assignments across Scopus, Web of Science, Dimensions, and Microsoft Academic, arising from missing affiliations, divergent organization-unification rules, and incorrect assignments. Coverage also varies systematically by field, language, and region; Scopus is broader than Web of Science but both overrepresent English-language and biomedical or natural-science journals relative to many social-science and humanities outlets (Mongeon & Paul-Hus, 2016). Baas et al. (2020) describe Scopus as a curated source appropriate for quantitative science studies, but curation does not make every author–affiliation link complete or error-free.

The consequences for mobility can be direct. Mitchell et al. (2026) show that affiliation errors can generate distorted researcher-mobility histories in both OpenAlex and Scopus and advocate targeted record-level validation with uncertainty quantification. Their result supports a crucial design principle: mobility should be inferred from author-linked affiliation evidence, not from an undifferentiated set of all countries appearing on a multi-author publication.

2.3 Validation and the observation process

Alternative sources measure different constructs. Surveys can capture motivations, residence, and self-reported moves but are selective and expensive. Curricula vitae provide event histories and have been used to disentangle migration from scientific mobility (Woolley & Turpin, 2009). ORCID contains self-reported employment and education affiliations and has supported large-scale mapping of scientific mobility (Gomez et al., 2020). ORCID is not a census: adoption and completeness vary by country, field, career stage, and individual disclosure choices. It is nevertheless valuable for external-status sensitivity analysis because it supplies an affiliation source outside the publication-address sequence.

Recent work illustrates the value of combining sources. Paul et al. (2026) linked OpenAlex histories to interviews with migrant scientists in Saudi Arabia, Qatar, and the UAE and treated dual affiliations as ambiguous. Carneiro et al. (2026) similarly argue that scientific-diaspora research should combine bibliometric traces with curricula vitae, institutional databases, documentary sources, and qualitative evidence. For return mobility, Zhao et al. (2022) used Scopus histories to model both departure and subsequent return among German-affiliated researchers, demonstrating why return should be retained as a nonabsorbing state rather than erased by a first-versus-last comparison. These studies reinforce the distinction between positive publication evidence and the physical or organizational reality it imperfectly proxies.

Publication visibility is part of the data-generating process. An exit can be observed only if an author continues publishing after leaving and the new output remains indexed and correctly linked. Consequently,

$$P(\text{terminal foreign-affiliation transition}) = P(\text{departure}) \times P(\text{continued observable publication} \mid \text{departure}),$$

Observing a foreign-only publication after target-country participation requires several jointly operating conditions: an underlying affiliation change, continued publication, database coverage, a foreign-only indexed address, correct author and affiliation linkage, and the absence of a simultaneous target-country address. False-positive affiliations and later

returns are also possible. Country contrasts therefore combine mobility and observation processes and cannot be read as national retention rankings.

2.4 Competing events and identification limits

Competing-risk methods distinguish the instantaneous rate of a cause among event-free observations from the absolute probability of experiencing that cause. A Kaplan–Meier curve (Kaplan & Meier, 1958) that censors a dominant competing event is a cause-specific construct, not the proportion of the original cohort known to remain (Fine & Gray, 1999; Latouche et al., 2013). The Aalen–Johansen cumulative-incidence function and empirical multistate transition matrix preserve probability mass across competing endpoints (Allignol et al., 2011) and are therefore the primary descriptive estimands here.

The identification framework has three levels. First, positive later foreign-only affiliation evidence identifies a terminal foreign-affiliation transition. Second, a permissive last-affiliation rule treats every early ending as departure and supplies an upper bibliographic classification, not a validated migration estimate. Third, external employment information can reallocate a subset of disappearance spells and produce a narrower, explicitly conditional region. Right-censored spells remain unresolved; none of these quantities is interpreted as a physical-migration rate without assumptions.

3. Data and methods

3.1 Scopus extraction and cohorts

Scopus records were retrieved through the Elsevier APIs on 23–25 August 2026. Country searches covered 2000–2025. The analytical cohorts comprise authors whose first positionally aligned target-country publication occurred during 2010–2020. Complete post-entry histories were retrieved through 2025 for every provisional cohort Author ID. The final cohorts contain 9,942 Oman spells, 12,028 Qatar spells, and 27,617 UAE spells: 49,587 country-specific spells belonging to 49,270 unique Author IDs. An Author ID may contribute a spell to more than one country; 314 IDs do so.

Records were deduplicated first by EID and then by normalized DOI when EID was unavailable. The final post-entry files contain 76,232 Oman, 143,555 Qatar, and 235,655 UAE publication records. Licensed records and author-level identifiers remain restricted; code, synthetic fixtures, aggregate tables, queries, and verification manifests are separated for release.

3.2 Positional author–affiliation–country alignment

The core parsing rule is deliberately conservative. For each record, the ordered Author ID, affiliation, and country arrays are aligned by position, retaining empty slots. A country is assigned to an author only when the record structure supplies a defensible author-specific link. When arrays are unequal or a slot is missing, the observation is flagged as ambiguous and cannot establish target entry, terminal foreign-affiliation transition, destination, or return. The set of all countries on a publication is never broadcast to every coauthor.

Annual author states are target-only, target-and-foreign, foreign-only, and country-unknown. A target-and-foreign state records co-affiliation but does not establish physical location. A return is observed when a target-country affiliation reappears after an intervening foreign-only publication year.

3.3 Outcomes and event timing

Entry is the first clearly linked target-country publication year. The primary first-transition event is the first later year with a clearly linked foreign-only publication state; later target-country reappearance is retained as return and does not revoke the event. The primary observation-process event occurs when three consecutive calendar years have elapsed without an indexed publication after a publication year; later publication is retained as resumption and does not revoke the completed-gap event. When both are possible, the first occurring event determines the prospective competing-event endpoint. Administrative censoring occurs at the end of 2025.

The prospective completed-gap event is dated at the third publication-free year using only information available by that date. It does not require the preceding publication to remain the final publication through 2025. Thus a researcher who publishes in year t, remains unobserved in t + 1 through t + 3, and publishes again in t + 4 experiences a completed-gap event at t + 3 followed by publication resumption at t + 4. Two- and five-year gaps remain sensitivity analyses.

The primary annual representation is nonabsorbing: first foreign-only transition may be followed by target-country return, and first completed gap may be followed by publication resumption. A separate secondary analysis classifies terminal foreign-affiliation transition and terminal disappearance using the complete record through 2025. Both secondary terminal endpoints are retrospective and window-defined because later observations can alter their classification. Accordingly, the first-event Aalen–Johansen estimand and the annual transition matrix are not competing summaries: the former estimates the first observed event, while the latter restores later nonabsorbing trajectories.

3.4 Secondary terminal-status and external-status sensitivity analyses

The secondary terminal-transition classification requires a foreign-only publication after the final target-containing year and no later target-containing publication through 2025. Terminal disappearance requires a completed gap that remains unresolved at window end. Both use later information and are reported only as fixed-window sensitivity descriptions. A permissive last-affiliation ceiling and ORCID-informed current-status scenarios are likewise assumption-dependent classifications, not population identification bounds.

3.5 Field and academic age

The Scopus Author Retrieval API was queried in batches for all 49,270 unique Author IDs using the ENHANCED view. Every profile returned a publication range; 49,581 of 49,587 country spells received a nonunknown dominant ASJC field. Academic age at entry is entry year minus the profile's first publication year and is winsorized at the 99th percentile (30.14 years) to limit the influence of implausibly long or merged histories. Broad ASJC field is derived from the leading two digits of the profile's frequency-weighted dominant ASJC code. Because current profiles incorporate the author's accumulated career, field adjustment is a sensitivity control rather than a strictly pre-entry characteristic; entry-year publication covariates remain fixed at baseline.

3.6 Descriptive competing-event analysis

Discrete-time Aalen–Johansen estimates are primary for the first of two prospective events: first foreign-only publication state or first completed three-year publication gap. Ninety-five percent intervals use 1,000 author-level bootstrap replicates within country. Risk sets are reported at 1, 3, 5, and 10 years. Later target return and publication resumption are summarized as subsequent transitions. Retrospective terminal-status curves through 2025 and a 2023 window-end analysis are secondary sensitivities and are explicitly labeled fixed-window descriptions. The Aalen–Johansen analysis estimates the first observed event only: once a spell first reaches foreign-only publication or a completed three-year gap, that event remains in its cumulative-incidence history. The separate annual transition

matrix restores the later nonabsorbing trajectories by showing target-country return after foreign-only publication and publication resumption after a completed gap.

3.7 Discrete-time models

Annual timing creates tied event times by design. We therefore replace continuous-time Cox models (Cox, 1972) with person-year generalized linear models using a complementary log-log link, a standard formulation for discrete event-history data (Tutz & Schmid, 2016). Separate models predict terminal foreign-affiliation transition and publication disappearance. Each includes flexible year-since-entry indicators, country-by-time interactions, standardized log entry-year target publications, entry-year international-collaboration share, standardized log mean team size, entry cohort, entry co-affiliation, winsorized academic age, and broad ASJC field. Standard errors are clustered by Author ID, accounting for repeated person-years and cross-country spells. Country-by-time interactions avoid imposing a single proportional country coefficient.

A second visibility-conditioned model uses only occasions on which another indexed publication is observed after a target-containing publication and predicts whether the next observed state is foreign-only. It includes the same adjustment set. Country-by-time interactions are evaluated jointly with Author-ID-cluster-robust Wald tests; time-specific adjusted odds-ratio contrasts with 95% confidence intervals are reported for interpretation. This estimand does not recover physical departure, but tests whether country contrasts persist conditional on continued bibliographic visibility.

3.8 ORCID external-status sensitivity sample

A reproducible stratified random sample was drawn from publication-disappearance spells using seed 20260825. To obtain 75 Scopus-linked ORCID cases per country, the deterministic batch search screened 2,575 of 6,267 Oman cases, 2,175 of 6,133 Qatar cases, and 2,350 of 16,328 UAE cases in randomized order, using 25-ID batches and retaining the first 75 linked records per country. The final N = 225 public ORCID records were processed by a rule-based country classifier into current foreign-only, target-only, target-and-foreign, or no public current employment; ambiguous or absent current employment was not manually recoded. Author-level data remain restricted because they contain persistent identifiers.

This sample is external to the publication-affiliation trace but is neither representative of disappeared researchers nor temporally aligned with disappearance. It is conditional on a Scopus-linked ORCID iD and public self-reported employment, and status is observed in 2026 rather than at the event year. A target-containing record may be stale, honorary, nonresidential, or follow a departure and return; a foreign-only record may begin after the ten-year horizon. Accordingly, the exercise cannot validate event-time departure. Exploratory scenarios count current foreign-only employment in the lower reclassification and additionally count missing current employment in the upper reclassification. Exact Clopper–Pearson intervals quantify sample uncertainty but not ORCID selection. Duration-specific 1–5, 6–10, and 11 + year tabulations are reported without using sparse strata as population calibration weights. ORCID adoption varies strongly by discipline and rank, and many profiles remain incomplete (Porter et al., 2025).

3.9 Verification and sensitivity analyses

Tests cover CSV/XLSX ingestion, EID/DOI deduplication, empty-slot preservation, nonbroadcasting of country sets, return detection, and hand-calculated event estimators. Every terminal foreign transition requires clear target evidence at the final target year and clear foreign evidence at the transition year. Record-level validation used a reproducible diagnostic sample of 175 records: 25 records from each of seven prespecified strata (unequal or missing arrays, observed return, foreign-only, terminal transition, first foreign-only transition, multiple affiliations, and target-only),

sampled with seed 20260830 and shuffled with seed 20260831. The sample contained 56 Oman, 53 Qatar, and 66 UAE records and therefore diagnoses parser behavior in difficult and substantively important cases rather than estimating population-wide error prevalence.

The validation used a browser-assisted first pass followed by an independent human inspection of each visible Scopus record, including the author superscripts and the Author affiliations display. The reviewer coded the observed annual state, author-affiliation alignment, country mapping, and error category using the row-specific target country. All 75 records with at least one discrepant field were adjudicated from visible Scopus evidence. Twelve observed-state disagreements were reopened directly in Scopus; the remaining discrepancies were resolved using the human reviewer's detailed record notes together with saved first-pass browser evidence. The separate targeted reconciliation of 11 previously unresolved Oman Author IDs also remained in place: all profiles resolved, merging 532 directly retrieved records and rerunning the 2000–2025 analysis left the core outputs byte-identical (Supplementary Table S8).

The seed/full-history audit identified complete-history target author-document pairs absent from the country seed: 197 in Oman (0.39%), 399 in Qatar (0.38%), and 2,737 in the UAE (1.41%). The UAE miss rate was 3.7 times Oman's and Qatar's, consistent with systematic differences such as branch-campus names or affiliation-ID variants rather than homogeneous retrieval noise. Of the missing pairs, 162/35/0 in Oman, 340/59/0 in Qatar, and 2,266/471/0 in the UAE fell before/at/after the author's complete-history final target year. Because outcomes use the complete-history files, no pair occurred after that endpoint and the union extended the final target year for zero authors. Follow-up is also truncated at 2023 as an indexing-lag sensitivity analysis.

4. Results

4.1 Cohorts, fields, and observation endpoints

Medicine was the largest dominant field in all three cohorts (3,153 Oman spells, 5,013 Qatar spells, and 6,081 UAE spells). The next-largest profiles differed: Oman included Engineering (891), Social Sciences (852), and Energy (718); Qatar included Social Sciences (1,086), Engineering (1,027), and Computer Science (747); the UAE included Engineering (3,748), Social Sciences (2,634), Energy (2,438), and Computer Science (2,094). These contrasts substantiate the need for field adjustment.

In the prospective competing-event analysis, Oman recorded 1,349 first foreign-only events, 7,479 first completed gaps, and 1,114 administratively censored spells. Qatar recorded 2,652, 7,253, and 2,123; the UAE recorded 4,690, 18,770, and 4,157. Later publication resumed after 1,669 Oman, 1,454 Qatar, and 3,239 completed-gap events. Target-country affiliation returned after 353, 560, and 989 first foreign-only events. These subsequent transitions demonstrate why neither primary event is treated as terminal.

4.2 Primary prospective competing-event estimates

At ten years, first-foreign-only cumulative incidence was 14.3% (95% bootstrap CI [13.6%, 15.1%]) in Oman, 23.4% [22.6%, 24.1%] in Qatar, and 18.0% [17.5%, 18.5%] in the UAE. First-completed-gap incidence was 77.4% [76.5%, 78.4%], 62.8% [61.9%, 63.7%], and 70.6% [70.0%, 71.2%]. Event-free probability was 8.2%, 13.9%, and 11.4%. Table 1 reports the prospective estimates; later return and resumption are reported in Section 4.7.

In the secondary fixed-window analysis, terminal-transition cumulative incidence at ten years was 15.0% in Oman, 24.4% in Qatar, and 19.2% in the UAE. These retrospectively classified values are close to, but conceptually distinct from, the primary first-transition estimates. The permissive terminal-transition-to-last-affiliation ranges were 15.0%–79.2%, 24.4%–77.2%, and 19.2%–80.6% and are reported only as bibliographic sensitivity ranges.

4.3 Follow-up support and primary risk sets

The primary estimates in Table 1 retain every first event even when a later state is observed. At ten years, risk sets contained 333 Oman, 687 Qatar, and 1,169 UAE spells. Later cohorts contribute only to shorter durations; therefore, Fig. 1 displays prospective event accumulation and Fig. 2 displays cohort support rather than implying a population retention probability.

Table 1
Ten-year prospective first-event cumulative incidence and risk sets

Country	First foreign-only CIF (95% CI)	First completed-gap CIF (95% CI)	Event-free (95% CI)	At risk
Oman	14.3% (13.6%–15.1%)	77.4% (76.5%–78.4%)	8.2% (7.6%–8.9%)	333
Qatar	23.4% (22.6%–24.1%)	62.8% (61.9%–63.7%)	13.9% (13.1%–14.5%)	687
United Arab Emirates	18.0% (17.5%–18.5%)	70.6% (70.0%–71.2%)	11.4% (10.9%–11.8%)	1,169

Note. Percentile intervals use 1,000 author-level bootstrap replicates within country. Event-free is one minus the two cumulative incidences.

The numbers at risk decline with event occurrence and administrative follow-up. Ten-year risk sets were 333, 687, and 1,169 for Oman, Qatar, and the UAE. Figure 2 displays the cohort-specific support; no affiliation half-life is reported.

Pairwise bootstrap contrasts confirmed differences in the terminal-transition fixed-window cumulative incidence as an observed-data estimand: Oman minus Qatar was − 9.4 percentage points (95% CI [− 10.5, − 8.3]); Oman minus UAE was − 4.2 [− 5.1, − 3.3]; and Qatar minus UAE was 5.3 [4.3, 6.2]. These tests do not establish differences in physical departure because visibility remains part of the outcome.

4.4 ORCID-informed external-status sensitivity

Among 75 ORCID-linked disappearance cases per country, current foreign-only employment appeared for 7 Oman, 16 Qatar, and 12 UAE researchers. Current target-only employment appeared for 38, 26, and 31; target-and-foreign employment for 3, 4, and 2; and no public current employment for 27, 29, and 30. Across all 225 cases, 35 (15.6%) had current foreign-only employment, 104 (46.2%) retained a target-country employment link, and 86 (38.2%) were unresolved because no public current employment was available.

Applying the prespecified current-status rules to the secondary terminal analysis produced conditional reclassification ranges of 21.0%–44.1% for Oman, 35.7%–56.1% for Qatar, and 29.0%–53.6% for the UAE. Marginal Clopper–Pearson limits for each scenario component were propagated separately, yielding sample-uncertainty intervals of 17.5%–51.7%, 31.1%–62.0%, and 24.4%–60.6%. These are not joint 95% confidence regions and do not correct ORCID selection or temporal mismatch. Table 2 reports the country-specific scenario components and intervals, and Fig. 3 displays the corresponding bibliographic ceilings and ORCID-informed external-status sensitivity ranges.

Table 2
Secondary ORCID current-employment sensitivity scenarios

Country	Foreign only	Target + foreign	Target only	No public current	Conditional reclassification range	Sample-uncertainty interval
Oman	7	3	38	27	21.0%–44.1%	17.5%–51.7%
Qatar	16	4	26	29	35.7%–56.1%	31.1%–62.0%
United Arab Emirates	12	2	31	30	29.0%–53.6%	24.4%–60.6%

Note. n = 75 disappearance spells per country. Sample-uncertainty intervals propagate marginal exact Clopper–Pearson limits separately through the point disappearance estimate; they are not joint 95% confidence regions and do not account for ORCID selection.

The selection diagnostics reinforce caution. Among 86 records with no current employment classification, 69 had no public employment section and 17 contained only ended entries. Duration-specific current foreign-only counts were Oman 5/39, 2/31, and 0/5; Qatar 9/38, 6/29, and 1/8; and UAE 12/53, 0/17, and 0/5 for disappearance at 1–5, 6–10, and 11 + years after entry. Exact intervals were wide, including 0%–52.2% for each 0/5 cell. The heterogeneity and sparse late strata preclude exchangeable event-time calibration; pooled scenarios are retained only as exploratory summaries.

4.5 Visibility-conditioned comparison

The next-publication analysis contained 94,128 transition opportunities from 28,448 authors. Relative to Oman, adjusted Qatar odds ratios at years 1, 3, 5, and 10 were 0.95 (95% CI [0.85, 1.07]), 1.19 [0.99, 1.43], 1.13 [0.93, 1.37], and 1.09 [0.72, 1.64]; UAE ratios were 0.94 [0.85, 1.05], 0.99 [0.84, 1.18], 0.91 [0.77, 1.09], and 0.77 [0.52, 1.13]. Author-cluster-robust joint Wald tests rejected equality of all country-by-time terms for Qatar versus Oman ($\chi^2(15) = 26.03$, p=.038), UAE versus Oman ($\chi^2(15) = 32.87$, p=.005), and jointly ($\chi^2(30) = 76.27$, p<.001). The valid joint test replaces the correlated-interval sign test. It supports time-varying country differences conditional on another indexed publication, but not a stable rank ordering or a causal visibility mechanism. Figure 4 and Table 3 report selected contrasts; complete interval estimates appear in Supplementary Table S1.

Table 3
Selected adjusted visibility-conditioned country-by-time odds ratios

Country contrast	Year after entry	Adjusted OR	95% CI	Observed opportunities
Qatar vs Oman	1	0.95	0.85–1.07	16,822
Qatar vs Oman	3	1.19	0.99–1.43	11,975
Qatar vs Oman	5	1.13	0.93–1.37	10,216
Qatar vs Oman	10	1.09	0.72–1.64	2,801
United Arab Emirates vs Oman	1	0.94	0.85–1.05	16,822
United Arab Emirates vs Oman	3	0.99	0.84–1.18	11,975
United Arab Emirates vs Oman	5	0.91	0.77–1.09	10,216
United Arab Emirates vs Oman	10	0.77	0.52–1.13	2,801

Entry-year international collaboration remained associated with a foreign-only next publication (OR 1.09, 95% CI 1.06–1.12), as did older academic age at entry (OR 1.30, 95% CI 1.27–1.33) and entry co-affiliation (OR 1.53, 95% CI 1.44–1.64). More entry-year target publications predicted a lower probability (OR 0.72, 95% CI 0.70–0.74). These are publication-state associations, not causal mobility effects.

4.6 Field- and age-adjusted discrete-time models

In the field- and age-adjusted discrete-time terminal-transition model, a one-standard-deviation increase in entry-year international-collaboration share was associated with an 18% higher terminal-transition hazard (HR 1.18, 95% CI 1.15–1.21). More entry-year target publications (HR 0.89, 95% CI 0.87–0.91) and larger entry teams (HR 0.96, 95% CI 0.93–0.98) predicted lower hazards. Entry co-affiliation (HR 1.76, 95% CI 1.65–1.87) and academic age (HR 1.43, 95% CI 1.41–1.46) predicted greater observability of a later terminal foreign-only publication. The estimation unit was 49,587 country-specific spells from 49,270 unique Author IDs, expanded to 244,845 person-years and clustered by Author ID.

For publication disappearance, entry publications (HR 0.66, 95% CI 0.65–0.67), international collaboration (HR 0.93, 95% CI 0.92–0.94), entry co-affiliation (HR 0.60, 95% CI 0.56–0.64), and academic age (HR 0.56, 95% CI 0.54–0.58) were associated with lower hazards, while larger teams predicted a higher hazard (HR 1.11, 95% CI 1.09–1.12). Table 4 reports reader-facing covariate labels; the full field coefficients (reference: Medicine) and the no-academic-age sensitivity appear in Supplementary Tables S2–S3.

Table 4
Discrete-time complementary log-log models

Outcome	Covariate	Hazard ratio	95% CI	Person-years	Spells	Authors
Terminal foreign-affiliation transition	Entry publications (log, z)	0.89	0.87–0.91	244,845	49,587	49,270
Terminal foreign-affiliation transition	International collaboration (z)	1.18	1.15–1.21	244,845	49,587	49,270
Terminal foreign-affiliation transition	Mean team size (log, z)	0.96	0.93–0.98	244,845	49,587	49,270
Terminal foreign-affiliation transition	Entry co-affiliation	1.76	1.65–1.87	244,845	49,587	49,270
Terminal foreign-affiliation transition	Academic age (winsorized, z)	1.43	1.41–1.46	244,845	49,587	49,270
Disappearance	Entry publications (log, z)	0.66	0.65–0.67	244,845	49,587	49,270
Disappearance	International collaboration (z)	0.93	0.92–0.94	244,845	49,587	49,270
Disappearance	Mean team size (log, z)	1.11	1.09–1.12	244,845	49,587	49,270
Disappearance	Entry co-affiliation	0.60	0.56–0.64	244,845	49,587	49,270
Disappearance	Academic age (winsorized, z)	0.56	0.54–0.58	244,845	49,587	49,270

Note. Models also include flexible time, country-by-time interactions, cohort, and broad ASJC field. Standard errors are clustered by Author ID. No causal interpretation is intended.

4.7 Circulation and returns

Among primary first foreign-only events, later target-country return occurred in 353 of 1,349 Oman spells (26.2%), 560 of 2,652 Qatar spells (21.1%), and 989 of 4,690 UAE spells (21.1%). Publication resumed after 1,669 of 7,479 Oman completed gaps (22.3%), 1,454 of 7,253 Qatar gaps (20.0%), and 3,239 of 18,770 UAE gaps (17.3%). Table 5 reports these subsequent transitions and Fig. 5 the annual state matrices. First destinations remain descriptive and appear in Supplementary Table S4. These displays answer complementary questions: the Aalen–Johansen estimates summarize which event is observed first, whereas the transition matrices retain the later movements that a first-event estimand cannot represent.

Table 5
Selected nonabsorbing transitions and returns

Country	Target only → target + foreign	Foreign only → target only	Foreign only → target + foreign	Returns among first foreign-only events, n (%)
Oman	774	275	406	353 (26.2%)
Qatar	1,671	352	996	560 (21.1%)
United Arab Emirates	2,826	709	1,387	989 (21.1%)

4.8 Sensitivity analyses

Changing the publication-free threshold alters the time and frequency of the prospective completed-gap event. Truncating follow-up at 2023 affects administrative support and the separate retrospective terminal classification. Supplementary Tables S5–S6 report these sensitivities with explicit estimand labels.

4.9 Record-level validation

The two passes agreed on observed state for 163 of 175 records (93.1%; Cohen's kappa = 0.892). After adjudication, the exported author-affiliation alignment was supported for 168 records (96.0%) and rejected for 7 (4.0%). Country mapping was supported for 149 records (85.1%), rejected for 12 (6.9%), and remained unclear from the visible display for 14 (8.0%); among the 161 determinate cases, 92.5% were correct. An observable state could be assigned in 166 records (94.9%), while 9 remained country-unknown. The final error categories were none (139), ambiguous display (13), wrong country (11), unequal arrays (6), and missing affiliation (6). Full inter-pass agreement and adjudicated counts appear in Supplementary Table S9. Because diagnostic strata were sampled equally, these percentages evaluate the selected validation set and are not population-weighted error rates.

Agreement was lower for the intermediate accuracy judgments: 69.7% for author-affiliation alignment and 76.6% for country mapping. The corresponding kappa for alignment was − 0.016 because almost all adjudicated decisions were Yes and the browser-assisted first pass was deliberately conservative; raw agreement and the adjudicated result are therefore more informative than kappa for that imbalanced item. No aggregate mobility estimate, table, or figure was changed by either the 175-record validation or the 11-profile reconciliation.

5. Discussion

The central contribution is to represent observed publication-affiliation histories without forcing them into terminal migrant/stayer labels. First foreign-only transition and first completed publication gap are prospective events that

remain recorded when return or resumption follows. Terminal 2025 status and ORCID current employment are secondary sensitivity analyses rather than formal bounds on migration.

The ORCID external-status sensitivity analysis is informative but does not validate event-time departure. Nearly half of the linked sample reported a current target-country employment link, a smaller group reported current foreign-only employment, and more than one-third lacked public current employment. This demonstrates heterogeneity in later external status after bibliographic disappearance. Because current status may follow departure, return, stale reporting, or delayed foreign employment, the resulting ranges remain conditional sensitivity scenarios rather than population bounds.

This distinction is consequential for country comparison. Qatar had the highest terminal-transition fixed-window cumulative incidence and Oman the lowest, while the disappearance ordering was nearly reversed. Terminal foreign-affiliation transition requires continued publication after departure, so systems whose researchers remain more visible can appear to have higher exit even when their physical departure probabilities are similar or lower.

The visibility-conditioned model clarifies why the descriptive ordering cannot be converted into a retention ranking. The joint country-by-time test was significant, yet time-specific adjusted contrasts were heterogeneous and often imprecise. The UAE's contrast was below Oman at several intervals despite its higher terminal-transition cumulative incidence, which is consistent with—though does not prove—country differences in the observation process. We therefore report country curves as database-observed trajectories rather than policy-performance rankings.

A further interpretive complication is disciplinary composition. Field adjustment materially improves comparability because the systems have distinct disciplinary portfolios. Medicine dominates each cohort, but engineering, energy, computer science, and social sciences have different weights. Publication frequency and database coverage vary across these fields, altering both the opportunity to observe an exit and the probability of disappearance (Mongeon & Paul-Hus, 2016). Arts and Humanities had the largest estimated exit hazard among identified fields (HR 1.86), which is coherent with the comparatively incomplete Scopus coverage documented for those fields: lower coverage can make affiliation histories appear to terminate or change more readily. The coefficient is therefore further evidence that the model captures visibility as well as mobility.

Academic age is consequential, but it is partly a visibility parameter rather than a clean career-stage exposure. Older publishing careers were more likely to generate an observable foreign-only transition and less likely to disappear, consistent with both accumulated networks and the continued publishing required to observe an exit. Because the current Scopus profile’s first publication year is vulnerable to split and merged identities, the model without academic age is an important sensitivity. Entry co-affiliation behaved similarly. Its large coefficient should not be read as evidence that dual affiliation causes migration; an existing foreign link makes a later foreign-only publication easier to observe and may persist in metadata.

Taken together, these findings have direct implications for research evaluation. National dashboards should separate at least four quantities: positively evidenced foreign transition, continued target participation, publication disappearance, and administrative censoring. Co-affiliation and return should be shown as states rather than forced into mutually exclusive migrant/stayer labels. Every chart should name its estimand; “terminal-transition fixed-window cumulative incidence” is defensible, whereas “emigration rate” is not.

The completed record-level validation supports the core state classification while also showing where uncertainty enters. Observed-state agreement was high, and adjudication supported 96.0% of author-affiliation alignments and 92.5% of determinate country mappings. The remaining problems clustered in ambiguous displays, wrong countries,

unequal arrays, and missing affiliations, especially in structurally difficult records. Because the sample deliberately balanced diagnostic strata and the first pass was browser-assisted rather than a second human coding, these values are evidence about parser behavior in the audited cases, not population-wide precision or recall. Future indicator production should repeat stratified human validation and, where feasible, link samples to institutional directories, CVs, or administrative personnel records.

The author-specific alignment rule is immediately portable. Many exports provide a publication-level list of countries but do not guarantee that each country belongs to each coauthor. Broadcasting that set fabricates international affiliations and transitions. Preserving positional arrays and uncertainty flags should be a minimum standard for researcher-level mobility analysis.

6. Limitations

Scopus Author IDs can split or merge careers, and current Author Retrieval profiles may reflect corrections made after the historical publications. Dominant ASJC is a whole-profile field, not a purely pre-entry characteristic; it is used as a compositional control, not a causal exposure. Publication-range start years may be affected by profile splitting or incomplete early coverage, although 99th-percentile winsorization limits extreme influence.

Affiliation indicates a publication address rather than residence, nationality, contractual employment, or physical presence. Honorary, visiting, and strategic affiliations cannot be distinguished consistently. The ORCID external-status sensitivity sample is conditional on a Scopus-linked ORCID iD and current public employment, and required screening 2,575, 2,175, and 2,350 randomized disappearance cases to retain 75 per country. Current status may be stale or may follow departure and return; it does not identify employment at the disappearance year. Duration-specific cells are sparse, and exact intervals expose but cannot remove selection or temporal mismatch.

Disappearance is an administrative endpoint dated only after a completed publication-free gap. The terminal foreign-transition class is also retrospective because later target reappearance through the window end can revoke it. The Aalen–Johansen curves should therefore be read as fixed-window trajectory summaries, while first transitions and returns are represented in the nonabsorbing annual multistate analysis. The 2023 truncation and alternative gaps assess, but cannot eliminate, window dependence and informative observation loss.

The discrete-time models are associational. Field, career stage, collaboration, and co-affiliation improve adjustment but do not remove unobserved differences in institution type, appointment status, nationality, funding, or selection into Gulf employment. Country-by-time interactions avoid a single proportional country coefficient, but other covariate effects are summarized multiplicatively and may still vary over time. The 175-record validation was diagnostic and stratified rather than prevalence-weighted; its first pass used browser-assisted validation with independent human verification, and only the 12 observed-state disagreements were reopened directly during adjudication. The remaining disagreement decisions relied on the independent human reviewer's detailed Scopus notes and saved first-pass evidence. These constraints limit claims about population-wide parser accuracy, although the final state agreement and adjudicated alignment results support the analytical rule in the audited cases.

7. Conclusion

Publication-affiliation traces reveal observable research-system participation, not complete migration histories. Prospective first events show that foreign-only publication transitions and completed publication gaps are often followed by return or resumption. Secondary terminal and ORCID analyses remain useful sensitivity descriptions but

do not identify physical departure. The defensible output is a multistate account of observable transitions with explicit uncertainty and observation limits.

The defensible policy product is therefore not a single retention percentage. It is a transparent decomposition of observed transition, continued participation, co-affiliation, return, disappearance, and censoring, accompanied by validation and explicit uncertainty. Mobility indicators become more useful—not less—when they state precisely what the data can and cannot identify.

Declarations

Competing interests

The author has no relevant financial or non-financial interests to disclose.

Ethics approval

and consent

Use of generative AI and AI-assisted technologies in manuscript preparation

OpenAI's ChatGPT and Codex were used to improve the clarity, readability, organization, and presentation quality of the manuscript. The author reviewed and verified all suggested revisions and takes full responsibility for the accuracy, originality, and integrity of the final content.

Funding

No funding was received for conducting this study or preparing this manuscript.

Author Contribution

The sole author was responsible for conceptualization, methodology, software, validation, formal analysis, investigation, data curation, visualization, writing—original draft, and writing—review and editing, and approved the final manuscript.

Data Availability

Licensed Scopus records, Author IDs, the 175-record Scopus audit workbook, and author-level ORCID external-status sensitivity data cannot be redistributed. The release package contains synthetic Scopus-format fixtures, aggregate country and transition tables, aggregate audit counts, figure-source tables, query templates, verification manifests, and instructions for authorized reconstruction. Only nonidentifying aggregate validation results are included in the public package.

Code availability

The Scopus Mobility/Transience Analyzer (version 0.2.1), including the source code, documentation, tests, synthetic test data, aggregate outputs, figure-source tables, query templates, and verification manifests, is openly available on Zenodo at https://doi.org/10.5281/zenodo.22228590 under the MIT License. Licensed Scopus records, API credentials, persistent author identifiers, and person-level ORCID data are not included in the public archive.

References

1. Allignol, A., Schumacher, M., & Beyersmann, J. (2011). Empirical transition matrix of multi-state models: The etm package. *Journal of Statistical Software*, *38*(4), 1–15. https://doi.org/10.18637/jss.v038.i04
2. Al-Marzouqi, A. H., & Arabi, A. A. (2022). Research performance of the GCC countries: A comparative analysis of quantity and quality. *Heliyon*, *8*(11), e11309. https://doi.org/10.1016/j.heliyon.2022.e11309
3. Aman, V. (2018). Does the Scopus author ID suffice to track scientific international mobility? A case study based on Leibniz laureates. *Scientometrics*, *117*(2), 705–720. https://doi.org/10.1007/s11192-018-2895-3
4. Appelt, S., van Beuzekom, B., Galindo-Rueda, F., & de Pinho, R. (2015). *Which factors influence the international mobility of research scientists?* (OECD Science, Technology and Industry Working Papers No. 2015/02). OECD Publishing. https://doi.org/10.1787/5js1tmrr2233-en
5. Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, *1*(1), 377–386. https://doi.org/10.1162/qss_a_00019
6. Carneiro, A. M., Pereira, C. A., Meireles, F., & Brito, R. R. (2026). Studying the scientific diaspora: Going beyond bibliometric methods. *Scientometrics*. Advance online publication. https://doi.org/10.1007/s11192-026-05723-w
7. Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–220. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x
8. Czaika, M., & Orazbayev, S. (2018). The globalisation of scientific mobility, 1970–2014. *Applied Geography*, *96*, 1–10. https://doi.org/10.1016/j.apgeog.2018.04.017
9. El-Ouahi, J., Robinson-Garcia, N., & Costas, R. (2021). Analyzing scientific mobility and collaboration in the Middle East and North Africa. *Quantitative Science Studies*, *2*(3), 1023–1047. https://doi.org/10.1162/qss_a_00149
10. Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, *94*(446), 496–509. https://doi.org/10.1080/01621459.1999.10474144
11. Franzoni, C., Scellato, G., & Stephan, P. (2012). Foreign-born scientists: Mobility patterns for 16 countries. *Nature Biotechnology*, *30*(12), 1250–1253. https://doi.org/10.1038/nbt.2449
12. Gomez, C. J., Herman, A. C., & Parigi, P. (2020). Moving more, but closer: Mapping the growing regionalization of global scientific mobility using ORCID. *Journal of Informetrics*, *14*(3), 101044. https://doi.org/10.1016/j.joi.2020.101044
13. Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53*(282), 457–481. https://doi.org/10.1080/01621459.1958.10501452
14. Latouche, A., Allignol, A., Beyersmann, J., Labopin, M., & Fine, J. P. (2013). A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *Journal of Clinical Epidemiology*, *66*(6), 648–653. https://doi.org/10.1016/j.jclinepi.2012.09.017
15. Mitchell, M., Johnson, S. L., & Porter, M. D. (2026). Affiliation errors and distorted researcher mobility: Evidence from OpenAlex and Scopus. *Scientometrics*, *131*(7), 4743–4762. https://doi.org/10.1007/s11192-026-05689-9

16. Moed, H. F., Aisati, M., & Plume, A. (2013). Studying scientific migration in Scopus. *Scientometrics*, *94*(3), 929–942. https://doi.org/10.1007/s11192-012-0783-9
17. Moed, H. F., & Halevi, G. (2014). A bibliometric approach to tracking international scientific migration. *Scientometrics*, *101*(3), 1987–2001. https://doi.org/10.1007/s11192-014-1307-6
18. Momeni, F., Karimi, F., Mayr, P., Peters, I., & Dietze, S. (2022). The many facets of academic mobility and its impact on scholars' career. *Journal of Informetrics*, *16*(2), 101280. https://doi.org/10.1016/j.joi.2022.101280
19. Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, *106*(1), 213–228. https://doi.org/10.1007/s11192-015-1765-5
20. Netz, N., Hampel, S., & Aman, V. (2020). What effects does international mobility have on scientists' careers? A systematic review. *Research Evaluation*, *29*(3), 327–351. https://doi.org/10.1093/reseval/rvaa007
21. OECD. (2026). *Bibliometric indicators of implied bilateral mobility flows* [Data set]. OECD Data Explorer. Retrieved August 25, 2026, from https://data-explorer.oecd.org/vis?df%5Bag%5D=OECD.STI.STP&df%5Bds%5D=DisseminateFinalDMZ&df%5Bid%5D=DSD_BIBLIO_F%40DF_BIBLIO_F
22. Paul, A. M., Hamid, Z., Kuehne, K. W., Sahakyan, M., & AlShebli, B. (2026). The role of institutional context in shaping migrant scientists' new research collaborations. *Journal of the Knowledge Economy*. Advance online publication. https://doi.org/10.1007/s13132-026-03340-9
23. Porter, S. R., Umbach, P. D., & Willis, C. (2025). Understanding ORCID adoption among academic researchers. *Scientometrics*, *130*(5), 2783–2797. https://doi.org/10.1007/s11192-025-05300-7
24. Purnell, P. J. (2022). The prevalence and impact of university affiliation discrepancies between four bibliographic databases—Scopus, Web of Science, Dimensions, and Microsoft Academic. *Quantitative Science Studies*, *3*(1), 99–121. https://doi.org/10.1162/qss_a_00175
25. Robinson-Garcia, N., Sugimoto, C. R., Murray, D., Yegros-Yegros, A., Larivière, V., & Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics*, *13*(1), 50–63. https://doi.org/10.1016/j.joi.2018.11.002
26. Sugimoto, C. R., Robinson-Garcia, N., Murray, D. S., Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, *550*(7674), 29–31. https://doi.org/10.1038/550029a
27. Tutz, G., & Schmid, M. (2016). *Modeling discrete time-to-event data*. Springer. https://doi.org/10.1007/978-3-319-28158-2
28. Woolley, R., & Turpin, T. (2009). CV analysis as a complementary methodological approach: Investigating the mobility of Australian scientists. *Research Evaluation*, *18*(2), 143–151. https://doi.org/10.3152/095820209X441808
29. Zhao, X., Aref, S., Zagheni, E., & Stecklov, G. (2022). Return migration of German-affiliated researchers: Analyzing departure and return by gender, cohort, and discipline using Scopus bibliometric data 1996–2020. *Scientometrics*, *127*(12), 7707–7729. https://doi.org/10.1007/s11192-022-04351-4

Figures

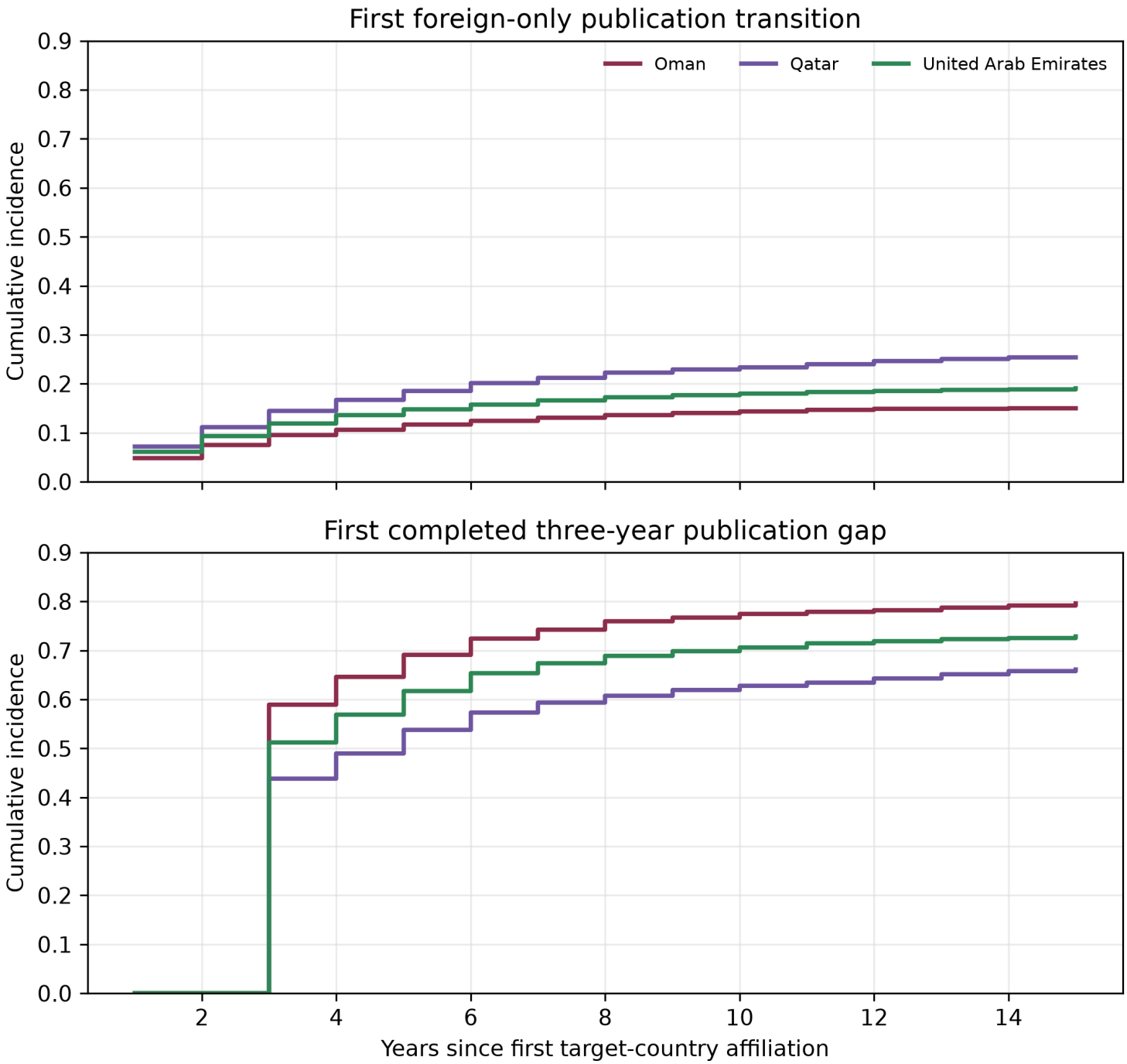


Figure 1

Prospective cumulative incidence of first foreign-only transition and first completed publication gap

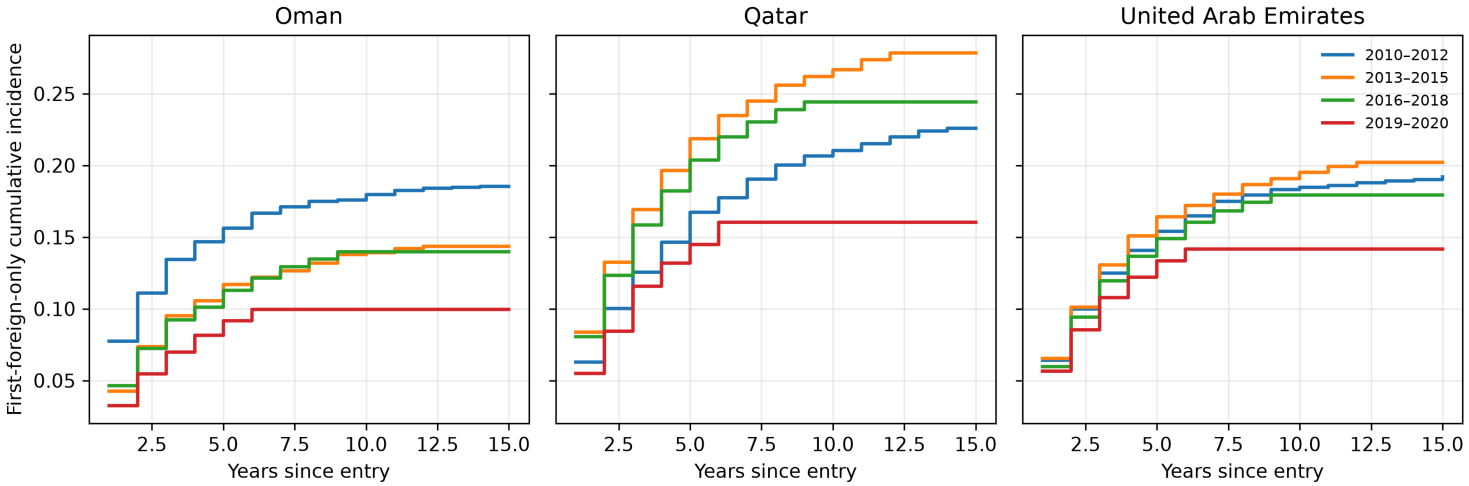


Figure 2

First-foreign-only cumulative incidence by entry cohort

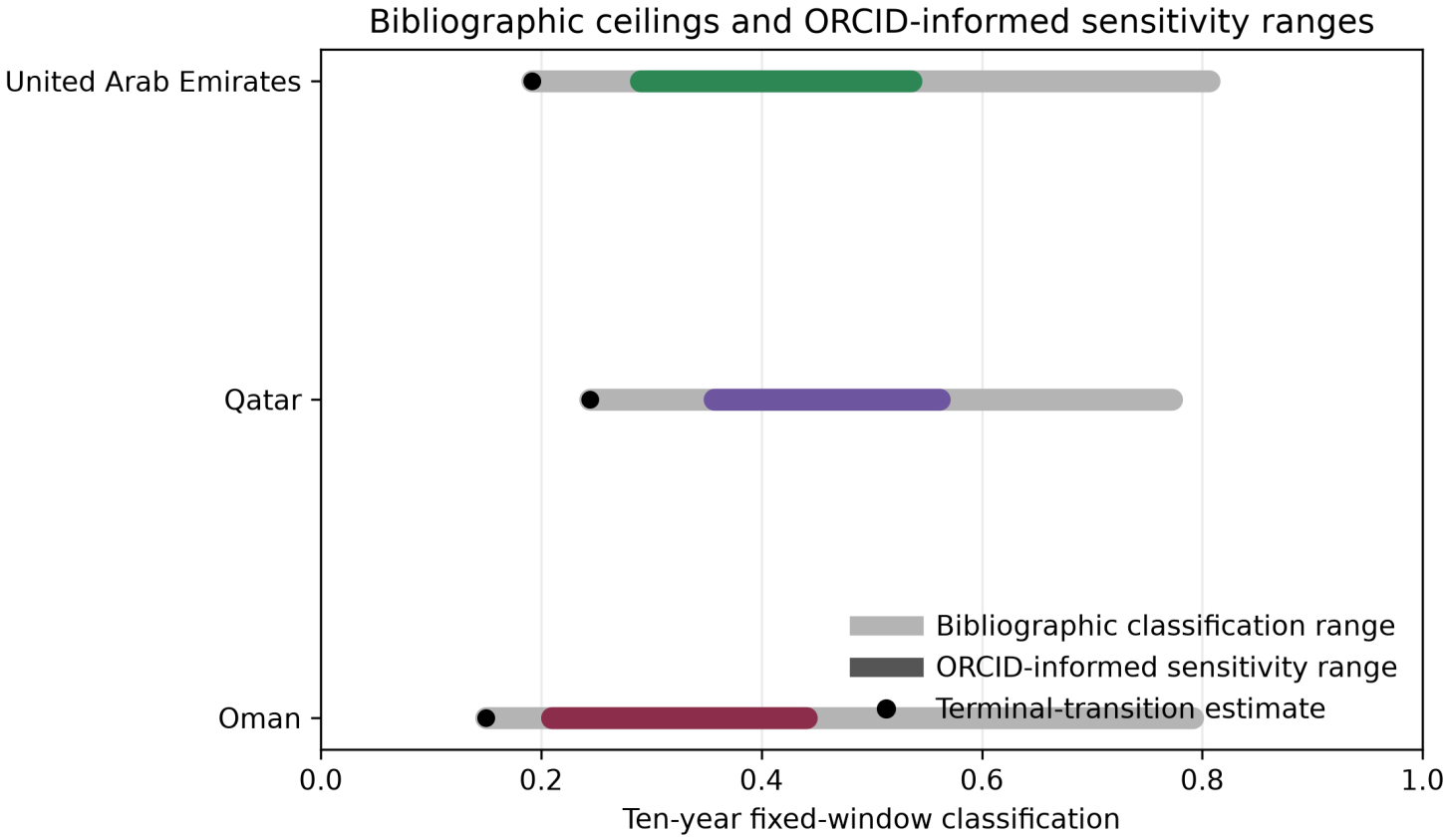


Figure 3

Bibliographic ceilings and ORCID-informed sensitivity ranges

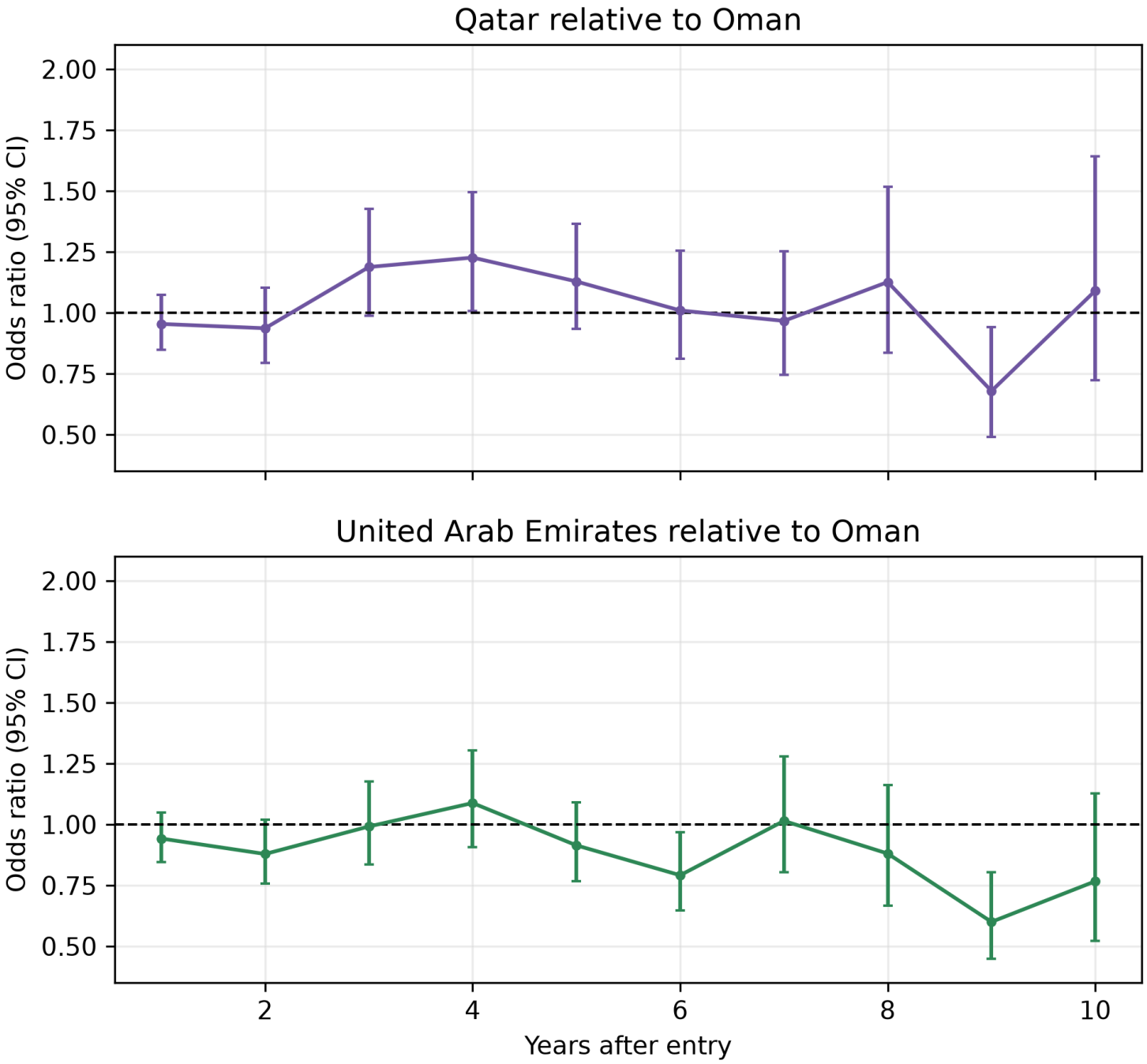


Figure 4

Adjusted visibility-conditioned country contrasts during the first ten years after entry

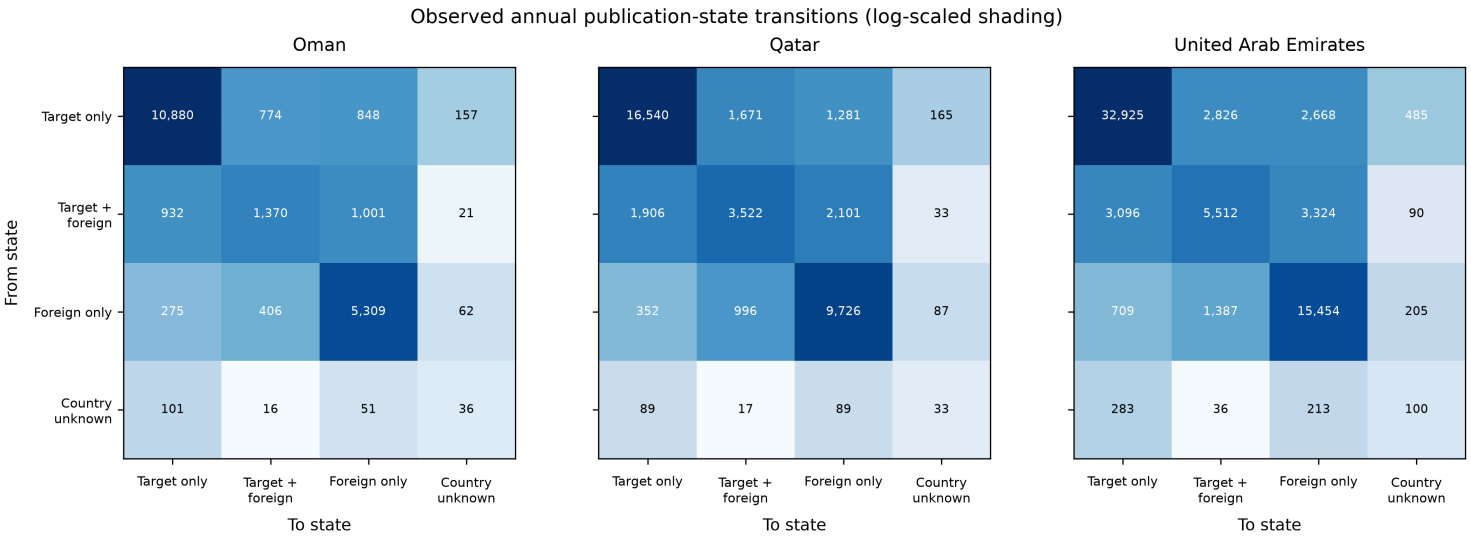


Figure 5

Observed annual publication-state transitions

Supplementary Files

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

- 03SupplementaryMaterialQSS.docx