Brain Circulation and Knowledge Creation: Evidence from Elite Chinese Scientists

Shihang Ru* Heiwai Tang† Bernard Yeung‡ Xinbei Zhou§

October 2026

Abstract

This paper examines how brain circulation contributes to frontier scientific development by analyzing the return of overseas Chinese scientists from 2003 to 2022. Using bibliometric data from leading scientific journals, we document China's rising presence in global science alongside a substantial influx of returnees, particularly from the United States. We show that returnees are positively selected for research quality, while their post-return performance varies substantially across fields. Local scholars exhibit a persistent premium in cumulative citation-share-based quality after first collaborating with returnees, including among scholars already internationally connected. Institutional exposure is associated with suggestive quality gains among scholars with no prior returnee or non-local international collaboration. At receiving institutions, predicted exposure is positively associated with incumbent quality, research participation, and output. These findings are consistent with contributions through joint research and broader institutional channels.

JEL Classification: F22, J61, O31, O33, I23

Keywords: Return migration; high-skilled migration; scientific productivity; knowledge spillovers; research collaboration; Chinese science

*Yale University, New Haven, United States. (haley.ru@yale.edu)

†HKU Business School, The University of Hong Kong, Hong Kong SAR, China. (tangecon@hku.hk)

‡National University of Singapore, Singapore; Southern University of Science and Technology, China. (byeung@nus.edu.sg)

§HKU Business School, The University of Hong Kong, Hong Kong SAR, China. (xbeizhou@connect.hku.hk).

"When our thousands of Chinese students abroad return home, you will see how China will transform itself."

— Deng Xiaoping (as recounted in Robert Lenzner, "The China Hand," Forbes, 31 October 2005, 79)

1 Introduction

No country has more at stake in knowledge migration than China, which has sent more students abroad than any other country since the reform era reopened overseas study in 1978.[1] Historically, the flow ran overwhelmingly one way: only a minority of those who went abroad had returned by the early 2000s (Cao, 2008)—especially the scientists: 92% of the Chinese who earned U.S. science and engineering doctorates in 2002 remained in the United States in 2007, the highest stay rate of any country, on the eve of China's talent programs (Finn, 2010). Stay rates nevertheless remain high: among temporary-visa holders from China who earned U.S. science and engineering doctorates in 2017–2019, 83% were still residing in the United States in 2023, approximately five years after graduation (National Center for Science and Engineering Statistics, 2025).

These stakes turn on an old question: when a country's brightest minds leave, does knowledge leave with them—or come back multiplied? Talent is more productive at the frontier, so departures drain; but returnees carry the frontier home, a phenomenon known as "brain circulation" in the social sciences (Prato, 2025; Saxenian, 2005).[2] Whether brain circulation can offset the losses from brain drain is therefore an empirical question, shaped by both the scientists who return and the environments that receive them.

Against this backdrop, China rolled out a series of policies to draw these scientists back, following a familiar playbook at an unfamiliar scale.[3] The 2006 *Medium- and Long-Term*

[1] Roughly 6.6 million Chinese students had studied overseas by 2019, and China remains the largest origin country of internationally mobile students today; of those who had completed their studies by 2019, 86% had returned—a cumulative share driven largely by the high return rates of recent, much larger cohorts (Ministry of Education of the People's Republic of China, 2020; UNESCO, 2025).

[2] Frontier location raises scientists' output, particularly for those from lower-income countries (Agarwal and Gaule, 2020; Kahn and MacGarvie, 2016). Emigration can also benefit origin countries without any return, by raising the payoff to human capital investment at home (Beine et al., 2008) and through diaspora networks that transmit knowledge across borders (Agrawal et al., 2011; Kerr, 2008; Zweig et al., 2004); see Docquier and Rapoport (2012) and Gibson and McKenzie (2011) for surveys, and Shin and Moon (2018) on the term "brain circulation." In our context, the circulating "brains" are scientists who studied or worked at overseas institutions and published in high-quality scientific outlets.

[3] South Korea's state-led repatriation of overseas scientists from the 1960s and Taiwan's recruitment of U.S.-trained engineers into Hsinchu are the canonical precedents (Saxenian, 2005; Yoon, 1992); return

Plan established the broader policy agenda (Cao et al., 2006), and two flagship recruitment programs put it into practice. The *Thousand Talents Program*, launched in 2008, targeted established Chinese-born scholars at leading overseas universities, while the *Young Thousand Talents Program*, introduced in 2011, extended the model to researchers under age 40, typically at the postdoctoral or assistant-professor stage (Jia, 2018). Together, these programs sought to bring not only scientists, but also their overseas training, frontier experience, and international networks into Chinese research institutions (Zweig et al., 2004)

This recruitment drive unfolded amid a broader transformation of science in China. Over the past two decades, China has moved from the periphery of elite science to its center. Since 2016, it has been the world's largest producer of scientific papers (National Science Board, 2018), and since 2023 it has led the Nature Index annual tables in the natural sciences (Woolston, 2023). Whether, and through which channels, the returning diaspora contributed to this rise is the subject of this paper. Specifically, we ask three sets of questions. First, is China's rise reflected in the growing presence of Chinese scholars in elite journals, and has this rise been accompanied by an increasing flow of returnees? Second, does China attract its top researchers back? Where do returnees lie in the global research-quality distribution, and how do they perform after returning? Third, do returnees improve science in China? If so, do they do so through direct collaboration with local coauthors or through broader institutional spillovers?

To study these questions, we use bibliometric data from leading scientific journals to document the rise in the quantity, prominence, and quality of scientific output by Chinese authors from 2003 to 2022. [4] Our sample comprises journals included in the Nature Index.[5] We focus on the five highest-ranked journals by the 2022 Journal Impact Factor (JIF) in the Journal Citation Reports[6] with complete Scopus[7] coverage over the sample period in Chemistry, Physics, and Life Sciences, and three multidisciplinary journals spanning all

incentives remain a live policy instrument elsewhere (Bassetto and Ippedico, 2024).

[4]In the final years of our sample, rising US–China tensions and the tightening research-security environment in the United States may also have pushed some scholars toward return (Flynn et al., 2024; Jia et al., 2024; Kitajima and Okamura, 2025; Xie et al., 2023); disentangling this push from China's pull requires a different empirical design, and we leave it to future work.

[5]The Nature Index tracks institutional contributions to research articles published in a selected list of high-quality journals using "Share" and "Count" metrics, updated monthly through a twelve-month rolling window. For the "Share" metric, each article's contribution is split equally between all authors, adjusted for multiple affiliations, to measure proportional contribution. The included journals are selected by an independent panel of active leading researchers through a reputation-based survey, regardless of publisher. The current revised list was released in June 2023, including 145 journals in natural and health sciences. See *https://www.nature.com/nature-index/faq* for more details.

[6]https://jcr.clarivate.com/jcr/browse-journals.

[7]https://www.scopus.com/

scientific fields, which we label as "General". [8] By focusing on these elite journals, we capture contributions at the forefront of science. From these records, we construct the paper's central measure of scholar quality: an author's cumulative citation share—the citations earned by her papers, split equally among coauthors and expressed as a share of all citations to papers published in the same years. Adjusting for coauthorship prevents double-counting credit, and normalizing within publication year makes quality comparable across cohorts whose papers have had different lengths of time to accumulate citations.

We begin by establishing four stylized facts.[9] First, Chinese scholars' presence, defined as the share of papers in these elite outlets with at least one China-affiliated co-author, rose sharply from 2003 to 2022, particularly in Chemistry. Second, Chinese scholars affiliated with Chinese institutions account for most of the increase in Chinese authors' citation-share contribution. Third, high-quality Chinese scholars became increasingly likely to return from the U.S. to China, a pattern driven by Chemistry, absent in Physics, and reversed in Life Sciences. Finally, post-return performance is also field-specific: returning chemists outperform both matched U.S.-based comparison groups, returning life scientists and General-journal authors underperform, and returning physicists perform similarly to ethnically Chinese U.S. stayers while showing somewhat weaker performance relative to non-Chinese stayers. These patterns show that China's rise in frontier science coincided with a growing and, in aggregate, increasingly positively selected returnee flow, while revealing sharp cross-field differences in who returned and how they performed.

We then turn to how returnees affect the local scientific communities they join. The first channel is direct collaboration. Our main event studies distinguish scholars who already had non-local international coauthors before their first returnee collaboration from those who had not yet collaborated internationally. In the prior-international group, internationally connected scholars who never coauthor with a returnee serve as controls, allowing us to examine the returnee-collaboration increment conditional on existing international ties and network evolution. We use Sun–Abraham as our preferred estimator, report TWFE as a benchmark, and examine Callaway–Sant'Anna and custom pairwise comparisons in the appendix. Both groups exhibit positive post-collaboration coefficients, and the displayed leads are jointly insignificant under the preferred and TWFE estimators. Among already internationally connected scholars, the preferred coefficient rises from 0.077 in the collaboration year to 0.259 at $\tau = 3$, corresponding to cumulative Author Quality approximately 8.0% and 29.6% higher relative to the comparison path. When returnee-coauthored papers are

[8] Appendix Table A1 reports the complete list of journals.

[9] Throughout the paper, "Chinese" applied to individual scholars or authors denotes name-based ethnicity, independent of location or citizenship, while "China-affiliated" and "Chinese institution" refer to institutional affiliation in a given year. Section 3 defines both categorizations precisely.

excluded, the contemporaneous increase disappears in both groups, while the later non-joint trajectory is sensitive to network adjustment and comparison construction. The evidence therefore points most clearly to gains embodied in joint research.

A second channel operates through broader exposure within institutions. We consider local scholars who, as of the preceding year, had no observed returnee or non-local international coauthorship. This baseline risk set can include the year of first collaboration, so we also examine a stricter sample censored before either network entry. Because returnees disproportionately sort into stronger institutions, simple comparisons would confound returnee exposure with preexisting institutional quality. We construct a capacity-aligned instrument by interacting each institution's pre-policy share of national publication capacity within a field with cumulative post-2009 returnee inflows in that field, excluding arrivals at the institution itself. Our preferred specification jointly absorbs author-by-field, institution-by-field, and field-by-year fixed effects. The 2SLS estimate is 0.122 (s.e. 0.061, $p = 0.047$) among local scholars without prior collaboration, while the Anderson–Rubin test has $p = 0.070$. Using the same baseline specification among already-connected scholars yields 0.042 (s.e. 0.078); a joint test does not reject equality with the unconnected-sample coefficient ($p = 0.429$). Adding collaboration history and annual intensity leaves the connected-sample estimate small and imprecise at 0.034 (s.e. 0.076, $p = 0.650$). At the institution level, predicted exposure is positively associated with incumbent quality, active local authors, annual entrants into the observed publication panel, active post-2008 participants, fractional publication volume, and citation-weighted output in reduced-form level specifications.

The evidence therefore suggests that brain circulation can contribute to scientific performance beyond the publications of returning scientists alone. The clearest collaborator gains are embodied in joint research. We also find suggestive evidence of institutional quality gains outside prior collaboration networks. The institution-level results extend this picture to observed participation and aggregate research output. These reduced-form associations are consistent with broader institutional channels.

This paper contributes to three related literatures. First, it adds to the work documenting China's growing role in global science. Earlier studies show the rapid expansion of Chinese scientific output and R&D capacity (Tollefson, 2018; Xie and Freeman, 2018; Xie et al., 2014; Zhou and Leydesdorff, 2006). Recent work shows that this rise extends to elite outlets, especially in engineering and the physical sciences, but also emphasizes field heterogeneity and uneven international uptake: Nagaraj and Yao (2026) document China's growing share of top-journal publications while showing that Chinese research remains less globally integrated in topics and citations, and Qiu et al. (2022) document cross-border citation frictions in Chemistry. Related measurement work cautions that citation-based rankings are sensitive to

field normalization and home-country citation bias (Qiu et al., 2024; Wagner et al., 2022). We provide complementary evidence by focusing on elite scientific journals across major natural science fields and by linking China's growing frontier presence, as proxied by publication citation shares, to the mobility, selection, and post-return performance of Chinese scientists.

Second, the paper contributes to research on Chinese diaspora scientists, return migration, and brain circulation. Early work emphasizes that overseas study and work generate "transnational human capital" through foreign training, technology transfer, and international networks (Zweig et al., 2004). Subsequent studies document both the ambition and limits of China's efforts to convert high-end brain drain into brain gain (Cao, 2008; Zweig and Wang, 2013; Zweig et al., 2020). More recent bibliometric evidence shows that returning scientists publish higher-impact work than domestic counterparts and help embed China in international collaboration networks (Cao et al., 2020), while related studies examine Chinese diaspora scientists (Xie and Freeman, 2020, 2023) and overseas returnees (Shi et al., 2023). Related quasi-experimental evidence links the return of U.S.-trained African scientists to higher publication output among local colleagues (Fry, 2023) and temporary international mobility among Chinese doctoral students to larger collaboration networks and greater research output (Gu et al., 2024). Most closely related, Giorcelli et al. (2026) use historical exit restrictions and bilingual bibliometric data to study how the Great Return Wave of 1940–1957 shaped Chinese scientific output and field development. We add to this literature by studying contemporary return migration and jointly measuring returnee selection, post-return performance, direct collaboration effects on China-based scholars, and broader institutional exposure, while also accounting for changes in the observed local-author pool.

Third, the paper relates to the economics literature on high-skilled migration, scientific networks, and innovation (Docquier and Rapoport, 2012). Studies of forced migration from Nazi Germany show that the loss of scientific human capital depressed research output in sending countries (Waldinger, 2016) while émigrés reshaped invention in receiving countries (Moser et al., 2014), and Becker et al. (2024) show that professional networks facilitated high-skilled emigration by transmitting non-public information across destinations. A recent review of causal evidence finds that high-skilled emigration can benefit origin countries through human-capital formation, linkages, knowledge transfer, and return migration, with effects that depend on skill type and origin-country institutions (Batista et al., 2025), and greater visibility of scientific achievement can reduce information frictions in the allocation of scientific talent (Hager et al., 2024). More broadly, high-skilled migrants and inventors are internationally mobile (Akcigit et al., 2016) and contribute to innovation in destination economies (Gaulé and Piacentini, 2013; Hunt and Gauthier-Loiselle, 2010; Prato, 2025; Terry et al., 2026), though policy-induced inflows do not mechanically translate into firm-level innova-

tion gains (Doran et al., 2022), and frontier location itself shapes scientists' productivity and return decisions (Gaulé, 2014; Kahn and MacGarvie, 2016). Our paper differs by studying return migration to a rapidly upgrading origin country and by estimating not only movers' own performance, but also their effects on local collaborators and recipient institutions.

The rest of the paper is organized as follows. Section 2 provides institutional background on China's talent policies. Section 3 describes the data and presents the stylized facts. Section 4 presents the main empirical analysis: direct collaboration, exposure among locals without prior collaboration, exposure among already-connected scholars, and institution-level research quality and output. Section 5 concludes.

2 Institutional Background

In 2006, China launched the *Medium- and Long-Term Plan for the Development of Science and Technology (2006–2020)* (MLP), committing the country to becoming an "innovation-oriented society" by 2020. To do so, the government set targets of raising R&D expenditure to 2.5% of GDP and reducing dependence on imported technology to below 30% (Cao et al., 2006). Central to the plan was the concept of *zizhu chuangxin* (indigenous innovation, 自主创新), reflecting a strategic pivot toward the development of frontier capabilities (Liu et al., 2011). Reinforcing this S&T agenda, the *National Medium- and Long-Term Talent Development Plan (2010–2020)*, jointly issued by the CCP Central Committee and the State Council in 2010, explicitly aimed to transform China's labor force from one driven by low-cost manufacturing into one anchored by high-skill, innovation-intensive sectors (Wang, 2011).

Talent recruitment was a central pillar of these plans. Two flagship programs illustrate the approach. The *Thousand Talents Program* (TTP, 千人计划, Qianren Jihua), launched in 2008 by the CCP Central Organization Department, targeted Chinese-born scholars holding full professorships or equivalent senior positions at leading overseas universities (Jia, 2018). Candidates were required to secure an institutional affiliation in China before applying, and only one domestic institution could sponsor a given candidate. According to the 2011 Detailed Rules for the Thousand Talents Program, the initiative offered competitive salaries, start-up research funding, and housing subsidies, aiming to recruit 2,000 overseas talents over five to ten years (Central Organization Department of the Communist Party of China [COD], 2011, trans. CSET). In 2011 the scheme was complemented by the *Young Thousand Talents Program* (YTTP, 青年千人计划, Qingnian Qianren Jihua), which extended the model to Chinese-origin researchers under age 40, typically at the postdoctoral or assistant-professor stage (Jia, 2018). By the end of their active period, the TTP and YTTP had collectively attracted several thousand overseas scientists to Chinese institutions. More than

7,000 individuals had been recruited under the TTP within a decade of its launch (Jia, 2018), and Shi et al. (2023) studied 721 awardees out of the 3,576 offers made under the YTTP between 2011 and 2017.

The emphasis on return migration reflects a broader view of overseas-trained scientists as vessels of transnational human and social capital rather than merely an additional labor supply. Foreign training, international publication experience, access to frontier technology, and overseas professional networks made returnees especially valuable as China sought to internationalize its science system (Zweig et al., 2004). Consistent with this view, bibliometric evidence similarly shows that returnees produce higher-impact work than comparable domestically trained scholars and remain more deeply embedded in international collaboration networks (Cao et al., 2020). However, the literature also cautions against viewing talent recruitment programs as mechanically generating brain gain. The attraction and integration of top scientists depended on research environments, academic governance, institutional autonomy, and organizational receptiveness, and many elite overseas scholars returned only part-time or chose not to return permanently (Cao, 2008; Zweig and Wang, 2013; Zweig et al., 2020). Consequently, institutions, not national policy alone, are the relevant unit for understanding how returnee inflows translate into local scientific production.

We therefore treat these talent attraction programs as an institutional backdrop for analyzing the return of overseas-trained scientists to China between 2003 and 2022. Although the policy initiatives were designed nationally, universities and research institutes implement them locally. Candidates required institutional sponsors, and the support available to returnees depended on the laboratories, research teams, and disciplinary environments into which they were embedded (COD, 2011; Jia, 2018). Prior evidence from both the Changjiang Scholar Program and evaluation of the YTTP shows substantial heterogeneity across universities in their ability and incentives to recruit returnees, reflecting differences in resources, prestige, host support, and the match between a candidate's expertise and local research priorities (Jia, 2018; Shi et al., 2023; Xin, 2006). Consequently, returnee inflows varied considerably across institutions and fields, despite a common national policy framework.

Ideally, one would observe each scientist's formal participation in the TTP or YTTP. In practice, however, these records are incomplete, inconsistently disclosed, and difficult to track because the programs were subsequently reorganized and renamed. Moreover, our objective is not to assess the effectiveness of these talent programs per se, but rather to evaluate the broader effects of brain circulation on China's scientific output. We therefore use 2008—the launch of the TTP—as the regime-change date, as it marks the beginning of a nationwide policy commitment to attracting overseas Chinese scholars back to China.

3 Data and Stylized Facts

We collect publication records from Scopus using selected journals' ISSNs and publication year (see Appendix Table A1). Using the Nature Index as the benchmark set of frontier scientific journals, we retain, within Chemistry, Physics, and Life Sciences, the five highest ranked journals by the 2022 Journal Impact Factor with complete Scopus coverage over the sample period and include three broad General journals. For each article, we retrieve the title, publication year, journal volume, author names, affiliations, citation counts (as of March 2025), and a brief description. We assign each publication, author, and institution a unique EID, author ID, and affiliation ID, respectively. Institution names and countries are linked through affiliation IDs. For authors listing multiple affiliations on a paper, we treat the first affiliation as the primary one. The final data set comprises 602,730 publications, 1,204,800 scholars, and 64,652 institutions.

Scholars' ethnicity is estimated by NamSor,[10] a European name-recognition software launched in February 2014. For a given full name, the NamSor Diaspora API estimates the most likely ethnicity, recognizing that a last name linked to an ethnicity may belong to a second- or third-generation migrant with a first name less common among that ethnicity's descendants.

Using unique author IDs, we assemble an author-year panel that tracks publication histories in the selected journals in our sample. Because authors may publish multiple papers in a given year and none in others, the panel is unbalanced. For the analysis, we classify authors' ethnicities and their affiliation countries as Asian and non-Asian groups. Each author-year observation is thus characterized by two dimensions: the author's ethnicity[11] and the country[12] of their affiliated institution at the time of publication. This structure allows us to examine collaboration patterns across ethnic and geographic groups and to document four stylized facts about China-affiliated research and ethnically Chinese scientists.

The four facts are organized around the first two questions posed in the introduction.

[10] https://namsor.app/. Independent validation exercises support the use of NamSor while also documenting classification error. Bursztyn et al. (2024) find that fewer than 1 percent of 250,000 voter records were misclassified in the error category relevant to their origin analysis. Comparing NamSor with alternative race and ethnicity inference methods against self-reported data, Rieke et al. (2022) find such inferences broadly informative but imperfect. In an affiliation-based benchmark of 88,699 researchers, Sebo (2023) report 85 percent correct ethnicity classification for researchers affiliated with Chinese institutions.

[11] Asian ethnicities include Bangladeshi, Bhutanese, Burmese, Cambodian, Chinese, Indian, Indonesian, Japanese, Korean, Laotian, Malays, Mongolian, Nepalese, Pakistani, Sri Lankan, Thai, and Vietnamese. The top six most common Asian ethnicities among unique scholars are: Chinese (21.45%), Japanese (5.60%), Indian (4.30%), Korean (3.43%), Vietnamese (0.52%), Bangladeshi (0.39%)

[12] Asian countries include most East, Southeast, and South Asian countries: Bangladesh, Bhutan, Brunei Darussalam, Cambodia, China, India, Indonesia, Japan, Laos, Malaysia, Maldives, Mongolia, Myanmar, Nepal, Pakistan, Philippines, Singapore, South Korea, Sri Lanka, Thailand, and Viet Nam.

Facts 1 and 2 take up the first: China's rise at the frontier is real, sustained, and driven by China-based scholars. Facts 3 and 4 take up the second: return migration grew and increasingly drew high-quality scholars, and returnees' post-return success is concentrated in the field where China is already strongest.

Throughout the paper, we distinguish ethnicity from research location. An *ethnically Chinese author* is an author classified as Chinese by NamSor; this name-based classification is probabilistic and does not measure citizenship or nationality. A *China-affiliated author* is an author whose assigned affiliation country in a given author-year is China: we first identify the primary affiliation on each paper and then use the modal affiliation country when an author publishes multiple papers in the same year. A *China-affiliated paper* has at least one China-affiliated author. When both the ethnicity and affiliation conditions hold in an author-year, we refer to a *China-affiliated ethnically Chinese author*.

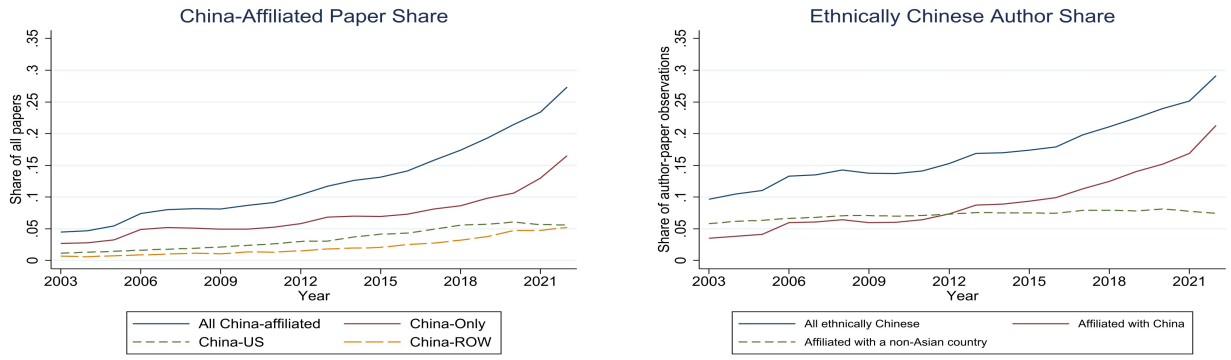
3.1 Fact 1: Rising presence of Chinese scholars

Between 2003 and 2022, Chinese science moved from the margins of elite journals toward their center: the share of papers with at least one China-affiliated author rose more than fivefold, and by the end of the sample nearly half of Chemistry papers had one (Figure 1). We measure this presence on two margins. We define the *paper share* as the fraction of papers in a given year with at least one China-affiliated author, and the *author share* as the fraction of author-paper observations accounted for by ethnically Chinese authors.

First, the China-affiliated paper share rose drastically, from 4.47% of all papers published in 2003 to 27.34% in 2022. We categorize China-affiliated papers into three mutually exclusive groups: "China-Only" (all authors are China-affiliated), "China-US" (at least one author is China-affiliated and at least one is U.S.-affiliated), and "China-ROW" (at least one author is China-affiliated and at least one is affiliated with a non-China, non-U.S. institution, but none is U.S.-affiliated). Figure 1a shows that the China-Only paper share rose from 2.66% in 2003 to 16.51% in 2022. The shares of China-US and China-ROW papers also increased, though more slowly. The growth of these internationally coauthored papers indicates the increasing integration of China-affiliated researchers into the global scientific community.

Second, Figure 1b shows that the share of all author-paper observations accounted for by China-affiliated ethnically Chinese authors rose from 3.50% to 21.29%, pointing to a substantial strengthening of domestic research capacity. By contrast, the share of author-paper observations accounted for by ethnically Chinese authors affiliated with non-Asian institutions rose more modestly, from 5.79% to 7.42%.

Moreover, these trends exhibit substantial cross-field heterogeneity. Figure 2 shows that



(A) China-Affiliated Paper Share (B) Chinese Author Share by Affiliation

Note: Figure 1a shows the annual share of papers with at least one China-affiliated author and decomposes it into three mutually exclusive groups: China-Only, China-US, and China-ROW. Figure 1b shows the annual share of author-paper observations accounted for by ethnically Chinese authors, together with two selected affiliation subsets: China-affiliated ethnically Chinese authors and ethnically Chinese authors affiliated with institutions in non-Asian countries.

FIGURE 1
China-Affiliated Papers and Ethnically Chinese Authors

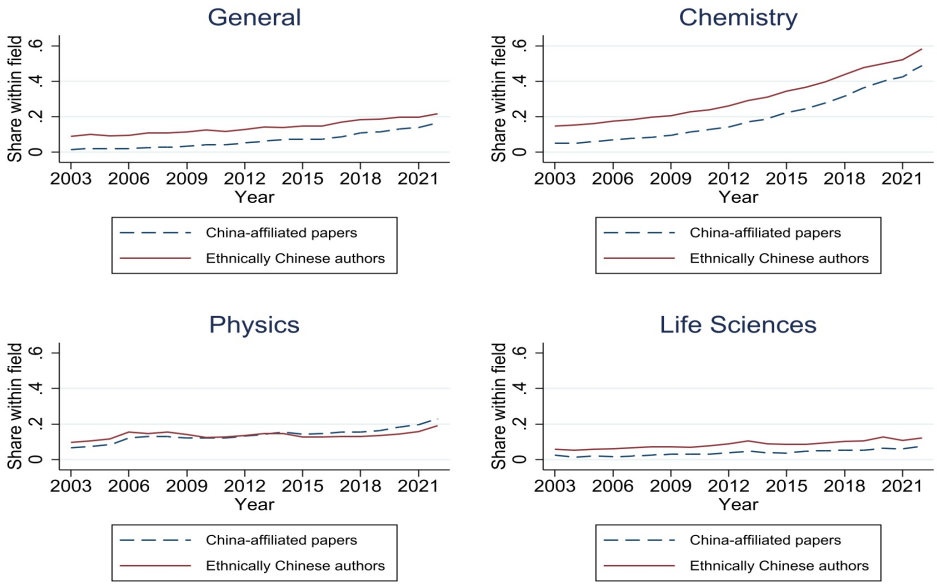
the increase is concentrated in Chemistry, where the China-affiliated paper share reaches 48.91% and the ethnically Chinese author share reaches 58.31% by the end of the sample period. The growth in General and Physics is more moderate, while that in Life Sciences remains the slowest, with the China-affiliated paper share increasing from 2.28% in 2003 to 7.46% in 2022.

3.2 Fact 2: Increasing contribution of Chinese scholars

Presence, however, is not influence: does the publication surge translate into citation-weighted contribution? It does, and the growth is home-grown: the citation share of China-affiliated ethnically Chinese authors rose roughly tenfold over the sample period, while the share of their U.S.-affiliated counterparts stayed broadly flat (Figure 3). Following standard practice in the literature, we use citation counts as a proxy for paper quality and scientific influence in our dataset. To quantify contributions by author-ethnicity, we construct a Citation Share measure. Assuming equal contributions across coauthors, we compute the Citation Share for ethnicity e in each publication year as follows:

$$CS_t^{e(k)} = \frac{\sum_{i=1}^{N_t} c_{it} * s_{it}^{e(k)}}{\sum_{i=1}^{N_t} c_{it}} = \frac{\sum_{i=1}^{N_t} c_{it} * \frac{a_{it}^{e(k)}}{a_{it}}}{\sum_{i=1}^{N_t} c_{it}} \tag{1}$$

where c_{it} is the total citation count of paper i published in year t, measured as of March 2025. Because Scopus reports cumulative citations as of the data retrieval date rather than annual



Note: In each field, "China-affiliated papers" denotes the annual share of papers with at least one China-affiliated author, and "Ethnically Chinese authors" denotes the annual share of author-paper observations accounted for by ethnically Chinese authors.

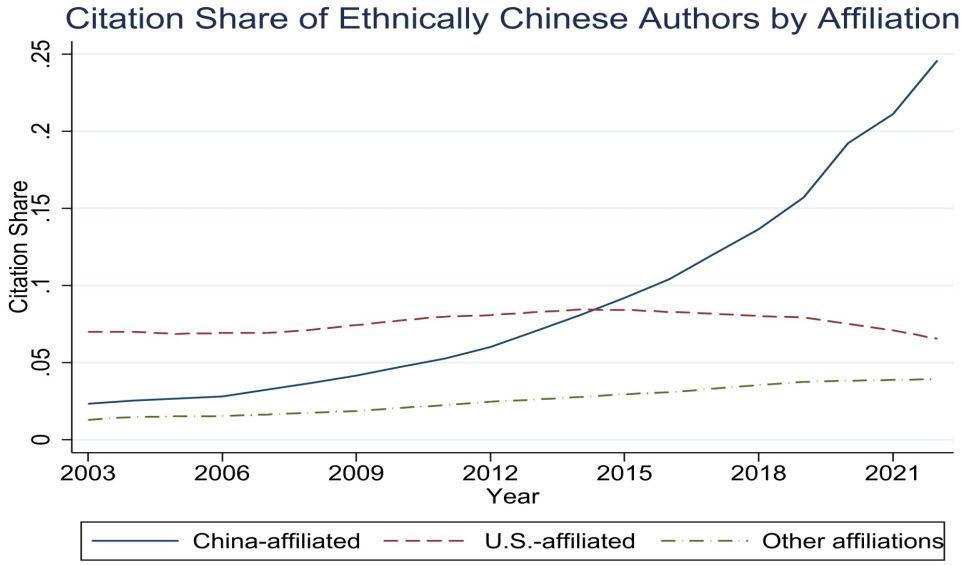
FIGURE 2
China-Affiliated Paper and Ethnically Chinese Author Shares by Field

flows, c_{it} is a fixed value for each paper i with publication year t; it is not a time-varying measure. s_{it}^{e} is the share of authors with ethnicity e in paper i: a_{it}^{e} is the number of authors with ethnicity e on paper i published in year t, a_{it} is the paper's total number of authors, and N_t is the total number of papers published in year t. Eq. (1) measures the share of total citations in publication year t attributable to authors of ethnicity e. Because citations proxy for scientific influence and paper quality, the measure captures ethnicity e's share of high-impact contributions to scientific discovery. We can apply the same approach at the institutional level by replacing ethnicity e with institution k, and measuring the Citation Share of scholars affiliated with institutions across different countries.

We examine the contributions of ethnically Chinese scientists by grouping them according to their institutional affiliation at the time of publication. Specifically, we compare the Citation Share of ethnically Chinese authors affiliated with institutions in China, the United States, and all other countries, which we collectively label "Rest of the World (ROW)". Figure 3 plots the year-by-year evolution of these Citation Share series, calculated using Eq. (1).

The blue line in Figure 3 shows that the Citation Share of China-affiliated ethnically Chinese authors rose sharply from 0.023 in 2003 to 0.246 in 2022. The Citation Share

of ethnically Chinese authors affiliated with institutions outside China and the U.S. also increased, although more modestly. In contrast, while U.S.-affiliated ethnically Chinese authors accounted for the largest Citation Share at the beginning of the sample period, their share remained broadly stable and declined slightly in recent years. Overall, the combined Citation Share of ethnically Chinese authors increased by 230.23% over the sample period, underscoring their increasingly important role in advancing frontier scientific research.



Note: This figure compares the Citation Share of ethnically Chinese authors affiliated with institutions in China, the U.S., and all other countries ("ROW"), calculated using Eq. (1).

FIGURE 3
Citation Share of Ethnically Chinese Authors by Affiliation Country

3.3 Fact 3: High-quality Chinese scholars are more likely to return

Behind this home-grown rise stands a growing—and increasingly selected—flow of returnees. To show this, we first construct an individual author-quality measure and then examine return migration among ethnically Chinese authors.

To compare research performance across scholars, we construct an Author Quality measure based on each scholar's cumulative citation share across all papers included in our sample. Assuming equal contributions among coauthors, the Author Quality of author j in year t is defined as follows:

$$AQ_{jt} = \sum_{s=2003}^{t} \frac{\sum_{i=1}^{n_{js}} \frac{c_{is}}{a_{is}}}{\sum_{i=1}^{N_s} c_{is}} \tag{2}$$

where c_{is} is the total lifetime citation count of paper i published in year s, measured as of

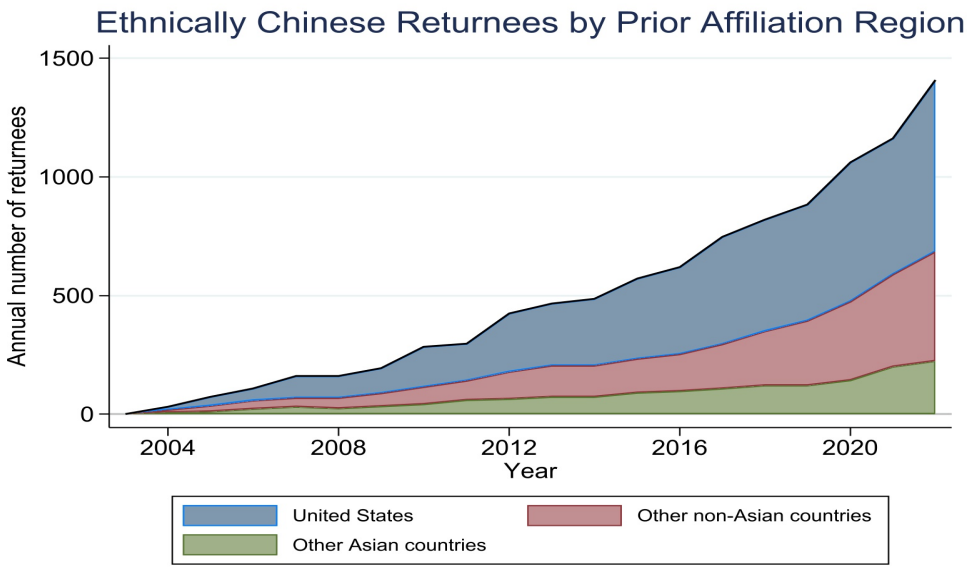
March 2025, the end of our time window; a_{is} is the number of coauthors of paper i in year s; n_{js} is the number of papers published by author j in year s; and N_s is the total number of papers published in year s. To facilitate comparisons across authors within a publication-year cohort, author j's citation share is normalized by total citations in the same year s. Accordingly, author j's quality index in year t is the cumulative citation share of all papers she published up to year t. For each year t, we compute the median Author Quality, med_t, among all active authors and define $HighQuality_{jt} = \mathbb{1}\{AQ_{jt} > med_t\}$. Thus, $HighQuality_{jt}$ identifies authors in the upper half of the active-author quality distribution in year t. We also construct field-specific medians to measure an author's quality relative to peers within the same field.

We infer authors' cross-country mobility from changes in the countries of their publication affiliations.[13] This affiliation-based approach follows recent large-scale work using bibliometric records to measure international scholarly mobility (Sanliturk et al., 2023). We define a *returnee* as an ethnically Chinese author whose assigned affiliation changes from an institution outside China to one in China within our observed publication panel. This definition records observed publication-affiliation mobility rather than a scholar's complete migration or employment history. Figure 4 plots annual inflows of returnees by origin country. Two features stand out. First, inflows rise in every year of the sample, from 31 returnees in 2004 to 1,409 in 2022. Second, most returnees were previously affiliated with U.S. institutions. Appendix Table A2 reports the corresponding annual counts and cumulative totals; by 2022, we identify 9,963 ethnically Chinese scholars with an observed move from an overseas institution to China, including 5,504 whose immediately preceding affiliation was in the U.S.

Because the U.S. has long been the world's leading destination for scientific education and research, all subsequent analyses of returnee selection, performance, collaboration, and institutional exposure focus on the 5,504 scholars whose assigned affiliation changes from the U.S. to China. We refer to these scholars as *U.S.-to-China returnees* and, unless otherwise stated, use "returnees" to mean this group throughout the remainder of the paper. We define $Return_{jt}$ as an indicator equal to one in the year scholar j becomes a U.S.-to-China returnee, zero during the scholar's observed U.S.-affiliated pre-return years, and missing after the return; this designation facilitates our subsequent estimation of the probability of becoming a returnee.[14]

[13] We treat the first affiliation as primary for authors with multiple affiliations in one paper. If an author has different primary affiliations for papers published in the same year, we keep the modal affiliation country within a year as the main location and assign by alphabetical order if there are ties. We focus on the last movement for authors with multiple cross-country movements in their publishing history.

[14] The regression sample consists of all author-year observations during scholar j's observed U.S. tenure. For scholars who first appear outside the U.S. and later move to the U.S., observations before the first observed U.S. affiliation are excluded; for U.S.-to-China returnees, post-return years are right-censored. For



Note: This figure shows annual inflows of ethnically Chinese returnees by the region of their assigned affiliation immediately before moving to China. The three mutually exclusive origin groups are the U.S., other non-Asian countries, and other Asian countries. Appendix Table A2 reports the corresponding annual counts and cumulative totals.

FIGURE 4
Annual Inflows of Ethnically Chinese Returnees by Prior Affiliation Region

We investigate whether the propensity of ethnically Chinese scientists to return from the U.S. to China varies with author quality. Since migration decisions are endogenous, we do not seek to identify a causal effect. Instead, we estimate the following linear probability model to describe the quality ranking of U.S.-to-China returnees:

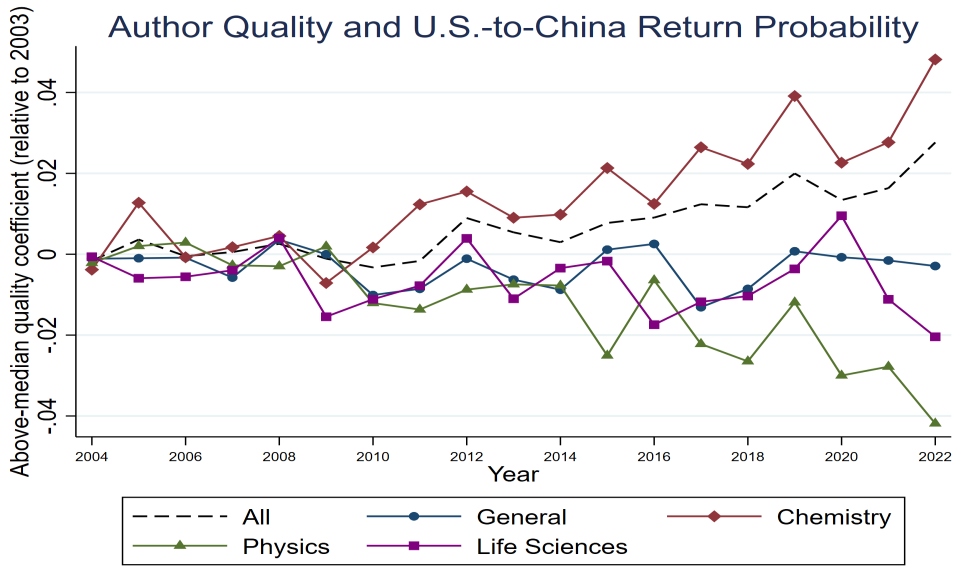
$$\begin{aligned} Return_{jt} = \alpha + \sum_{s=2004}^{2022} \beta_s HighQuality_{jt} \mathbb{1}\{t = s\} \\ + \lambda HighQuality_{jt} + \sum_{s=2004}^{2022} \gamma_s \mathbb{1}\{t = s\} + \eta_1 CareerAge_{jt} + \epsilon_{jt} \end{aligned} \tag{3}$$

For this migration specification, we compute the median defining $HighQuality_{jt}$ after excluding all author-year observations following an observed U.S.-to-China return, while retaining the return year itself. Our parameter of interest is β_s, the coefficient on the interaction between above-median quality and the year indicator. In this linear probability model, β_s is the percentage-point difference in the probability of returning from the U.S. to China between above-median- and below-median-quality ethnically Chinese authors in year s, relative to the 2003 difference. To proxy for seniority, we include $CareerAge_{jt}$, defined as year

scholars who have always been affiliated with U.S. institutions, the variable takes 0 for all years in the sample.

t minus the year of author j's first observed publication.

We estimate Equation (3) using simple Ordinary Least Squares (OLS) with clustered standard errors at the individual author level. Figure 5 plots the estimated β_s coefficients over time as a black dashed line; the corresponding all-field regression table appears in Appendix Table A3. Overall, the estimated β coefficients trend upward, implying that higher-quality U.S.-affiliated ethnically Chinese authors become increasingly likely to return to China relative to their lower-quality counterparts, especially toward the end of the sample period. For example, in 2022, the return probability gap between higher- and lower-quality U.S.-affiliated ethnically Chinese authors was approximately 2.8 percentage points larger than in 2003.[15]



Note: This figure plots the estimated coefficients β_s in Equation (3) for ethnically Chinese authors in the observed U.S.-affiliated risk set. Each coefficient is the above- versus below-median difference in the probability of becoming a U.S.-to-China returnee in year s, relative to the corresponding difference in 2003. The yearly median used to define above-median quality excludes post-return author-year observations.

FIGURE 5
Author Quality and U.S.-to-China Return Probability

We further examine field heterogeneity by separately estimating Equation (3) using subsamples of different fields $f \in \{\text{General}, \text{Chemistry}, \text{Physics}, \text{Life Sciences}\}$, also using OLS

[15]One might worry that the rising $\hat{\beta}_s$ reflects a growing pool of high-quality ethnically Chinese scientists at U.S. institutions rather than genuine quality selection in return migration. Since $HighQuality_{jt}$ is defined relative to the within-year median and year fixed effects absorb aggregate time trends in return rates, this concern is largely addressed by the regression design. Consistent with this, Appendix Figure A1 shows that the count of high-quality ethnically Chinese scholars affiliated with U.S. institutions remained relatively stable over the sample period.

with clustered standard errors at the individual author level. Figure 5 also plots the estimated β coefficients for each field. Among U.S.-affiliated ethnically Chinese authors, high-quality chemists are significantly more likely to return to China from 2010 onward. In 2022, the return probability gap between higher- and lower-quality chemists was about 4.7 percentage points larger than in 2003. By contrast, the return probability gap for authors publishing in General journals shows no significant change. Moreover, the corresponding gaps for physicists and life scientists decline significantly over time. In 2022, the quality gaps for physicists and life scientists were about 4.2 and 2 percentage points smaller than in 2003, respectively.

3.4 Fact 4: Returnee performance varies across fields

In this subsection, we compare the post-return performance of returnees with that of similar U.S.-based scientists. Because some authors publish in multiple fields, the analysis is conducted at the author–field level: an author publishing in both Chemistry and Physics contributes a separate history in each field.

We construct a complete author–field–year panel, assigning zero annual contribution in years without a publication in our selected journals. Let AQ_{jft} denote author j's cumulative Author Quality in field f through year t, constructed as the field-specific counterpart of Equation (2). For a returnee who moves to China in year t_j, we measure post-return performance over two common windows: the return year through $t_j + 3$ and the return year through $t_j + 5$. We denote the returnee's Author Quality accumulated over horizon h by $AQ^h_{jf} = AQ_{jf,t_j+h} - AQ_{jf,t_j-1}$, where $h \in \{3, 5\}$. The common windows ensure that returnees and their comparisons have the same amount of post-return time in which to publish.

We identify two matched U.S. comparisons for each returnee–field pair: one ethnically Chinese scientist and one non-Chinese scientist. Matching is with replacement and exact within field–return-year cells. Within each field–year cell, we use one-nearest-neighbor Mahalanobis matching based on cumulative pre-return quality, quality in the year before return, quality and publication counts over the preceding three years, and career age. Comparison scholars must be U.S.-affiliated before the return year and remain observed at U.S. institutions through the relevant three- or five-year window; they may move afterward.

For a given horizon h and comparison group $g \in \{\text{EC}, \text{NC}\}$, let j' denote the U.S.-based scientist matched to returnee j in field f, where EC and NC denote ethnically Chinese and non-Chinese U.S. scientists, respectively. We calculate the matched scientist's Author Quality over the same window as $AQ^h_{j'f} = AQ_{j'f,t_j+h} - AQ_{j'f,t_j-1}$ and define the matched difference as $\Delta^{h,g}_{jf} = 10{,}000\left(AQ^h_{jf} - AQ^h_{j'f}\right)$. We estimate an intercept-only OLS regression, $\Delta^{h,g}_{jf} = \alpha^{h,g}_f + \varepsilon^{h,g}_{jf}$, separately for each field f, horizon h, and comparison group g. Thus,

$\widehat{\alpha}_f^{h,g}$ is simply the mean matched returnee–stayer difference in that cell. Standard errors are two-way clustered by returnee j and matched-scientist identity to account for returnees appearing in multiple fields and for matching with replacement. These estimates are matched descriptive comparisons of post-return performance.

Figure 6 shows a sharp and stable field pattern. Over the three-year window, returning chemists outperform ethnically Chinese U.S. stayers by 0.257 and non-Chinese U.S. stayers by 0.331. The corresponding five-year gaps widen to 0.627 and 0.693. By contrast, returning life scientists underperform both groups: the three-year gaps are -0.234 and -0.383, and the five-year gaps are -0.260 and -0.507. General-journal returnees also underperform both comparison groups over both horizons. Returning physicists occupy the middle ground. Their performance is statistically indistinguishable from that of ethnically Chinese U.S. stayers over both horizons. Relative to non-Chinese U.S. stayers, the gap is -0.103 and statistically significant over three years, but is similar in magnitude and imprecisely estimated over five years. Appendix Table A4 reports all estimates.

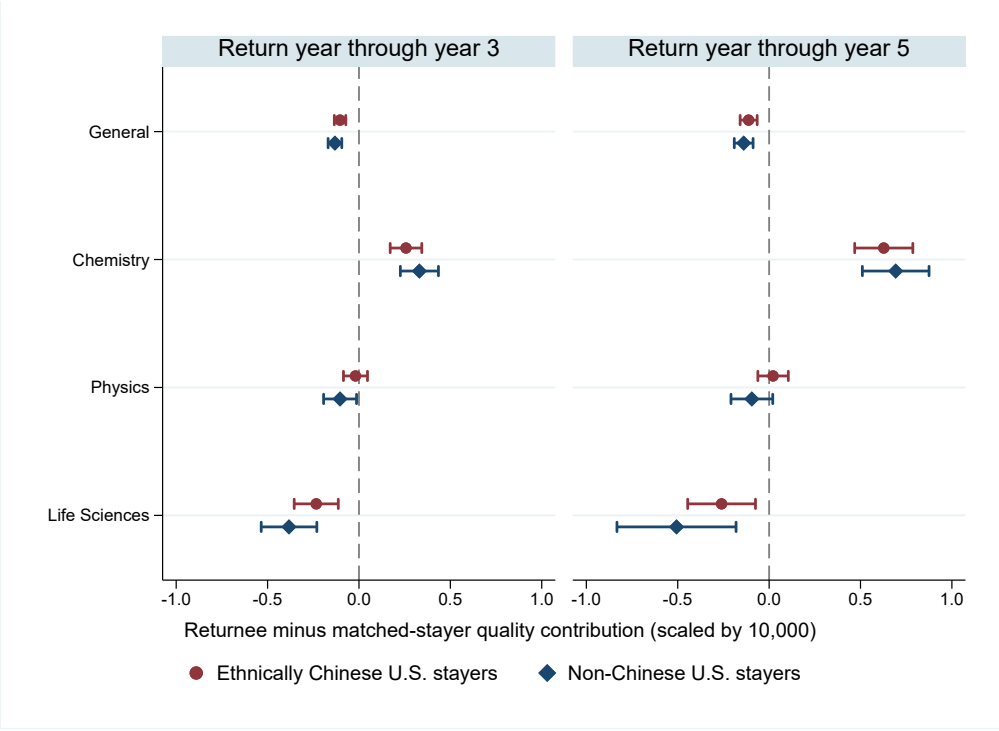
One possible explanation for the weaker performance of General-journal returnees is that some had been embedded in research teams with prominent U.S.-based scientists before returning, while operating more independently after returning may make it harder to maintain publication records in these most competitive outlets. The broader cross-field correspondence is notable: Chemistry, the field in which return migration increasingly selected high-quality scholars, is also the field in which returnees display the strongest post-return performance, while Physics and Life Sciences show no comparable advantage (Figure 5).

Taken together, the four facts show that China's growing contribution to frontier science is increasingly driven by China-based researchers and accompanied by rising return migration. Returnees have become more positively selected over time, but their post-return performance varies substantially across fields. These patterns motivate the analysis in Section 4, where we examine how returnees contribute to local scientists' research performance through direct collaboration and their broader presence within research institutions.

4 Brain Circulation of Returnees

Having documented who returns and how returnees perform, we next ask how returnees contribute to China's research system. We move from their closest research links to broader outcomes among local scholars and institutions.

We define a *local China-based scholar* as an ethnically Chinese author whose assigned publication affiliation remains in China throughout every author-year observed in our sample. "Local" therefore refers to the observed publication panel, not necessarily the scholar's com-



Note: The figure plots $\widehat{\alpha}_f^{h,g}$ from the intercept-only regressions described in the text. Each point is the mean difference between a returnee's and a matched U.S.-based scientist's cumulative normalized citation contribution over the indicated post-return horizon, multiplied by 10,000. Positive values indicate higher performance by the returnee. Matching is with replacement and exact within field and return year; Mahalanobis distance is calculated using cumulative pre-return quality, quality in the previous year, quality and publication counts over the previous three years, and career age. Bars show 95% confidence intervals based on standard errors two-way clustered by returnee and matched-scientist identities.

FIGURE 6
Field Heterogeneity in the Post-Return Performance of Returnees

plete career or residence history.

Table 1 summarizes collaboration histories among local scholars with a sustained presence in the publication panel, i.e. if she publishes in at least two distinct years. A single appearance provides little information about how her collaboration network develops and supplies no within-author variation for the fixed-effect designs used below, so we omit such authors from this descriptive tabulation. We classify each scholar once at the end of the sample period (2022), using all collaborations observed during 2003–2022. The four groups indicate whether the scholar ever coauthored with returnees, non-local international scholars, both, or neither during this window; the two types of collaboration need not occur in the same year. In this broad composition of the local-scholar population, we see that connections beyond the purely local network are common. Of the 46,358 repeatedly observed local scholars, 39.3% had collaborated with both types by the end of the sample period.

The remaining local scholars are divided fairly evenly among those with no such connection, those connected only to returnees, and those connected only to international scholars.[16]

TABLE 1
Collaboration Histories of Repeatedly Observed Local Scholars

	Unique count	Share
No returnee or international coauthor	9,893	21.3%
Returnee coauthor(s) only	8,270	17.8%
International coauthor(s) only	9,972	21.5%
Both types of coauthors	18,223	39.3%
All repeatedly observed local scholars	46,358	100.0%
Returnees	5,504	—

Notes: The table counts each scholar once over 2003–2022 and restricts the descriptive population to scholars observed publishing in at least two distinct years. Local scholars are ethnically Chinese authors whose assigned publication affiliation remains in China in every observed author-year. The four local-scholar categories are mutually exclusive and assigned once at the end of the sample period (2022), using all collaborations observed during 2003–2022. A scholar belongs to the "both" category if at least one collaboration of each type is observed during this window, even if the two types occur in different years. The "only" categories require that no collaboration of the other type is observed anywhere in the window. Classification does not require the scholar to publish in 2022. A returnee coauthor is a U.S.-to-China returnee observed in or after the return year; a non-local international coauthor is assigned an affiliation outside China in the year of the joint paper. The first row includes 19 scholars with no observed coauthors. Shares are calculated among repeatedly observed local scholars; returnees are reported separately.

The analysis proceeds in four parts. First, we estimate how local scholars' research quality changes after they begin coauthoring with returnees, distinguishing the gains associated with returnee collaboration from those associated with international collaboration. Second, among scholars with no prior returnee or non-local international collaboration, we use a capacity-aligned shift-share instrument to study institution–field returnee exposure while addressing the endogenous sorting of returnees across institutions. Third, we apply the same exposure design to scholars with prior returnee or international collaboration, conditioning on their collaboration histories and annual collaboration intensity. Note that the full-period collaboration categories in Table 1 do not map one-for-one onto the regression sample of

[16] Appendix Figure A2 plots annual shares among publishing scholars in the same population, using collaboration histories observed through each year rather than the table's fixed full-period categories. Scholars can therefore change groups. Pre-2003 histories are unavailable, and identifying a returnee requires an observed U.S.-to-China affiliation transition; the returnee-related shares consequently start at zero in 2003 by construction.

each analysis above, as they apply different restrictions based on timing, field, institution, covariate availability. Finally, we aggregate the data to the institution–year level and examine research quality, workforce composition, and scientific output.

4.1 Direct Collaboration with Returnees

We begin with the direct collaboration channel: whether local scholars' research quality changes after they first coauthor with a returnee, and whether that pattern depends on their prior international connectedness. The outcome is log Author Quality, $\log AQ_{jt}$, where AQ_{jt} is the cumulative citation-share-based measure defined in Eq. (2). As a broad benchmark, we first estimate an event study around first returnee collaboration using the full sample of local scholars. Appendix Equation (10) presents this specification, and Appendix Figure A3 reports positive post-collaboration estimates. Because first returnee collaboration can coincide with entry into broader international coauthorship networks, the benchmark does not by itself isolate a returnee-specific increment. Our main analysis therefore compares quality changes following the same event—first returnee collaboration—in two groups defined by whether the scholar already had a non-local international coauthor.

Risk sets and specification. Let T_j^{ret} and T_j^{int} denote the year of author j's first coauthorship with a returnee and with any non-local international author, respectively. The estimation sample excludes returnees. For treated authors, event time is $\tau_j = t - T_j^{ret}$. To estimate the returnee increment, we define the risk set

$$\mathcal{S}^{ret} = \{j : T_j^{int} < T_j^{ret}\} \cup \{j : T_j^{int} < \infty,\ T_j^{ret} = \infty\},$$

where the second set contains internationally connected authors who never coauthor with a returnee and therefore serve as controls. Within this author set, observations begin in the year of first international collaboration, so every treated and comparison observation comes from an internationally connected scholar. On this risk set, we estimate:

$$\begin{aligned}\log AQ_{jt} = &\sum_{\tau \neq -1} \beta_\tau \mathbf{1}\{\tau_j = \tau\} + \gamma \cdot \mathrm{IHS}(C_{j,t-1}^{intl}) + \delta \cdot \mathrm{IHS}(C_{j,t-1}^{all}) \\ &+ \xi_1(\overline{PreIntl}_j \times t) + \xi_2(\overline{PreIntl}_k \times t) + \mu_j + \lambda_t + \varepsilon_{jt} \end{aligned} \tag{4}$$

where $C_{j,t-1}^{intl}$ is the lagged cumulative count of distinct non-local international coauthors, absorbing ongoing foreign exposure; $C_{j,t-1}^{all}$ is the lagged cumulative count of all coauthors, controlling for overall collaboration scale. $\mathrm{IHS}(\cdot)$ is the inverse hyperbolic sine transforma-

tion,[17] which accommodates zeros. $\overline{PreIntl}_j$ and $\overline{PreIntl}_k$ denote author- and institution-level pre-policy (2003–2005) international-collaboration rates.[18] Both are interacted with the raw publication-year variable $t \in \{2003, \ldots, 2022\}$ to absorb differential linear trends by authors' and institutions' pre-policy international-collaboration rates. μ_j and λ_t are author and year fixed effects. We estimate the event study on the author–year panel using overall Author Quality. This avoids assigning authors who publish in multiple fields to a single field. The estimates therefore capture the average post-collaboration trajectory across fields rather than field-specific effects.

For scholars without prior international collaboration, we define the author–year risk set as

$$\mathcal{S}^{no\,intl} = \{(j,t) : t < T_j^{int}\}.$$

Treatment occurs at T_j^{ret} when $T_j^{ret} < T_j^{int}$; other scholars serve as controls until their first international collaboration. Simultaneous first events therefore contribute only earlier observations. We estimate Equation (4), omitting the international-coauthor stock, which is zero in this risk set. We also omit the author-level pre-policy international trend because its measurement window can extend beyond first returnee collaboration, making it potentially affected by the event being studied.

Our preferred implementation estimates separate event-time effects for each cohort defined by the calendar year of first returnee coauthorship and then aggregates the cohort-specific estimates using the interaction-weighted estimator of Sun and Abraham (2021). This avoids using already-treated authors as controls when effects differ across first-collaboration cohorts. Standard errors are clustered at the author level. Authors with no returnee treatment in the relevant risk set form the comparison cohort. For both risk sets, we report conventional two-way fixed-effects (TWFE) estimates as a benchmark, with standard errors clustered at author and institution level. Because the lagged collaborator stocks can themselves change following first returnee collaboration, we interpret these estimates as quality associations conditional on network evolution.

Results. Figure 7 reports the two returnee-collaboration event studies. We display exact event times from $\tau = -6$ through $\tau = 3$, omit $\tau = -1$, and group observations at $\tau \leq -7$ and $\tau \geq 4$ into endpoint categories. Panel A reports the returnee-collaboration increment among scholars with prior international collaboration, using 42,886 author-year observations from

[17] The inverse hyperbolic sine transformation is defined as $\text{IHS}(x) = \sinh^{-1}(x) = \ln\left(x + \sqrt{x^2+1}\right)$

[18] Let $InternationalCollab_{j\tau}$ equal one if author j coauthored with at least one non-local international scholar in year τ. Then $\overline{PreIntl}_j \equiv \frac{1}{3}\sum_{\tau=2003}^{2005} InternationalCollab_{j\tau}$. The institution-level measure $\overline{PreIntl}_k$ is the average of $InternationalCollab_{j\tau}$ across author-year observations at institution k during 2003–2005; k indexes author j's institution in the relevant publication year.

10,920 authors. The five displayed leads are jointly insignificant under both estimators. These tests provide no evidence of differential pre-collaboration trends over the displayed lead window, although they do not establish that collaboration timing is exogenous.

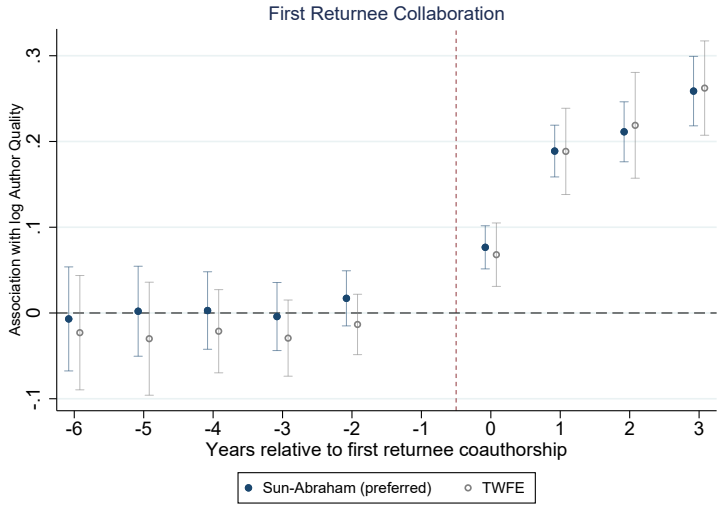
From the collaboration year onward, the preferred coefficients in Panel A are positive and precisely estimated: 0.077 at $\tau = 0$, 0.189 at $\tau = 1$, 0.211 at $\tau = 2$, and 0.259 at $\tau = 3$. These log-point estimates correspond to cumulative Author Quality approximately 8.0% higher in the collaboration year and 29.6% higher three years later, relative to the pre-collaboration year and net of contemporaneous changes among internationally connected authors who never collaborate with a returnee, conditional on the controls. The TWFE benchmark is close in both level and shape. The rising path indicates a widening gap in accumulated citation-share-based quality; it does not directly measure growth in annual research productivity.

Panel B examines first returnee collaboration among scholars who have not yet collaborated internationally, using 49,550 author-year observations from 18,406 authors and censoring observations at first international collaboration. The preferred estimates at event times 0 through 3 are 0.141, 0.060, 0.144, and 0.182, with a similar pattern under TWFE. The coefficients remain positive but, unlike Panel A, fall in the first post-collaboration year before rising again. The five displayed leads are jointly insignificant under both estimators. Thus, a positive post-collaboration association appears in both groups, but the different risk sets and controls prevent interpreting the difference between their coefficient paths as established heterogeneity.

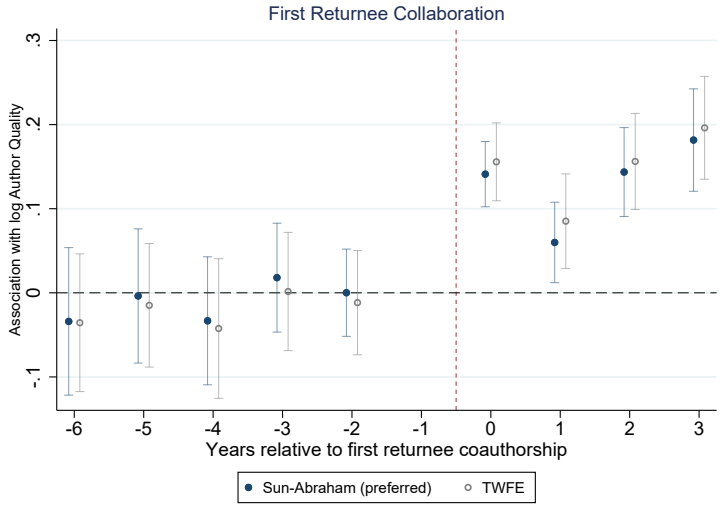
We first assess sensitivity to how collaboration exposure is controlled for in Panel A. The two lagged collaboration measures in Eq. (4) overlap because non-local international coauthors are included in the cumulative count of all coauthors. Appendix Table A6 replaces the all-coauthor count with a non-overlapping China-only collaborator stock through year $t-1$, while retaining the international-coauthor count. The preferred contemporaneous estimate is unchanged and the subsequent coefficients remain positive and precise, although they are approximately 25–30% smaller. Thus, the positive association persists when domestic and international collaborator stocks enter separately, while its longer-run magnitude is sensitive to this alternative parameterization of network exposure.

We next ask whether the quality gains extend beyond papers written jointly with returnees. We reconstruct cumulative Author Quality by assigning zero citation-share contribution to papers containing an eligible returnee in or after that returnee's reported return year, while leaving the publication-year citation denominator unchanged. In Appendix Figure A4, the contemporaneous coefficients become negative under both Sun–Abraham and TWFE in both groups. Later estimates turn positive among internationally connected schol-

Panel A: Scholars with prior international collaboration



Panel B: Scholars without prior international collaboration



Notes: Both panels estimate quality changes following first returnee collaboration and report Sun–Abraham (preferred) and TWFE estimates. Panel A conditions on prior non-local international collaboration and uses 42,886 author-year observations from 10,920 authors. Panel B retains scholars only while they have no international collaboration and uses 49,550 observations from 18,406 authors. Both specifications include author and year fixed effects and the lagged total-coauthor stock. Panel A also includes the lagged international-coauthor stock and author- and institution-level pre-policy international-collaboration-rate trends; Panel B retains the institution-level trend. Sun–Abraham standard errors are clustered by author; TWFE standard errors are two-way clustered by author and institution. Exact event times from −6 through 3 are displayed, with −1 normalized to zero; bars show 95% confidence intervals.

FIGURE 7
First Returnee Collaboration by Prior International Connectedness

ars, reaching 0.081 under Sun–Abraham at event time 3, but remain negative among scholars without prior international collaboration. These negative estimates do not establish that returnee collaboration harms other work: the event is dated by the first paper removed from the outcome, and logging the remaining quality excludes author–years with zero non-joint quality, particularly in Panel B. Appendix Figures A5 and A6 assess sensitivity to network adjustment and comparison construction for both overall quality and quality excluding joint papers. These diagnostics show that the later non-joint trajectory is specification-dependent; the growth-adjusted pairwise estimates do not precisely establish positive non-joint gains among internationally connected scholars at event times 1 through 3. The evidence therefore points most clearly to gains embodied in joint research. This pattern is consistent with returnees contributing expertise and research inputs through shared projects. The later positive estimates among internationally connected scholars could reflect greater capacity to apply those inputs elsewhere.

Finally, we reverse the sequence of network entry and examine first international collaboration among scholars already connected to returnees. Appendix Figure A7 reports positive and increasing preferred coefficients of 0.160, 0.294, 0.356, and 0.413 at event times 0 through 3. The path is descriptively steeper than the returnee increment in Panel A, but the different risk sets preclude ranking the effects of the two collaboration types. The positive increment suggests that direct international ties may provide access to expertise, collaborators, and research resources beyond those available through returnee partners.

4.2 Institutional Exposure among Locals Without Prior Collaboration

In addition to direct collaboration, returnees may affect local scholars' quality through indirect institutional channels, including seminar participation, shared research resources, mentoring, knowledge diffusion through colleagues, broader improvements in the institution's research environment, and competition over grants, lab space, and graduate students. In this section, we examine the net relationship between institutional returnee exposure and outcomes among local China-based scholars.

Section 3.4 shows that returnees' post-return performance is not uniform across scientific fields. Relative performance is strongest in Chemistry, negative in Life Sciences and General journals, and close to zero in Physics when ethnically Chinese U.S. stayers provide the benchmark. This heterogeneity motivates measuring both returnee exposure and local-author quality within fields. The IV strategy below predicts absolute returnee inflows into an institution–field cell by interacting its pre-policy share of China's publication capacity in

that field with the subsequent national returnee inflow.

Design. Let B_{kf}^{pre} be the number of distinct sampled papers involving institution k in field f during 2003–2005. Duplicate institution–paper–field records are removed before counting, so a paper is counted only once for an institution and field even when it contains multiple authors from that institution. Define institution k's share of China's pre-policy publication capacity in field f as $a_{kf} = \frac{B_{kf}^{pre}}{B_f^{pre}}$, where $B_f^{pre} = \sum_k B_{kf}^{pre}$. We call the resulting instrument *capacity-aligned* because it allocates the common national returnee wave to institution–field cells in proportion to their pre-policy publication capacity, aligning predicted absolute inflows with the observed institution–field returnee stock. The cumulative leave-own-institution-out national field shock from incoming returnees is $G_{f,t-1}^{-k} = \sum_{s=2009}^{t-1} \left(N_{f,s}^{ret} - r_{kf,s}\right)$, where $N_{f,s}^{ret}$ is the number of returnees entering any Chinese institution in field f in year s, and $r_{kf,s}$ is the number entering institution k in that field and year. Predicted absolute exposure is $C_{kf,t-1} = a_{kf} G_{f,t-1}^{-k}$, and the excluded instrument is

$$Z_{kf,t-1}^{cap} = \ln\left(1 + C_{kf,t-1}\right). \tag{5}$$

The corresponding endogenous treatment is $X_{kf,t-1} = \ln(1 + R_{kf,t-1})$, where $R_{kf,t-1}$ is the cumulative observed returnee stock in institution k and field f through year $t-1$. We apply the same log-plus-one transformation to observed stock and predicted exposure to allow diminishing marginal exposure while retaining zero-exposure cells (Chen and Roth, 2024).[19]

The exclusion restriction requires that, conditional on the controls and fixed effects below, an institution's pre-policy share of national capacity in a field affects subsequent local-author quality only through the institution–field returnee stock when interacted with the national leave-own-institution-out return wave. Field-by-year fixed effects absorb all national field-specific annual shocks. They do not absorb institution-specific responses to those shocks, such as funding targeted differentially toward particular institutions within a field; such responses remain the principal threat to the exclusion restriction.

Baseline estimates. Section 4.1 shows that scholars who directly coauthor with returnees experience large quality gains. To examine institutional exposure outside observed direct collaboration, we restrict to author–field–years in which the local scholar had not yet coauthored with either a returnee or a non-local international scholar by year $t-1$ (the

[19] Their concern is that treatment effects on log-like transformations of outcomes with zeros are sensitive to the outcome's units. Our use differs because the transformation is applied to a count-valued treatment and instrument, whose natural unit is one returnee, rather than to the outcome. We therefore interpret β^{IV} per unit change in log-plus-one returnee exposure, not as a unit-invariant percentage effect.

no-prior-collaboration risk set). Following the field-specific construction in Fact 4, an author publishing in multiple fields contributes a separate author–field history to the panel. The first-stage specification is:

$$X_{kf,t-1} = \pi Z^{cap}_{kf,t-1} + \mathbf{W}'_{jkt}\boldsymbol{\delta} + \mu_{jf} + \eta_{kf} + \lambda_{ft} + u_{jkft}, \tag{6}$$

where μ_{jf}, η_{kf}, and λ_{ft} denote author-by-field, institution-by-field, and field-by-year fixed effects. The control vector $\mathbf{W}_{jkt}$ includes author- and institution-level pre-policy international-collaboration rates interacted with t, lagged provincial log GDP per capita, and lagged provincial log public-budget education expenditure. For the exposure regressions, each pre-policy rate averages annual international-collaboration indicators over observed local-author–years in 2003–2005; an author or institution without a pre-policy observation is assigned zero. International collaboration requires an actual coauthor with a known non-China location and is constructed before institution-mapping restrictions.[20] Standard errors are clustered two-way, by author and by institution.

The baseline second-stage specification is:

$$\log AQ_{jft} = \beta^{\mathrm{IV}} \hat{X}_{kf,t-1} + \mathbf{W}'_{jkt}\boldsymbol{\gamma} + \mu_{jf} + \eta_{kf} + \lambda_{ft} + \varepsilon_{jkft}, \tag{7}$$

where AQ_{jft} is cumulative citation-share-based Author Quality within author j and field f. The field-by-year effects subsume separate field and publication-year effects. Identification comes from differential changes in predicted exposure across institutions within the same field and publication year, after removing persistent author–field and institution–field heterogeneity.

The regression retains 35,363 author–field–year observations, covering 13,644 scholars at 194 institutions. Table 2 reports a strong and positive first stage: the coefficient is 0.283 (s.e. 0.045), with a Kleibergen–Paap F-statistic of 40.54. Institution–field cells with larger pre-policy shares of national field capacity therefore receive more returnees when the national return wave in that field expands.

The reduced-form coefficient is 0.035 (s.e. 0.019, $p = 0.070$), and the corresponding 2SLS coefficient is 0.122 (s.e. 0.061, $p = 0.047$). The conventional 2SLS test is just below the 5% threshold, while the weak-instrument-robust Anderson–Rubin test has $p = 0.070$, so the evidence remains close to conventional significance thresholds. A one-unit increase in

[20] The regional controls are constructed from official province-year data. GDP per capita uses the National Bureau of Statistics' fifth-economic-census-revised series. Total public-budget education expenditure uses China Statistical Yearbook/NBS education-finance data, cross-checked against Ministry of Education funding bulletins. Institutions are matched to provinces using the institution–province crosswalk; amounts are expressed in current RMB where needed, then logged and lagged one year.

log-plus-one institution–field returnee stock is associated with a 0.122-log-point increase in cumulative field-specific Author Quality for variation induced by the instrument. Under the IV assumptions, this coefficient describes the response to exposure changes predicted by pre-policy field capacity and the national return wave. With continuous exposure, it has an instrument-induced response interpretation. Our hypothesis is that this estimate reflects intangible gains through seminars, mentoring, resource-sharing, or other institutional channels.

TABLE 2
Institutional Returnee Exposure for Locals Without Prior Collaboration

	(1) OLS	(2) First stage	(3) Reduced form	(4) 2SLS
$\ln(1 + R_{kf,t-1})$	0.086***			0.122**
	(0.021)			(0.061)
$Z^{cap}_{kf,t-1}$		0.284***	0.035*	
		(0.045)	(0.019)	
First-stage / KP F		40.54		40.54
KP underidentification LM				16.95
Anderson–Rubin $F(\beta = 0)$				3.31
Observations	35,363	35,363	35,363	35,363
Scholars	13,644	13,644	13,644	13,644
Institutions	194	194	194	194
Author–field FE	Yes	Yes	Yes	Yes
Institution–field FE	Yes	Yes	Yes	Yes
Field–year FE	Yes	Yes	Yes	Yes
Pre-policy international trends	Yes	Yes	Yes	Yes
Regional economic controls	Yes	Yes	Yes	Yes

Notes: The sample covers 2005–2022 and requires neither type of collaboration before the observation year. First-collaboration-year outcomes can remain in the sample. The unit is a local-scholar–field–year, and the outcome is log cumulative field-specific Author Quality. Exposure is $X_{kf,t-1} = \ln(1 + R_{kf,t-1})$, instrumented by $Z^{cap}_{kf,t-1}$ from Equation (5). Qualifying returnee coauthors are U.S.-to-China returnees in or after their return year; international coauthors have a known assigned location outside China. Histories are defined across fields. Controls include author- and institution-level 2003–2005 international-collaboration rates interacted with publication year and lagged provincial log GDP per capita and education expenditure. Standard errors are clustered by scholar and institution. The KP underidentification test has $p < 0.001$; the Anderson–Rubin test has $p = 0.070$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To examine whether the institutional-exposure relationship persists strictly before direct collaboration begins, we censor each scholar's observations from the first year of coauthorship with either a returnee or a non-local international scholar onward. This removes 6,470 first-collaboration-year observations; after fixed-effect singleton drops, the stricter regression uses 24,484 author–field–years. Appendix Table A7 reports a 2SLS coefficient of 0.097 (s.e. 0.082, $p = 0.241$), while the Kleibergen–Paap statistic remains strong at 42.95. The smaller

citation-quality estimate may reflect selection into collaboration: stronger scholars may be more likely to collaborate with returnees or international authors, so censoring observations once collaboration begins may leave a weaker pool.

Using log cumulative fractional publications as an alternative outcome, we see positive exposure coefficients in both the baseline sample (0.060, s.e. 0.045, $p = 0.190$) and the strictly pre-collaboration sample (0.103, s.e. 0.058, $p = 0.078$). Full results are reported in Appendix Table A8. This exercise provides suggestive evidence of an output benefit among scholars who have not yet entered either network.

Robustness Checks. We report three sets of diagnostics, detailed in Appendix A.10, that address distinct concerns about the capacity-aligned design. First, institution–field cells with greater pre-policy capacity shares might already have been on different quality trajectories before the post-2009 returnee inflow. To examine pre-policy trajectories over a longer horizon than the 2005–2022 IV sample, we recover the 2003–2022 publication histories of scholars who enter the corrected sample above and retain the same no-prior-collaboration author–field–year risk set for the diagnostic. Let $\widetilde{A}_{kf}$ denote the standardized log of institution k's pre-policy national capacity share in field f. We estimate:

$$\log AQ_{jft} = \sum_{\tau \neq 2008} \beta_\tau \left(\widetilde{A}_{kf} \times \mathbf{1}[t = \tau] \right) + \delta_1 \left(\overline{PreIntl}_j \times t \right) + \delta_2 \left(\overline{PreIntl}_k \times t \right) + \mathbf{Q}'_{k,t-1}\boldsymbol{\gamma} + \mu_{jf} + \eta_{kf} + \lambda_{ft} + \varepsilon_{jkft},$$

where $\mathbf{Q}_{k,t-1}$ contains the two lagged regional controls. The specification uses the 2004–2022 sample, omits 2008, and includes the same trends and fixed effects as the preferred model. Appendix Figure A8 plots the estimates. The four pre-policy coefficients for 2004–2007 are jointly insignificant, with $F = 1.046$ and $p = 0.384$, providing no joint evidence that higher-capacity institution–field cells were already on different quality trajectories before the policy period. A separate linear pre-trend and a six-year-lead placebo are also insignificant.

Second, we ask which fields provide the identifying variation. The conventional linear Rotemberg decomposition used for shift-share instruments (Goldsmith-Pinkham et al., 2020) is not mechanically available because $Z^{cap}_{kf,t-1}$ applies a nonlinear log transformation and each observation belongs to one field. We instead residualize the total log instrument and four mutually exclusive field components using the exact baseline controls and fixed effects, and calculate each component's share of the residualized first-stage covariance. Appendix Table A9 reports this Rotemberg-style decomposition. General receives 5.0% of the residualized first-stage covariance, Chemistry 20.1%, Physics 75.0%, and Life Sciences effectively zero weight. The component estimates are −0.406 for General, 0.058 for Chemistry, and

0.163 for Physics; the Life Sciences estimate is effectively unidentified. The weighted contributions reconstruct the pooled coefficient of 0.122 exactly. Thus, the positive relationship is driven predominantly by Physics rather than evenly by the four fields.

Third, because the weighting decomposition is concentrated in Physics, we test whether any field is individually essential to the result. In the author–field panel, the natural leave-one-out exercise deletes observations belonging to one field rather than subtracting one component from an aggregate institution-level instrument. Appendix Table A10 reports these estimates. Relative to the all-field coefficient of 0.122, the estimate is 0.149 without General, 0.145 without Chemistry, −0.028 without Physics, and 0.116 without Life Sciences. The instrument remains strong in every case, with KP statistics between 33.97 and 40.22. Removing Physics eliminates the pooled positive estimate, whereas excluding any other field leaves it positive and similar in magnitude.

Overall, the diagnostics show no evidence of differential pre-trends and confirm that the capacity-aligned instrument remains strong under the field exclusions. They also delimit the strength of the evidence: the pooled relationship is predominantly a Physics result and is not robust to removing that field.

4.3 Institutional Exposure among Already-Connected Scholars

We next ask whether institutional returnee exposure is associated with additional quality gains among local scholars who already have collaboration links to returnees or non-local international scholars. We retain author–field–years with at least one qualifying collaboration before year t: $D^{ret}_{j,t-1} = 1$ or $D^{intl}_{j,t-1} = 1$. These are lagged absorbing indicators, so scholars can contribute observations to Section 4.2 before entering either network and to this section afterward. The connected sample contains 47,147 author–field–years from 11,485 scholars at 198 institutions.

Specifically, we retain the outcome, exposure, capacity-aligned instrument, controls and fixed effects of Section 4.2, adding the two prior-collaboration indicators. The outcome specification is

$$\begin{aligned}\log AQ_{jft} = \beta \widetilde{X}_{kf,t-1} + \theta_{ret} D^{ret}_{j,t-1} + \theta_{intl} D^{intl}_{j,t-1} \\ + \mathbf{W}'_{jkt}\boldsymbol{\gamma} + \mu_{jf} + \eta_{kf} + \lambda_{ft} + \varepsilon_{jkft},\end{aligned} \tag{8}$$

where $\widetilde{X} = X$ for OLS and $\widetilde{X} = \hat{X}$ for 2SLS, with X instrumented by $Z^{cap}_{kf,t-1}$. The first stage includes the same prior-collaboration indicators, controls and fixed effects. Both panels use exactly the same connected sample and two-way clustering by scholar and institution.

Panel A reports a positive first-stage coefficient of 0.251 (s.e. 0.048), with a KP statistic

of 27.93, indicating that predicted exposure remains informative about institution–field returnee stock in the connected sample. However, the reduced-form coefficient is small and imprecise at 0.013 (s.e. 0.019, $p = 0.516$), as is the 2SLS exposure coefficient of 0.050 (s.e. 0.083, $p = 0.549$). We therefore find no statistically detectable additional quality gain associated with institutional exposure after accounting for collaboration history. By contrast, the 2SLS coefficients on prior returnee and international collaboration are positive and precisely estimated at 0.167 and 0.217. Because every observation already has at least one connection, these coefficients are identified from scholars acquiring the other type of connection, for example, an internationally connected scholar beginning to collaborate with returnees. Economically, these associations are consistent with access to both networks expanding the expertise and research resources available to a scholar. They do not identify causal returns to network entry, since collaboration histories are not instrumented and may reflect endogenous matching.

Panel B extends Equation (8) to account for annual collaboration intensity. Let $A^{g}_{j,t-1}$ be the number of distinct coauthors of type $g \in \{ret, intl\}$ active in the preceding calendar year, including repeat collaborators. Each annual count is zero when no qualifying joint publication occurs in that year; a scholar can therefore have prior history $D^{g}_{j,t-1} = 1$ but no active coauthor of that type. We add the two terms $\phi_{ret}(D^{ret}_{j,t-1}A^{ret}_{j,t-1})$ and $\phi_{intl}(D^{intl}_{j,t-1}A^{intl}_{j,t-1})$ to the outcome specification and include the corresponding intensity controls in the first stage.

Adding annual intensity in Panel B leaves the institutional-exposure estimate small and imprecise at 0.034 (s.e. 0.076, $p = 0.650$), while the KP statistic remains 28.25. The history coefficients are 0.134 for returnee collaboration and 0.202 for international collaboration; each describes the conditional history association when the corresponding annual coauthor count is zero. Conditional on these histories, each additional returnee coauthor active in the preceding year is associated with 0.035 log points higher cumulative field-specific quality, and each additional active international coauthor with 0.0047 log points; both coefficients have $p < 0.001$. Thus, annual collaboration activity is positively associated with quality beyond the recorded history of network entry. These coefficients remain descriptive because neither collaboration history nor intensity is instrumented.

To assess whether the small exposure estimates arise from conditioning on collaboration, Appendix Table A11 omits the collaboration-history and intensity controls while retaining the same sample and baseline controls. The exposure coefficient is already small and imprecise at 0.042 (s.e. 0.078), so adding collaboration controls does not reveal a substantial attenuation of the estimate.

Finally, we compare these exposure estimates for already-connected scholars with the

larger estimate of 0.122 for scholars without prior connections in Section 4.2. A joint IV comparison allowing controls and fixed effects to differ across the two groups does not reject equal exposure coefficients, either without collaboration controls ($p = 0.429$) or with connected-sample history and intensity controls ($p = 0.376$). See Appendix Table A12. Thus, the difference in point estimates does not establish that scholars without prior connections benefit more from institutional exposure.

TABLE 3
Institutional Returnee Exposure among Already-Connected Scholars

	(1) OLS	(2) First stage	(3) Reduced form	(4) 2SLS
Panel A: Controlling for collaboration history				
$\ln(1+R_{kf,t-1})$	0.022			0.050
	(0.017)			(0.083)
$Z^{cap}_{kf,t-1}$		0.251***	0.013	
		(0.048)	(0.019)	
Prior returnee collaboration	0.169***	0.066***	0.170***	0.167***
	(0.015)	(0.012)	(0.015)	(0.019)
Prior international collaboration	0.218***	0.021**	0.219***	0.217***
	(0.014)	(0.010)	(0.014)	(0.014)
First-stage / KP F		27.93		27.93
KP underidentification LM				15.61
Anderson–Rubin $F(\beta=0)$				0.42
Panel B: Adding annual collaboration intensity				
$\ln(1+R_{kf,t-1})$	0.022			0.034
	(0.017)			(0.076)
$Z^{cap}_{kf,t-1}$		0.251***	0.009	
		(0.047)	(0.018)	
Prior returnee collaboration	0.135***	0.061***	0.136***	0.134***
	(0.013)	(0.011)	(0.013)	(0.016)
Prior international collaboration	0.202***	0.022**	0.203***	0.202***
	(0.013)	(0.010)	(0.013)	(0.013)
Returnee history × active returnees	0.036***	0.005**	0.036***	0.035***
	(0.004)	(0.003)	(0.004)	(0.004)
International history × active intl. coauthors	0.0046***	-0.0008***	0.0046***	0.0047***
	(0.0011)	(0.0001)	(0.0011)	(0.0011)
First-stage / KP F		28.25		28.25
KP underidentification LM				15.46
Anderson–Rubin $F(\beta=0)$				0.23
Observations	47,147	47,147	47,147	47,147
Scholars	11,485	11,485	11,485	11,485
Institutions	198	198	198	198
Author–field FE	Yes	Yes	Yes	Yes
Institution–field FE	Yes	Yes	Yes	Yes
Field–year FE	Yes	Yes	Yes	Yes
Pre-policy international trends	Yes	Yes	Yes	Yes
Regional economic controls	Yes	Yes	Yes	Yes

Notes: Both panels use the same connected sample over 2005–2022: scholars had at least one qualifying returnee or international collaboration before the observation year. Panel A includes lagged absorbing history indicators; Panel B also interacts each history with the number of distinct corresponding coauthors active in the previous calendar year, including repeat collaborators. A year without qualifying joint publications has zero active coauthors. The unit is a local-scholar–field–year, and the outcome is log cumulative field-specific Author Quality. Exposure is $X_{kf,t-1}=\ln(1+R_{kf,t-1})$, instrumented by $Z^{cap}_{kf,t-1}$ from Equation (5). Qualifying returnee coauthors are U.S.-to-China returnees in or after their return year; international coauthors have a known assigned location outside China. Histories are defined across fields. Controls include author- and institution-level 2003–2005 international-collaboration rates interacted with publication year and lagged provincial log GDP per capita and education expenditure. Standard errors are clustered by scholar and institution. The KP underidentification tests have $p<0.001$; Anderson–Rubin p-values are 0.516 and 0.631. Collaboration coefficients are conditional associations. The unadjusted benchmark is in Appendix Table A11. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

4.4 Institutional Research Capacity and Output

The preceding analyses find a positive quality response among observed local scholars. We aggregate the current local-scholar population to the institution–year level and examine whether the author-level response appears in institutional outcomes. Before aggregating, we retain only author–field–year observations supported by the current capacity-aligned IV. We include both previously connected and unconnected scholars, as well as scholars observed publishing in only one year, because their participation and output contribute to institutional research capacity. This aggregation precedes author-level fixed-effect singleton removal and therefore does not inherit the estimation samples of Sections 4.2 and 4.3. We then sum exposure and outcomes across those supported fields, producing a balanced panel of 3,240 institution–years at 180 institutions.

We estimate reduced-form specifications using the same capacity-aligned exposure as in Sections 4.2 and 4.3:

$$Y_{kt} = \rho Z^{cap}_{k,t-1} + \eta_k + \lambda_t + \mathbf{X}'_{k,t-1}\delta + \varepsilon_{kt}, \tag{9}$$

where $Z^{cap}_{k,t-1} = \ln(1+\sum_f a_{kf} G^{-k}_{f,t-1})$ aggregates predicted exposure over the current IV's supported fields, η_k denotes institution fixed effects, and λ_t denotes year fixed effects. The vector $\mathbf{X}_{k,t-1}$ contains the lagged regional controls and the institution-level pre-policy international-network trend used above. Standard errors are clustered by institution.

Table 4 reports six outcomes in levels. Mean cumulative citation quality among pre-policy incumbents, defined as local scholars first observed by 2008, captures the intensive-margin performance of the established research workforce. Active local authors measure the institution's annual research workforce, fractional publications measure research volume, and citation-weighted output measures aggregate scientific impact. We also distinguish annual entrants—authors first observed in the current year—from the active post-2008 cohort, which counts all currently active authors whose first sampled publication occurred after 2008. For readability, mean cumulative quality is multiplied by 10^5 and citation-weighted output by 10^3; the other outcomes remain in counts or fractional-publication units.

The level estimates are positive across all six outcomes and are statistically significant at the 1% level. Because these are reduced-form level associations, their magnitudes are expressed in outcome units rather than percentage changes and should not be interpreted as causal effects without the identifying assumptions of the exposure design.

For policy, these margins answer different questions. Taken together, the level results are consistent with broader institutional expansion: higher predicted exposure is associated with a larger active local-author pool, more annual entrants, more post-2008 participants, greater fractional publication volume, and higher aggregate citation-weighted output, alongside a

positive association with incumbent quality. These institution-level results complement the author-level evidence by indicating that the exposure relationship extends beyond average quality among already observed scholars to participation and aggregate research output.

TABLE 4
Capacity-Aligned Exposure and Institutional Research Outcomes in Levels

	(1) Mean incumbent cumulative quality ($\times 10^5$)	(2) Active local authors	(3) Fractional publications	(4) Citation-weighted output ($\times 10^3$)	(5) Annual entrants	(6) Active post-2008 cohort
$Z^{cap}_{k,t-1}$	4.630***	45.735***	5.060***	1.129***	22.411***	64.964***
	(1.020)	(9.750)	(1.198)	(0.226)	(4.453)	(14.150)
Observations	2,409	3,240	3,240	3,240	3,240	3,240
Institutions	174	180	180	180	180	180
Institution fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional controls	Yes	Yes	Yes	Yes	Yes	Yes
International-network trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The unit of observation is an institution–year. The balanced current-IV-native panel contains 3,240 institution–years at 180 institutions; Column 1 has fewer observations because mean incumbent quality is unavailable in institution–years without an eligible pre-2009 incumbent quality observation. The aggregation retains eligible local scholars regardless of prior collaboration status or the number of observed publication years, including those observed in only one year, and precedes author-level fixed-effect singleton removal. Each column reports a reduced-form regression of a level outcome on the lagged aggregate capacity-aligned instrument $Z^{cap}_{k,t-1} = \ln(1 + \sum_f a_{kf} G^{-k}_{f,t-1})$, constructed over the current IV's supported fields. Column 1 reports 10^5 times mean cumulative citation-share quality among local scholars first observed by 2008. Columns 2 and 5 are counts of active local authors and annual entrants, respectively; Column 6 counts active local authors first observed after 2008. Column 3 is fractional publications. Column 4 reports 10^3 times total annual normalized citation contribution. All specifications include institution and year fixed effects, lagged provincial controls, and the institution-level pre-policy international-network trend, measured from coauthors' publication-year country affiliations. Standard errors, clustered by institution, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5 Conclusion

This paper examines how brain circulation contributes to frontier scientific development by studying the return of overseas Chinese scientists as China rises to be a significant global scientific contributor. Using bibliometric data from leading scientific journals from 2003 to 2022, we document a profound transformation in China's position in frontier science and investigate the role of returning scientists in that transformation.

Our findings lead to four main conclusions. First, China's growing presence in elite scientific journals has been accompanied by a substantial increase in the return of overseas

Chinese scientists, many from the United States. Second, these returnees are positively selected for research quality, though the degree of selection varies markedly across scientific fields. Third, post-return performance is similarly heterogeneous: returning chemists outperform both matched U.S.-based comparison groups, returning life scientists and General-journal authors underperform, and returning physicists perform similarly to ethnically Chinese U.S. stayers while showing somewhat weaker performance relative to non-Chinese stayers. Finally, the relationship between brain circulation and local scientific performance extends beyond the returnees themselves. Local scholars exhibit a positive and persistent premium in cumulative citation-share-based quality after first collaborating with a returnee. The immediate increase is concentrated in joint papers, while later gains in non-joint work depend on the specification. Capacity-aligned institutional exposure is associated with suggestive quality gains among scholars without prior returnee or international collaboration, but the pooled positive estimate is driven predominantly by Physics and becomes imprecise when first-collaboration years are excluded. The same exposure specification yields a small and imprecise estimate among already-connected scholars; a joint test does not establish a difference between the groups. Adjustment for collaboration history and intensity leaves that connected-sample result largely unchanged. At the institution level, predicted exposure is positively associated with incumbent quality, active local authors, annual entrants into the observed publication panel, active post-2008 participants, fractional publication volume, and citation-weighted output. These reduced-form associations are consistent with broader institutional expansion, subject to the exposure design's identifying assumptions.

More broadly, our results suggest that the economic value of high-skilled return migration depends not simply on bringing talented individuals home, but also on how they interact with local scientific communities. Gains embodied in joint research are consistent with returnees contributing expertise and research inputs through shared projects. The institutional-exposure results are also consistent with broader benefits, although the evidence does not separately identify mentoring, resource sharing, competition, or knowledge diffusion. Nor does the comparison across collaboration groups establish that existing networks reduce the marginal benefit of institutional exposure. The considerable heterogeneity across scientific fields further cautions against treating brain circulation as a uniform source of gains across research environments.

Although our analysis focuses on China, the implications extend well beyond a single country. As governments around the world increasingly compete for highly skilled scientists through immigration policies, diaspora engagement, and talent-recruitment programs, understanding how knowledge moves with people has become a central question in the economics of innovation. Our evidence suggests that evaluating such policies requires looking

beyond the careers of recruited scientists themselves to their associations with collaborators, incumbent scholars, workforce participation, and aggregate institutional output.

Several questions remain open. We measure scientific impact through publications in elite journals, but future work could investigate whether the collaboration and institutional relationships documented here extend to technological innovation, patenting, entrepreneurship, industrial productivity, or broader economic growth. More evidence is also needed on the organizational mechanisms through which returnees influence their institutions, including mentoring, laboratory formation, graduate-student training, and research leadership. Finally, our results reveal substantial differences across scientific fields that remain only partially understood. Explaining how research environments shape the relationship between brain circulation and sustained scientific advancement remains an important challenge for future research.

Ultimately, the traditional distinction between "brain drain" and "brain gain" is too narrow to capture the dynamics of modern scientific development. The international mobility of scientists can matter not only by reallocating talent across countries, but also by reshaping collaborative networks and the research environments into which scientists move. Our findings indicate that these broader relationships are visible in collaborator quality, workforce participation, aggregate output, and the performance and composition of receiving institutions, although the institution-level evidence remains reduced-form and does not identify the organizational mechanisms behind these relationships.

References

Agarwal, R., and Gaule, P. (2020). 'Invisible Geniuses: Could the Knowledge Frontier Advance Faster?', *American Economic Review: Insights*, Vol. 2(4), pp. 409–24. https://doi.org/10.1257/aeri.20190457

Agrawal, A., Kapur, D., McHale, J., and Oettl, A. (2011). 'Brain drain or brain bank? The impact of skilled emigration on poor-country innovation', *Journal of Urban Economics*, Vol. 69(1), pp. 43–55. https://doi.org/10.1016/j.jue.2010.06.003

Akcigit, U., Baslandze, S., and Stantcheva, S. (2016). 'Taxation and the international mobility of inventors', *American economic review*, Vol. 106(10), pp. 2930–2981.

Bassetto, J., and Ippedico, G. (2024). *Tax Incentives and Return Migration* (IZA Discussion Paper No. 17224). IZA Institute of Labor Economics. https://docs.iza.org/dp17224.pdf

Batista, C., Han, D., Haushofer, J., Khanna, G., McKenzie, D., Mobarak, A. M., Theoharides, C., and Yang, D. (2025). 'Brain drain or brain gain? Effects of high-skilled international emigration on origin countries', *Science*, Vol. 388(6749), pp. eadr8861. https://doi.org/10.1126/science.adr8861

Becker, S. O., Lindenthal, V., Mukand, S. W., and Waldinger, F. (2024). 'Persecution and escape: professional networks and high-skilled emigration from nazi Germany', *American Economic Journal: Applied Economics*, Vol. 16(3), pp. 1–43.

Beine, M., Docquier, F., and Rapoport, H. (2008). 'Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers', *The Economic Journal*, Vol. 118(528), pp. 631–652. https://doi.org/10.1111/j.1468-0297.2008.02135.x

Bursztyn, L., Chaney, T., Hassan, T. A., and Rao, A. (2024). 'The Immigrant Next Door', *American Economic Review*, Vol. 114(2), pp. 348–384. https://doi.org/10.1257/aer.20220376

Cao, C. (2008). 'China's brain drain at the high end: why government policies have failed to attract first-rate academics to return', *Asian population studies*, Vol. 4(3), pp. 331–345.

Cao, C., Baas, J., Wagner, C. S., and Jonkers, K. (2020). 'Returning scientists and the emergence of China's science system', *Science and Public Policy*, Vol. 47(2), pp. 172–183.

Cao, C., Suttmeier, R. P., and Simon, D. F. (2006). 'China's 15-year science and technology plan', *Physics Today*, Vol. 59(12), pp. 38–43. https://doi.org/10.1063/1.2435680

Central Organization Department of the Communist Party of China. (2011). *Detailed Rules for the "Thousand Talents Program" High-Level Foreign Expert Project* (Translated

by Etcetera Language Group / Ben Murphy (CSET), 8 July 2020. Original: Notice No. 45 [2011]). Center for Security and Emerging Technology. https://cset.georgetown.edu/publication/detailed-rules-for-the-thousand-talents-program-high-level-foreign-expert-project/

Chen, J., and Roth, J. (2024). 'Logs with Zeros? Some Problems and Solutions', *The Quarterly Journal of Economics*, Vol. 139(2), pp. 891–936. https://doi.org/10.1093/qje/qjad054

Docquier, F., and Rapoport, H. (2012). 'Globalization, Brain Drain, and Development', *Journal of Economic Literature*, Vol. 50(3), pp. 681–730. https://doi.org/10.1257/jel.50.3.681

Doran, K., Gelber, A., and Isen, A. (2022). 'The effects of high-skilled immigration policy on firms: Evidence from visa lotteries', *Journal of Political Economy*, Vol. 130(10), pp. 2501–2533.

Finn, M. G. (2010). *Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2007* (tech. rep.) (Five-year stay rate for the 2002 Chinese doctorate cohort: 92%, the highest of any country). Oak Ridge Institute for Science and Education. https://www.osti.gov/servlets/purl/971617

Flynn, R., Glennon, B., Murciano-Goroff, R., and Xiao, J. (2024, June). *Building a Wall Around Science: The Effect of U.S.-China Tensions on International Scientific Research* (Working Paper No. 32622). National Bureau of Economic Research. https://doi.org/10.3386/w32622

Fry, C. V. (2023). 'Bridging the Gap: Evidence from the Return Migration of African Scientists', *Organization Science*, Vol. 34(1), pp. 404–432. https://doi.org/10.1287/orsc.2022.1580

Gaulé, P. (2014). 'Who comes back and when? Return migration decisions of academic scientists', *Economics Letters*, Vol. 124(3), pp. 461–464. https://doi.org/10.1016/j.econlet.2014.07.014

Gaulé, P., and Piacentini, M. (2013). 'Chinese Graduate Students and U.S. Scientific Productivity', *The Review of Economics and Statistics*, Vol. 95(2), pp. 698–701. https://doi.org/10.1162/REST_a_00283

Gibson, J., and McKenzie, D. (2011). 'Eight Questions about Brain Drain', *Journal of Economic Perspectives*, Vol. 25(3), pp. 107–128. https://doi.org/10.1257/jep.25.3.107

Giorcelli, M., Huang, Y., and Li, B. (2026). *The Return(s) of Superstar Scientists* (Working Paper) (April 30, 2026). University of California, Los Angeles and Peking University. Retrieved July 22, 2026, from https://drive.google.com/file/d/1jmG8zrDL9dna4ai5Y5G-qOPnyO7ec8fn/view

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). 'Bartik instruments: What, when, why, and how', *American Economic Review*, Vol. 110(8), pp. 2586–2624.

Gu, J., Pan, X., Zhang, S., and Chen, J. (2024). 'International mobility matters: Research collaboration and scientific productivity', *Journal of Informetrics*, Vol. 18(2), pp. 101522. https://doi.org/10.1016/j.joi.2024.101522

Hager, S., Schwarz, C., and Waldinger, F. (2024). 'Measuring Science: Performance Metrics and the Allocation of Talent', *American Economic Review*, Vol. 114(12), pp. 4052–4090. https://doi.org/10.1257/aer.20230515

Hunt, J., and Gauthier-Loiselle, M. (2010). 'How much does immigration boost innovation?', *American Economic Journal: Macroeconomics*, Vol. 2(2), pp. 31–56.

Jia, H. (2018). 'What is China's Thousand Talents Plan?', [NatureJobs Career Guide: China]. *Nature*. https://doi.org/10.1038/d41586-018-00538-z

Jia, R., Roberts, M. E., Wang, Y., and Yang, E. (2024). 'The impact of US–China tensions on US science: Evidence from the NIH investigations', *Proceedings of the National Academy of Sciences*, Vol. 121(19), pp. e2301436121.

Kahn, S., and MacGarvie, M. J. (2016). 'How Important Is U.S. Location for Research in Science?', *The Review of Economics and Statistics*, Vol. 98(2), pp. 397–414. https://doi.org/10.1162/REST_a_00490

Kerr, W. R. (2008). 'Ethnic Scientific Communities and International Technology Diffusion', *The Review of Economics and Statistics*, Vol. 90(3), pp. 518–537. https://doi.org/10.1162/rest.90.3.518

Kitajima, K., and Okamura, K. (2025). 'The altering landscape of US–China science collaboration: from convergence to divergence', *Humanities and Social Sciences Communications*, Vol. 12(302). https://doi.org/10.1057/s41599-025-04550-3

Liu, F.-c., Simon, D. F., Sun, Y.-t., and Cao, C. (2011). 'China's innovation policies: Evolution, institutional structure, and trajectory', *Research Policy*, Vol. 40(7), pp. 917–931. https://doi.org/10.1016/j.respol.2011.05.005

Ministry of Education of the People's Republic of China. (2020, December). Statistics on Chinese Learners Studying Overseas in 2019 [6.56 million Chinese pursued overseas study between 1978 and 2019; of the 4.9 million who completed their studies, 4.23 million (86.3%) returned to China]. http://en.moe.gov.cn/news/press_releases/202012/t20201224_507474.html

Moser, P., Voena, A., and Waldinger, F. (2014). 'German Jewish émigrés and US invention', *American Economic Review*, Vol. 104(10), pp. 3222–3255.

Nagaraj, A., and Yao, R. (2026, January). *The Geography of Science* (Working Paper No. 34694). National Bureau of Economic Research. https://doi.org/10.3386/w34694

National Center for Science and Engineering Statistics. (2025). *Most U.S.-Trained Science and Engineering Doctorate Recipients on Temporary Visas Remain in the United States* (tech. rep. No. NSF 25-325). National Science Foundation. https://ncses.nsf.gov/pubs/nsf25325

National Science Board. (2018). *Science and Engineering Indicators 2018* (tech. rep. No. NSB-2018-1). National Science Board. Alexandria, VA. https://www.nsf.gov/statistics/2018/nsb20181/permissions

Prato, M. (2025). 'The global race for talent: Brain drain, knowledge transfer, and growth', *The Quarterly Journal of Economics*, Vol. 140(1), pp. 165–238.

Qiu, S., Steinwender, C., and Azoulay, P. (2022). *Who Stands on the Shoulders of Chinese (Scientific) Giants? Evidence from Chemistry* (Working Paper). National Bureau of Economic Research. https://doi.org/10.3386/w30772

Qiu, S., Steinwender, C., and Azoulay, P. (2024). *Paper Tiger? Chinese Science and Home Bias in Citations* (Working Paper). National Bureau of Economic Research. https://doi.org/10.3386/w32468

Rieke, A., Southerland, V., Svirsky, D., and Hsu, M. (2022). 'Imperfect Inferences: A Practical Assessment', *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 767–777. https://doi.org/10.1145/3531146.3533140

Sanliturk, E., Zagheni, E., Dańko, M. J., Theile, T., and Akbaritabar, A. (2023). 'Global patterns of migration of scholars with economic development', *Proceedings of the National Academy of Sciences*, Vol. 120(4), pp. e2217937120. https://doi.org/10.1073/pnas.2217937120

Saxenian, A. (2005). 'From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China', *Studies in Comparative International Development*, Vol. 40(2), pp. 35–61. https://doi.org/10.1007/BF02686293

Sebo, P. (2023). 'How Well Does NamSor Perform in Predicting the Country of Origin and Ethnicity of Individuals Based on Their First and Last Names?', *PLOS ONE*, Vol. 18(11), pp. e0294562. https://doi.org/10.1371/journal.pone.0294562

Shi, D., Liu, W., and Wang, Y. (2023). 'Has China's Young Thousand Talents program been successful in recruiting and nurturing top-caliber scientists?', *Science*, Vol. 379(6627), pp. 62–65.

Shin, G.-W., and Moon, R. (2018, February). *From Brain Drain to Brain Circulation and Linkage* (tech. rep.). Freeman Spogli Institute for International Studies, Stanford University. https://fsi.stanford.edu/publication/brain-drain-brain-circulation-and-linkage

Sun, L., and Abraham, S. (2021). 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics*, Vol. 225(2), pp. 175–199.

Terry, S. J., Chaney, T., Burchardi, K. B., Tarquinio, L., and Hassan, T. A. (2026). 'Immigration, innovation, and growth', *American Economic Review*, Vol. 116(3), pp. 828–861.

Tollefson, J. (2018). 'China declared world's largest producer of scientific articles.' *Nature*, Vol. 553(7686), pp. 390–391. https://doi.org/10.1038/d41586-018-00927-4

UNESCO. (2025). Higher Education: Figures at a Glance [UNESCO Institute for Statistics data: 6.9 million internationally mobile tertiary students in 2022, up from 2.1 million in 2000]. https://unesdoc.unesco.org/ark:/48223/pf0000394112

Wagner, C. S., Zhang, L., and Leydesdorff, L. (2022). 'A discussion of measuring the top-1% most-highly cited publications: quality and impact of Chinese papers', *Scientometrics*, Vol. 127(4), pp. 1825–1839.

Waldinger, F. (2016). 'Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge', *Review of Economics and Statistics*, Vol. 98(5), pp. 811–831.

Wang, H. (2011). *China's National Talent Plan: Key Measures and Objectives* (Policy Brief). Brookings Institution. https://doi.org/10.2139/ssrn.1828162

Woolston, C. (2023). 'Nature Index Annual Tables 2023: China tops natural-science table', *Nature*. https://doi.org/10.1038/d41586-023-01868-3

Xie, Q., and Freeman, R. B. (2018). *Bigger Than You Thought: China's Contribution to Scientific Publications* (Working Paper). National Bureau of Economic Research. https://doi.org/10.3386/w24829

Xie, Q., and Freeman, R. B. (2020). *The Contribution of Chinese Diaspora Researchers to Global Science and China's Catching Up in Scientific Research* (Working Paper). National Bureau of Economic Research. https://doi.org/10.3386/w27169

Xie, Q., and Freeman, R. B. (2023). *Creating and Connecting US and China Science: Chinese Diaspora and Returnee Researchers* (Working Paper). National Bureau of Economic Research. https://doi.org/10.3386/w31306

Xie, Y., Lin, X., Li, J., He, Q., and Huang, J. (2023). 'Caught in the crossfire: Fears of Chinese–American scientists', *Proceedings of the National Academy of Sciences*, Vol. 120(27), pp. e2216248120. https://doi.org/10.1073/pnas.2216248120

Xie, Y., Zhang, C., and Lai, Q. (2014). 'China's rise as a major contributor to science and technology', *Proceedings of the National Academy of Sciences*, Vol. 111(26), pp. 9437–9442.

Xin, H. (2006). 'Frustrations Mount Over China's High-Priced Hunt for Trophy Professors', *Science*, Vol. 313(5794), pp. 1721–1723. https://doi.org/10.1126/science.313.5794.1721

Yoon, B.-S. L. (1992). 'Reverse brain drain in South Korea: State-led model', *Studies in Comparative International Development*, Vol. 27(1), pp. 4–26. https://doi.org/10.1007/BF02687102

Zhou, P., and Leydesdorff, L. (2006). 'The emergence of China as a leading nation in science', *Research policy*, Vol. 35(1), pp. 83–104.

Zweig, D., Changgui, C., and Rosen, S. (2004). 'Globalization and transnational human capital: Overseas and returnee scholars to China', *The China Quarterly*, Vol. 179, pp. 735–757.

Zweig, D., Siqin, K., and Huiyao, W. (2020). ''The best are yet to come:'state programs, domestic resistance and reverse migration of high-level talent to China', *Journal of Contemporary China*, Vol. 29(125), pp. 776–791.

Zweig, D., and Wang, H. (2013). 'Can China bring back the best? The Communist Party organizes China's search for talent', *The China Quarterly*, Vol. 215, pp. 590–615.

A Appendices

A.1 Complete Journal List

TABLE A1
Journal List

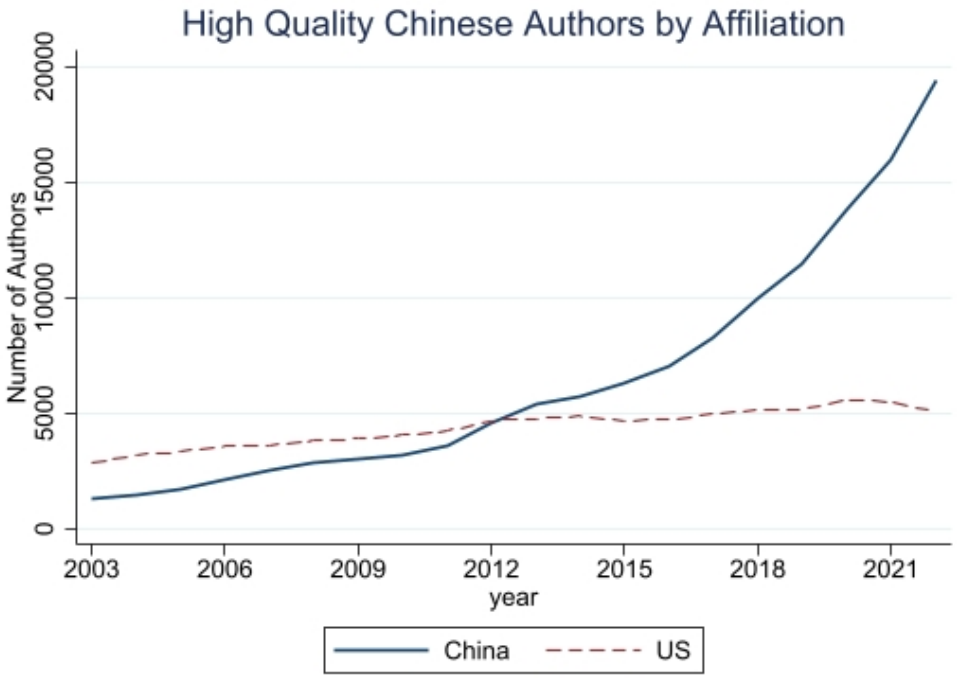
Journal Name	Published in	Impact Factor 2022
General		
Nature	UK	69.60
Science	USA	63.71
Proceedings of the National Academy of Sciences	USA	12.79
Chemistry		
Nature Materials	UK	47.66
Advanced Materials	Germany	32.09
Angewandte Chemie - International Edition in English	UK	16.82
Journal of the American Chemical Society	USA	16.38
Nano Letters	USA	12.62
Physics		
Physical Review Letters	USA	9.19
Astronomy and Astrophysics	Netherlands	6.24
Monthly Notices of The Royal Astronomical Society: Letters	UK	5.24
European Physical Journal C	Italy	4.99
Applied Physics Letters	USA	3.97
Life Sciences		
LANCET	UK	202.73
New England Journal of Medicine	USA	176.08
Journal of the American Medical Association	USA	157.34
Nature Medicine	UK	87.24
CELL	USA	66.85

A.2 Chinese Returnee Origin

TABLE A2
Chinese Moving to China by Origin

Year	U.S.	Non-Asia	Asia	Total
2004	13	10	8	31
2005	38	24	11	73
2006	52	38	19	109
2007	92	42	27	161
2008	93	51	18	162
2009	105	62	27	194
2010	169	82	33	284
2011	155	87	55	297
2012	246	120	58	424
2013	261	145	60	466
2014	280	148	58	486
2015	338	170	64	572
2016	367	177	76	620
2017	451	218	78	747
2018	472	276	73	821
2019	489	312	82	883
2020	586	399	76	1,061
2021	574	460	129	1,163
2022	723	558	128	1,409
Total	5,504	3,379	1,080	9,963

A.3 Composition of High-Quality Chinese Scientists in the U.S.



Note: This figure shows the number of high-quality Chinese authors affiliated with China and the U.S. over time. Each author's quality index is calculated using Equation (2). A "high-quality" author is one whose annual quality index exceeds the yearly median.

FIGURE A1
High-Quality Chinese Authors by Country of Affiliation

A.4 Return Probability and Above-Median Quality

Table A3 reports the all-field regression underlying the black dashed line in Figure 5. The above-median quality indicator is constructed using yearly medians that exclude post-return author-year observations.

TABLE A3
Return Probability and Above-Median Quality, All Fields

	(1) All fields
Above median x 2004	-0.00162
	(0.00138)
Above median x 2005	0.00362*
	(0.00194)
Above median x 2006	-0.000472
	(0.00218)
Above median x 2007	0.000530
	(0.00277)
Above median x 2008	0.00258
	(0.00258)
Above median x 2009	-0.00110
	(0.00273)
Above median x 2010	-0.00326
	(0.00332)
Above median x 2011	-0.00166
	(0.00301)
Above median x 2012	0.00895***
	(0.00328)
Above median x 2013	0.00541
	(0.00333)
Above median x 2014	0.00297
	(0.00340)
Above median x 2015	0.00773**
	(0.00384)
Above median x 2016	0.00906**
	(0.00394)
Above median x 2017	0.0124***
	(0.00411)
Above median x 2018	0.0116***
	(0.00398)
Above median x 2019	0.0200***
	(0.00401)
Above median x 2020	0.0134***
	(0.00413)
Above median x 2021	0.0163***
	(0.00416)
Above median x 2022	0.0276***
	(0.00486)
Observations	166602
Adjusted R-squared	0.034
Year fixed effects	Yes
Age proxy control	Yes
SE clustered by	Author
Median sample	Excludes post-return author-years

Notes: The dependent variable equals one when a Chinese scholar returns from the United States to China. High quality is defined using the yearly median cumulative publication quality after excluding post-return author-years from the median calculation. The 2003 interaction is omitted as the reference category. The model includes year fixed effects, the above-median indicator, and age proxy. Standard errors are clustered by author.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

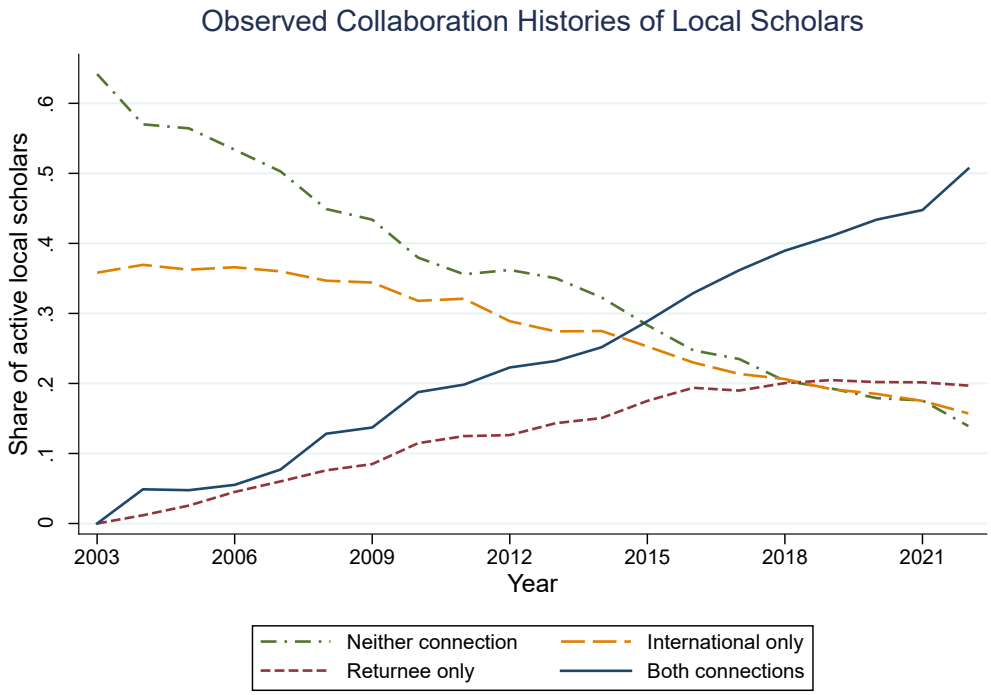
A.5 Field-Heterogeneous Post-Return Performance

TABLE A4
Post-Return Performance Relative to Matched U.S. Scientists

	Ethnically Chinese U.S. stayers	Non-Chinese U.S. stayers
Panel A: Return year through year 3		
General	−0.103***	−0.132***
	(0.016)	(0.019)
Chemistry	0.257***	0.331***
	(0.044)	(0.053)
Physics	−0.019	−0.103**
	(0.034)	(0.046)
Life Sciences	−0.234***	−0.383***
	(0.062)	(0.078)
Returnee–field pairs	4,445	4,445
Unique returnees	3,618	3,618
Panel B: Return year through year 5		
General	−0.112***	−0.139***
	(0.024)	(0.026)
Chemistry	0.627***	0.693***
	(0.081)	(0.093)
Physics	0.022	−0.094
	(0.042)	(0.058)
Life Sciences	−0.260***	−0.507***
	(0.095)	(0.166)
Returnee–field pairs	3,242	3,242
Unique returnees	2,658	2,658

Notes: Each entry reports the mean difference between a U.S.-to-China returnee's cumulative field-specific normalized citation contribution and that of a matched U.S.-based scientist over the indicated horizon, scaled by 10,000. Positive estimates indicate higher post-return performance by the returnee. Matching is with replacement, exact within field and return year, and based on cumulative pre-return quality, quality in the year before return, quality and publication counts over the preceding three years, and career age. Comparison scholars remain observed at U.S. institutions through the relevant horizon but may move afterward. Standard errors in parentheses are two-way clustered by returnee and matched-control identities. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.6 Evolution of Observed Local-Scholar Collaboration Histories



Notes: The eligible population consists of the 46,358 local scholars in Table 1, each observed publishing in at least two distinct years during 2003–2022. In each year, the denominator includes only members of this population observed publishing that year; scholars without a sampled publication that year are excluded from that year's counts and shares. The four mutually exclusive categories use collaboration histories observed from 2003 through the indicated year: neither qualifying connection, returnee only, international only, or both. A returnee connection is coauthorship with a returnee in or after the return year; an international connection is coauthorship with a non-local scholar in the joint-publication year. Connections remain recorded after their first occurrence, so scholars can change categories even though Table 1 assigns each scholar a fixed full-period category. "Neither" includes scholars who acquire connections later and those with no observed coauthors. Pre-2003 collaboration and migration histories are unavailable. Identifying a returnee requires an observed U.S.-to-China transition, so the earliest identifiable return is in 2004 and both returnee-related series start at zero in 2003 by construction. Changes reflect accumulated observed histories and turnover in the annual publishing population.

FIGURE A2
Evolution of Observed Collaboration Histories among Local Scholars

A.7 Direct-Collaboration Event-Study Robustness

The explicit risk set in Section 4.1 contains treated scholars whose international collaboration precedes their first returnee collaboration and internationally connected scholars who never coauthor with a returnee. Eligibility is imposed observation by observation, so an author enters only in or after the year of first international collaboration. Before outcome and control restrictions, the corrected risk set contains 75,006 author-year observations from 36,914 authors. Requiring the common outcome and controls leaves 60,411 observations from 28,445 authors; after the fixed-effect estimators remove singleton and unsupported observations, the main Sun–Abraham and TWFE sample contains 42,886 observations from 10,920 authors. Later event times are necessarily identified from earlier treatment cohorts.

Common-sample comparison. Table A5 re-estimates all three methods on the exact intersection of the corrected fixed-effect sample and the pair-supported Callaway–Sant'Anna sample: 42,197 observations from 10,839 authors. Sun–Abraham and TWFE are nearly identical on this sample, reaching 0.302 and 0.305 at $\tau = 3$. Callaway–Sant'Anna is modestly steeper after the first post-collaboration year, with estimates of 0.118, 0.269, 0.335, and 0.354 at event times 0 through 3. Thus, restricting estimation to a common observation pool does not eliminate the remaining difference in dynamics; the covariate adjustment and effective group-time comparisons remain estimator-specific. None of the three estimators rejects the joint null for the displayed leads ($p = 0.793$, 0.952, and 0.471 for Sun–Abraham, TWFE, and Callaway–Sant'Anna, respectively).

TABLE A5
Event-Study Estimators on an Exact Common Sample

Event time	Sun–Abraham	TWFE	Callaway–Sant'Anna
$\tau = -6$	-0.020	-0.006	0.024
	(0.031)	(0.035)	(0.028)
$\tau = -5$	-0.006	-0.008	-0.007
	(0.027)	(0.033)	(0.022)
$\tau = -4$	-0.003	0.001	0.010
	(0.023)	(0.023)	(0.023)
$\tau = -3$	-0.007	-0.005	0.017
	(0.020)	(0.024)	(0.021)
$\tau = -2$	0.016	0.014	0.032
	(0.016)	(0.021)	(0.018)
$\tau = 0$	0.099	0.103	0.118
	(0.013)	(0.015)	(0.014)
$\tau = 1$	0.220	0.225	0.269
	(0.015)	(0.016)	(0.020)
$\tau = 2$	0.256	0.259	0.335
	(0.018)	(0.018)	(0.028)
$\tau = 3$	0.302	0.305	0.354
	(0.021)	(0.022)	(0.031)
Joint lead-test p-value	0.793	0.952	0.471
Observations	42,197	42,197	42,197
Authors	10,839	10,839	10,839

Notes: All three estimators use the exact intersection of the corrected, observation-level risk set and the pair-supported Callaway–Sant'Anna sample. The 2004 treatment cohort is excluded because no prior base year is observed. Sun–Abraham and TWFE use the controls and clustering conventions from Figure 7. Callaway–Sant'Anna uses never-treated controls, varying base periods, pair balancing, regression adjustment for the lagged collaborator stocks and pre-policy network levels, and analytical pointwise standard errors. The joint tests cover the displayed leads from $\tau = -6$ through -2.

Control sensitivity. The lagged international-coauthor measure in Eq. (4) is a subset of the lagged all-coauthor measure. Table A6 therefore replaces the all-coauthor stock with a non-overlapping China-only stock, defined as the number of distinct coauthors not yet ob-

served collaborating from outside China. All other controls, the 42,886-observation sample, and the estimator-specific inference conventions are held fixed. The contemporaneous estimates are unchanged, while the estimates at event times 1 through 3 are approximately 25–30% smaller: under Sun–Abraham they decline from 0.189, 0.211, and 0.259 to 0.134, 0.148, and 0.192. TWFE yields almost the same alternative path. The post-collaboration coefficients remain positive and precisely estimated, and the displayed leads remain jointly insignificant under both estimators. The direction of the result is therefore robust, although its longer-run magnitude is sensitive to how prior collaborator exposure is parameterized.

TABLE A6
Alternative Definitions of Prior Coauthor Exposure

	Sun–Abraham		TWFE	
Event time	Intl. + all coauthors	Intl. + China-only	Intl. + all coauthors	Intl. + China-only
$\tau = 0$	0.077	0.077	0.068	0.067
$\tau = 1$	0.189	0.134	0.188	0.135
$\tau = 2$	0.211	0.148	0.219	0.159
$\tau = 3$	0.259	0.192	0.262	0.197
Joint lead-test p-value	0.871	0.761	0.882	0.951
Observations	42,886	42,886	42,886	42,886
Authors	10,920	10,920	10,920	10,920

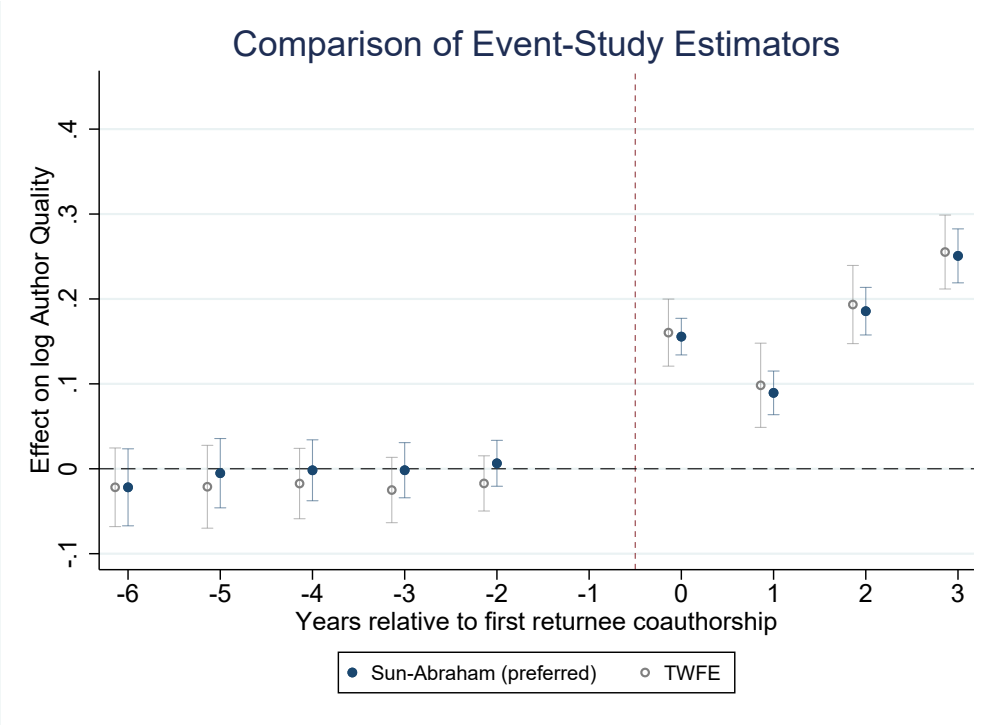
Notes: The baseline specification includes the lagged cumulative stocks of distinct non-local international coauthors and all distinct coauthors through year $t-1$. The alternative replaces the all-coauthor stock with the number of distinct coauthors not yet observed collaborating from outside China, calculated as the corrected all-coauthor stock minus the corrected international-coauthor stock. This produces international and China-only stocks that are non-overlapping by construction. Both specifications retain author and publication-year fixed effects and the two pre-policy international-collaboration-rate trends. Samples and inference conventions are identical within estimator. The joint tests cover the displayed leads from $\tau = -6$ through -2.

Broader unconditional benchmark. Figure A3 reports Sun–Abraham and TWFE estimates for the broader returnee design. The broader design uses the full always-China author sample and estimates:

$$\log AQ_{jt} = \sum_{\tau \neq -1} \beta_\tau \mathbf{1}\{\tau_j = \tau\} + \gamma \cdot \text{IHS}(C^{intl}_{j,t-1}) + \delta \cdot \text{IHS}(C^{all}_{j,t-1}) + \xi_1(t \times \overline{PreIntl}_j) + \xi_2(t \times \overline{PreIntl}_k) + \mu_j + \lambda_t + \varepsilon_{jt}. \quad (10)$$

Unlike the returnee-increment design, this specification imposes no restriction on prior international collaboration. The preferred post-collaboration estimates are 0.156, 0.089, 0.186,

and 0.251 at event times 0 through 3. The five displayed leads are jointly insignificant under Sun–Abraham ($p = 0.907$) and TWFE ($p = 0.862$).



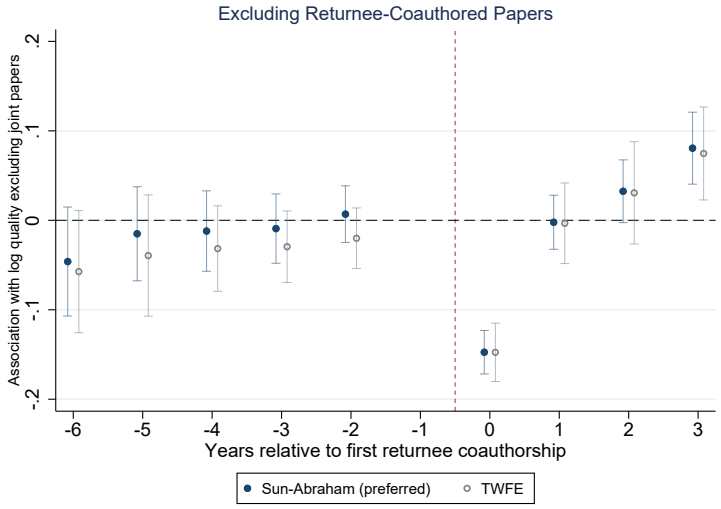
Notes: The figure compares the preferred Sun–Abraham estimator with conventional TWFE on a common complete-case sample of 125,748 observations from 37,406 authors. The specification controls for the lagged cumulative numbers of international and total coauthors, includes author and year fixed effects, and includes pre-policy international-collaboration-rate trends. Sun–Abraham standard errors are clustered by author; TWFE standard errors are clustered two-way by author and institution. Exact event times from −6 through 3 are shown, with event time −1 normalized to zero and omitted from the plot. Bars denote 95% confidence intervals.

FIGURE A3
Event Study in the Broader Returnee Design

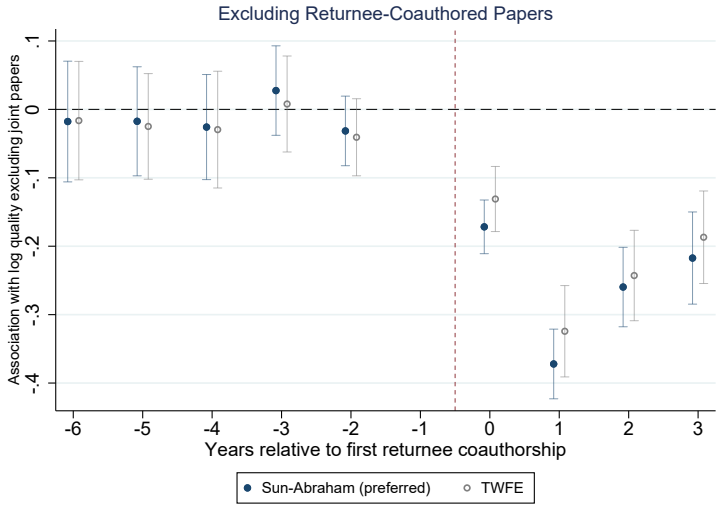
Excluding returnee-coauthored papers. We next ask whether the quality changes in Figure 7 extend beyond papers written jointly with returnees. We reconstruct cumulative Author Quality after assigning zero numerator contribution to every paper that contains an eligible returnee in or after that returnee's reported return year, while leaving the publication-year citation denominator unchanged. Figure A4 applies this outcome decomposition to both prior-network groups using their respective risk sets and controls.

In Panel A, the preferred Sun–Abraham estimate falls from 0.077 to −0.147 at event time 0 after excluding returnee-coauthored papers; TWFE is almost identical at −0.148. Sun–Abraham and TWFE are approximately zero at event time 1 and reach 0.081 and 0.075,

Panel A: Scholars with prior international collaboration



Panel B: Scholars without prior international collaboration



Notes: The outcome is logged cumulative Author Quality after setting the focal author's citation-share contribution to zero for papers containing an eligible returnee; the publication-year citation denominator is unchanged. Both panels report Sun–Abraham and TWFE estimates using the evolving-network controls and fixed effects from Figure 7. Panel A retains the prior-international risk set and uses 42,885 observations from 10,920 authors. Panel B retains the no-prior-international risk set and its censoring at first international collaboration, using 43,847 observations from 16,044 authors. The baseline Panel B contains 49,550 observations from 18,406 authors; the log exclusion outcome is undefined when cumulative quality from non-returnee-coauthored papers is zero. Sun–Abraham standard errors are clustered by author; TWFE standard errors are two-way clustered by author and institution. Exact event times from −6 through 3 are displayed, with −1 normalized to zero; bars show 95% confidence intervals. The displayed leads are jointly insignificant under both estimators: in Panel A, $p = 0.683$ and 0.582; in Panel B, $p = 0.592$ and 0.734 for Sun–Abraham and TWFE, respectively.

Figure A4
Author Quality Excluding Returnee-Coauthored Papers, by Prior International Connectedness

respectively, at event time 3. The immediate baseline increase is therefore concentrated in joint work. Although the preferred specifications show a later positive association in other work, the growth-adjusted pairwise comparisons below do not precisely establish a positive association at event times 1 through 3.

In Panel B, the preferred coefficients are -0.172, -0.372, -0.260, and -0.217 at event times 0 through 3; TWFE gives -0.131, -0.324, -0.243, and -0.187. The negative coefficients should not be interpreted as evidence that returnee collaboration harms other work. Treatment is dated by the first paper removed from the outcome, subsequent returnee-coauthored output remains excluded, and the log outcome drops author–years with no remaining cumulative quality. The sensitivity checks below show that later coefficients change substantially with the network adjustment and comparison construction. Thus, the immediate baseline increase is concentrated in joint work, while the later non-joint trajectory remains specification-dependent.

Sensitivity to network adjustment and pairwise comparisons. Figures A5 and A6 compare five specifications for the prior-international and no-prior-international risk sets, respectively. Each figure reports overall quality and quality excluding returnee-coauthored papers. The Sun–Abraham and TWFE series retain the preferred evolving-network specifications. Standard Callaway–Sant'Anna, labeled "CS (pre-event controls)," uses regression adjustment based on covariates in the earlier comparison period. For post-treatment comparisons, those values are measured in year $g-1$ for cohort g, so the lagged network stocks describe collaborators accumulated through $g-2$. Unlike the preferred fixed-effect regressions, standard CS does not update these controls to account for subsequent network growth.

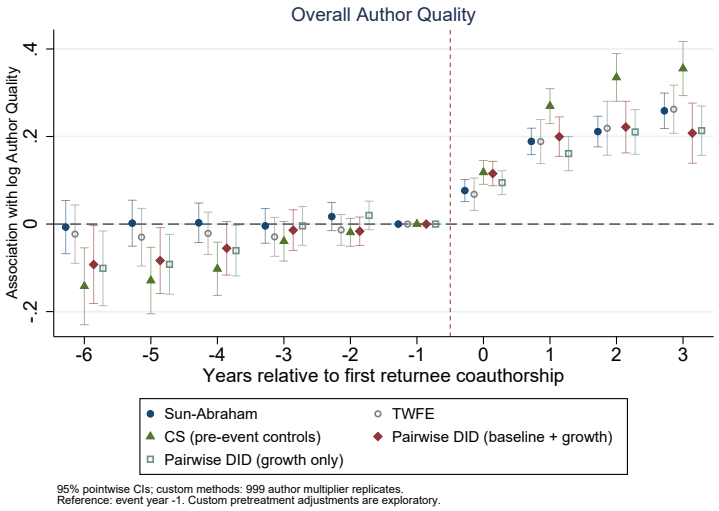
To assess this distinction, we construct two supplementary pairwise DID specifications. For each supported cohort–year cell, the outcome change is measured between $g-1$ and the endpoint year, and treated and never-treated comparison scholars must be observed in both years. "Pairwise DID (baseline + growth)" fits the outcome-change regression using never-treated scholars only, controlling for baseline network levels and changes in the lagged, IHS-transformed network stocks. It also includes baseline pre-policy network levels and changes in the corresponding baseline-by-year trends. The estimate averages treated scholars' outcome changes minus the fitted comparison-group changes. "Pairwise DID (growth only)" instead fits a pooled outcome-change regression on treatment status and changes in the evolving controls, equivalent to a two-period author fixed-effect specification. Both use the original risk-set-specific controls and CS cohort weights, normalized over supported cells. These are custom sensitivity specifications, not additional implementations of the standard CS estimator.

For internationally connected scholars, accounting for network growth brings the pairwise estimates closer to the preferred specifications. At event time 2, overall-quality estimates are 0.211 under Sun–Abraham, 0.335 under standard CS, 0.222 under baseline-plus-growth adjustment, and 0.210 under growth-only adjustment. The corresponding non-joint estimates are 0.032, 0.159, 0.045, and 0.039. The custom non-joint confidence intervals include zero at event times 1 through 3. For scholars without prior international collaboration, the overall-quality estimate at event time 2 changes from 0.371 under standard CS to 0.194 under baseline-plus-growth adjustment, compared with 0.144 under Sun–Abraham. The non-joint estimate changes from 0.141 to −0.038, compared with −0.260 under Sun–Abraham; the baseline-plus-growth intervals include zero at event times 1 and 2. Network adjustment therefore explains a substantial part of the numerical differences, but does not make the full-panel and pairwise comparisons equivalent.

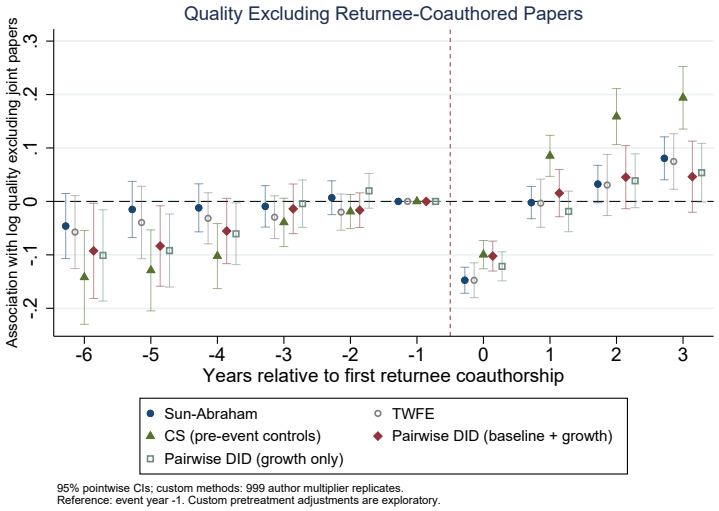
All sensitivity comparisons use $g-1$ as the outcome reference, including the displayed leads. This universal reference differs from the varying-reference CS leads in the common-

sample table. Some universal-reference lead estimates are negative and precisely estimated. The custom pretreatment adjustments use baseline network information measured at $g-1$ even for earlier endpoints and are exploratory rather than a validated parallel-trends test. Custom uncertainty is calculated jointly using 999 author-level multiplier-bootstrap replicates, refitting each cell and recomputing cohort weights while preserving shared authors across comparisons. The intervals are pointwise rather than simultaneous bands. Sparse pair support, differences in regression restrictions, and the possibility that collaboration itself changes subsequent network growth limit a causal interpretation. In particular, the adjusted estimates are not a mediation decomposition or proof that spillovers are absent.

Panel A: Overall Author Quality



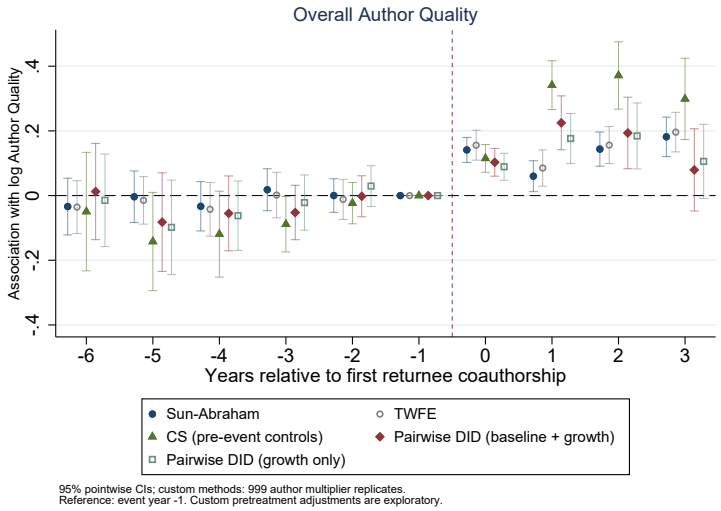
Panel B: Quality excluding returnee-coauthored papers



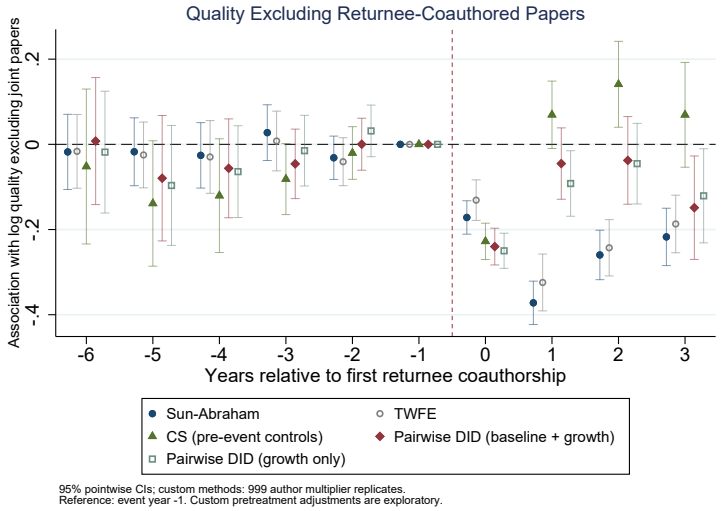
Notes: Both panels use scholars with prior international collaboration. Sun–Abraham and TWFE retain their original full-panel samples and evolving-network controls. Standard CS and the custom pairwise comparisons use never-treated controls and authors observed in both the reference and endpoint years; the 2004 cohort is excluded. The pairwise specifications differ in baseline-level adjustment and whether control slopes are fitted to comparison scholars only or to both groups. Original SA and TWFE intervals retain author and two-way author–institution clustering, respectively; standard CS uses analytical inference; the custom intervals use a joint author-level multiplier bootstrap. Event year −1 is normalized to zero. Bars show 95% pointwise confidence intervals. The custom methods are sensitivity diagnostics, and their pretreatment adjustments are exploratory.

Figure A5
Network-Adjustment Sensitivity among Internationally Connected Scholars

Panel A: Overall Author Quality



Panel B: Quality excluding returnee-coauthored papers



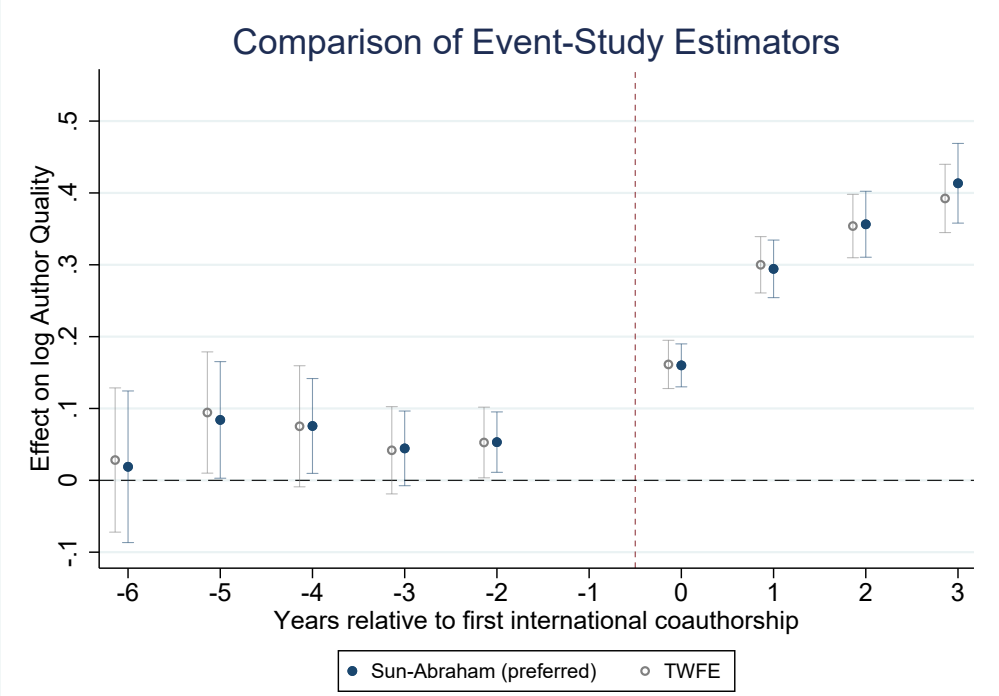
Notes: Both panels retain scholars only before first international collaboration. The evolving controls are the lagged total-coauthor stock and institution-level pre-policy network trend. Standard CS adjusts using earlier-period covariates; the custom methods additionally adjust for changes in those controls. Reference periods, comparison construction, and inference follow Figure A5. The pairwise analysis has 1,111, 320, 220, and 142 treated pairs at event times 0 through 3; this support differs from the full-panel SA/TWFE samples. Bars show 95% pointwise confidence intervals. Custom pretreatment adjustments are exploratory, and conditional non-joint estimates do not identify causal spillovers.

Figure A6
Network-Adjustment Sensitivity among Scholars Without Prior International Collaboration

International increment. Figure A7 estimates the international-collaboration increment for authors in

$$\mathcal{S}^{int} = \{j : T_j^{ret} < T_j^{int} < \infty\} \cup \{j : T_j^{ret} < \infty,\ T_j^{int} = \infty\}.$$

The first set contains authors whose returnee connection precedes their international collaboration; the second contains returnee-connected authors who never collaborate internationally during the sample period and serve as controls. Within this author set, eligibility is imposed observation by observation: only author-years at or after the first returnee connection enter the analysis. The specification replaces the returnee event indicators with event times relative to first international coauthorship and controls for the lagged cumulative stock of returnee coauthors. The resulting sample contains 26,328 observations from 8,716 authors. Sun–Abraham and TWFE are close in this design. The preferred estimates are 0.160, 0.294, 0.356, and 0.413 at event times 0 through 3. The five displayed leads are jointly insignificant at the 5% level, although the Sun–Abraham test is borderline ($p = 0.076$; TWFE $p = 0.180$). This path is descriptively larger than the returnee increment in Panel A of Figure 7, but the two designs use different risk sets, so their magnitudes are not directly comparable.



Notes: The figure compares Sun–Abraham (preferred) with conventional TWFE for first international collaboration among scholars already connected to returnees. The sample contains 26,328 author-year observations from 8,716 authors. Both specifications include author and year fixed effects, the lagged cumulative returnee-coauthor stock, and pre-policy international-collaboration-rate trends. Sun–Abraham standard errors are clustered by author; TWFE standard errors are clustered two-way by author and institution. Exact event times from −6 through 3 are displayed, with event time −1 normalized to zero; bars show 95% confidence intervals.

FIGURE A7
International-Collaboration Increment among Returnee-Connected Scholars

A.8 Censoring at First Direct Collaboration

The Section 4.2 baseline includes scholars with no prior returnee or non-local international coauthor as of year $t-1$. Consequently, an outcome observed in the first direct-collaboration year can remain in that sample. We define the first collaboration year at the author level across fields and retain an author–field–year only when its publication year strictly precedes both first-collaboration years. This rule does not select scholars according to whether they ever collaborate later; it censors their observations when direct collaboration starts. Table A7 compares the resulting estimates with the exact baseline specification. The stricter sample retains a strong first stage and a positive but imprecisely estimated 2SLS coefficient. The change in precision reflects both the removal of first-collaboration-year observations and the resulting loss of author–field histories that become singletons under the fixed effects.

Table A7
Institutional Exposure Before Any Direct Collaboration

	Baseline risk set	Strictly before collaboration
OLS coefficient	0.086	0.092
	(0.021)	(0.028)
First-stage coefficient	0.284	0.285
	(0.045)	(0.043)
Reduced-form coefficient	0.035	0.028
	(0.019)	(0.026)
Reduced-form p-value	0.070	0.288
2SLS coefficient	0.122	0.097
	(0.061)	(0.082)
2SLS p-value	0.047	0.241
Kleibergen–Paap F	40.54	42.95
KP underidentification LM	16.95	15.07
Observations	35,363	24,484
Authors	13,644	9,349
Institutions	194	185

Notes: The unit of observation is a local-scholar–field–year. The baseline risk set requires no returnee or non-local international coauthorship through year $t-1$. The stricter sample drops observations from the first year of either collaboration onward, including simultaneous first events. This removes 6,470 first-collaboration-year rows from the baseline sample; fixed-effect singleton drops reduce the resulting 28,893 candidate rows to 24,484 estimation observations. Both columns use the same log cumulative field-specific Author Quality outcome, lagged log-plus-one institution–field returnee stock, capacity-aligned instrument, controls, fixed effects, and two-way clustering by author and institution as Table 2. Standard errors are in parentheses. The instrument remains strongly relevant in the stricter sample, but its 2SLS coefficient is imprecise.

A.9 Institutional Exposure and Fractional Publication Output

Table A8 replaces log cumulative citation-share Author Quality with log cumulative fractional publications in the Section 4.2 specification. Each unique paper contributes $1/n_i$ to its author, where n_i is the paper's author count. We sum these credits within author–field–years and accumulate them from 2003 onward before applying the estimation-sample restrictions. The measure captures publication output rather than citation quality and is not divided by field–year publication totals. Panels A and B use exactly the baseline and strictly pre-collaboration samples in Table A7, retaining the same exposure, instrument, controls, fixed effects and clustering. Both exposure coefficients are positive; the strict-sample coefficient is more precise, but its conventional p-value is 0.078 and its Anderson–Rubin p-value is 0.102. These results are suggestive and do not establish that censoring selects weaker scholars or that the exposure coefficients differ between samples.

TABLE A8
Institutional Returnee Exposure and Fractional Publication Output

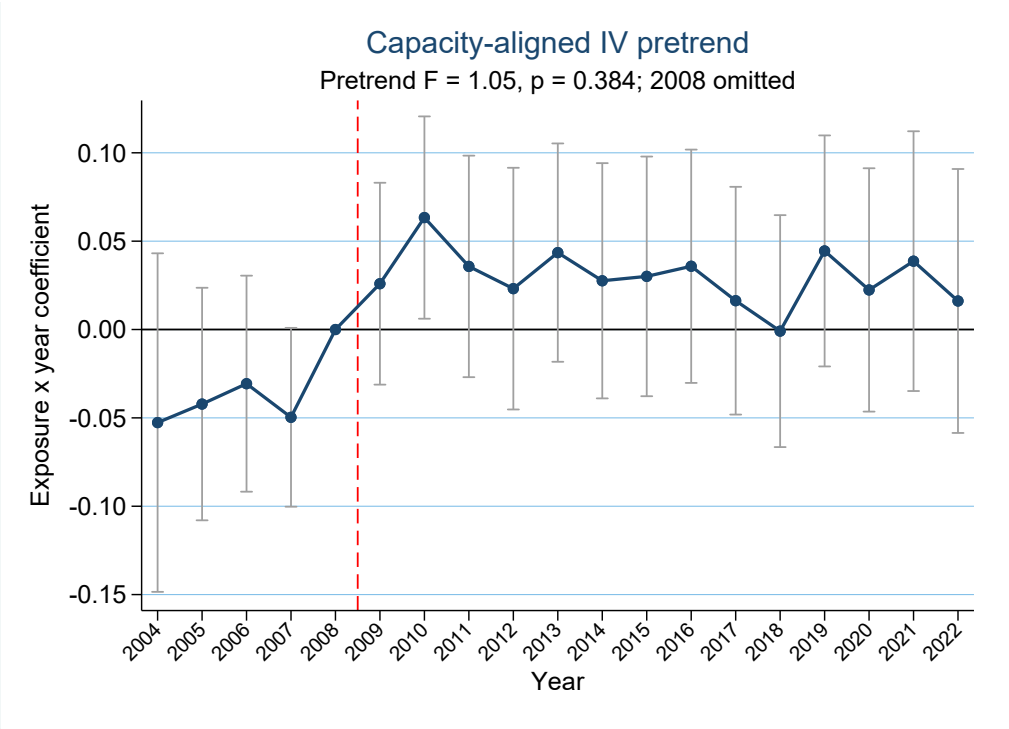
	(1) OLS	(2) First stage	(3) Reduced form	(4) 2SLS
Panel A: Baseline no-prior-collaboration sample				
$\ln(1+R_{kf,t-1})$	0.060***			0.060
	(0.014)			(0.045)
$Z^{cap}_{kf,t-1}$		0.284***	0.017	
		(0.045)	(0.013)	
First-stage / KP F		40.54		40.54
KP underidentification LM				16.95
Anderson–Rubin $F(\beta=0)$				1.64
Anderson–Rubin p-value				0.201
Observations	35,363	35,363	35,363	35,363
Scholars	13,644	13,644	13,644	13,644
Institutions	194	194	194	194
Panel B: Strictly pre-collaboration sample				
$\ln(1+R_{kf,t-1})$	0.065***			0.103*
	(0.018)			(0.058)
$Z^{cap}_{kf,t-1}$		0.285***	0.029	
		(0.043)	(0.018)	
First-stage / KP F		42.94		42.95
KP underidentification LM				15.07
Anderson–Rubin $F(\beta=0)$				2.69
Anderson–Rubin p-value				0.102
Observations	24,484	24,484	24,484	24,484
Scholars	9,349	9,349	9,349	9,349
Institutions	185	185	185	185
Author–field FE	Yes	Yes	Yes	Yes
Institution–field FE	Yes	Yes	Yes	Yes
Field–year FE	Yes	Yes	Yes	Yes
Pre-policy international trends	Yes	Yes	Yes	Yes
Regional economic controls	Yes	Yes	Yes	Yes

Notes: The unit is a local-scholar–field–year over 2005–2022. The outcome in Columns 1, 3 and 4 is log cumulative fractional publications: each unique paper contributes $1/n_i$, accumulated within author–field histories from 2003 onward. Column 2's outcome is lagged log-plus-one institution–field returnee stock. Panel A requires no returnee or non-local international collaboration through $t-1$ and can include first-collaboration-year outcomes. Panel B drops observations from the first year of either collaboration onward, including simultaneous first events, and removes fixed-effect singletons separately. Collaboration histories are defined across fields. Returnee coauthors are eligible U.S.-to-China returnees in or after their return year; international coauthors have a known assigned location outside China. Exposure, the capacity-aligned instrument, controls and fixed effects match Table 2. Standard errors, in parentheses, are clustered by scholar and institution. Conventional 2SLS p-values are 0.190 in Panel A and 0.078 in Panel B. Both KP underidentification tests have $p < 0.001$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.10 Shift-Share Exposure Diagnostics

This appendix reports three diagnostics for the capacity-aligned institution–field IV used in Section 4.2. Figure A8 tests whether higher-capacity institution–field cells were already on different quality trajectories before the post-2009 return wave. Table A9 reports a resid-

ualized field-component weighting decomposition, and Table A10 reports leave-one-field-out estimates. The earlier comparison across additive author and institution fixed effects is no longer informative for the preferred design, which instead absorbs author-by-field, institution-by-field, and field-by-year fixed effects throughout.



Notes: The figure plots coefficients on standardized log pre-policy institution–field capacity interacted with publication-year indicators; 2008 is omitted. The figure uses the recovered 2003–2022 publication histories of scholars who enter the corrected Section 4.2 sample and retains the same no-prior-collaboration author–field–year risk set. The regression absorbs author-by-field, institution-by-field, and field-by-year fixed effects; includes the baseline international-network trends and regional controls; and clusters standard errors two-way by author and institution. A joint test of the 2004–2007 interactions yields $F = 1.046$, $p = 0.384$. The separate pre-period linear trend is 0.0033 (s.e. 0.0115, $p = 0.774$), and a six-year-lead placebo is 0.0298 (s.e. 0.0990, $p = 0.764$).

FIGURE A8
Capacity-Aligned IV Pre-Trend Diagnostic

TABLE A9
Field-Component Weighting for the Capacity-Aligned IV

Field	Covariance weight	Component IV	Standard error	Weighted contribution
General	0.0498	-0.4057***	0.1335	-0.0202
Chemistry	0.2005	0.0575	0.1636	0.0115
Physics	0.7504	0.1631**	0.0687	0.1224
Life Sciences	-0.0008	-10.6230	31.3924	0.0083
Reconstructed pooled estimate	1.0000			0.1220

Notes: The conventional linear Rotemberg decomposition is not mechanically available because the capacity-aligned instrument applies the nonlinear transformation $\ln(1 + a_{kf} G^{-k}_{f,t-1})$ and each observation belongs to one field. This table instead reports a Rotemberg-style decomposition based on four mutually exclusive field components of the log instrument. Outcome, treatment, total-instrument, and component residuals absorb the preferred controls, author-by-field, institution-by-field, and field-by-year fixed effects. A field's weight is its share of the residualized first-stage covariance, and its weighted contribution is the weight multiplied by the component IV. The contributions reconstruct the current-IV-native pooled coefficient exactly. The Life Sciences component has effectively zero weight and is effectively unidentified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A10
Leave-One-Field-Out Diagnostics for the Capacity-Aligned IV

Estimation sample	Observations	KP F	2SLS coefficient	Standard error
All fields	35,363	40.54	0.1220**	0.0609
Exclude General	34,240	37.74	0.1494**	0.0632
Exclude Chemistry	20,458	34.55	0.1454**	0.0662
Exclude Physics	16,421	33.97	-0.0278	0.1211
Exclude Life Sciences	34,970	40.22	0.1158*	0.0609

Notes: Each diagnostic deletes observations belonging to the named field and re-estimates the preferred capacity-aligned specification on the current-IV-native sample. This is the natural leave-one-out exercise for the author–field panel and differs from subtracting a field component from an aggregate institution-level instrument. All models include the baseline regional controls and international-network trends, absorb author-by-field, institution-by-field, and field-by-year fixed effects, and cluster standard errors two-way by author and institution. Two-sided p-values are 0.047, 0.019, 0.029, 0.819, and 0.059 in row order. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.11 Institutional Exposure among Connected Scholars: Unadjusted Benchmark

Table A11 reports the connected-sample exposure specification without collaboration-history or intensity controls. It uses exactly the same observations as Panels A and B of Table 3. The 2SLS coefficient is 0.042 (s.e. 0.078, $p = 0.587$), and the KP statistic is 27.31. The exposure estimate is therefore already small and imprecise before conditioning on collaboration. The adjustment does not provide a mediation decomposition or establish that indirect spillovers are absent.

TABLE A11
Institutional Exposure among Connected Scholars: Unadjusted Benchmark

	(1) OLS	(2) First stage	(3) Reduced form	(4) 2SLS
Panel A: Without collaboration-history or intensity controls				
$\ln(1 + R_{kf,t-1})$	0.032*			0.042
	(0.018)			(0.078)
$Z^{cap}_{kf,t-1}$		0.251***	0.011	
		(0.048)	(0.018)	
First-stage / KP F		27.31		27.31
KP underidentification LM				15.78
Anderson–Rubin $F(\beta = 0)$				0.34
Observations	47,147	47,147	47,147	47,147
Scholars	11,485	11,485	11,485	11,485
Institutions	198	198	198	198
Author–field FE	Yes	Yes	Yes	Yes
Institution–field FE	Yes	Yes	Yes	Yes
Field–year FE	Yes	Yes	Yes	Yes
Pre-policy international trends	Yes	Yes	Yes	Yes
Regional economic controls	Yes	Yes	Yes	Yes

Notes: The sample is identical to Table 3, but collaboration-history and intensity controls are omitted. The unit is a local-scholar–field–year, and the outcome is log cumulative field-specific Author Quality. Exposure is $X_{kf,t-1} = \ln(1+R_{kf,t-1})$, instrumented by $Z^{cap}_{kf,t-1}$ from Equation (5). Qualifying returnee coauthors are U.S.-to-China returnees in or after their return year; international coauthors have a known assigned location outside China. Histories are defined across fields. Controls include author- and institution-level 2003–2005 international-collaboration rates interacted with publication year and lagged provincial log GDP per capita and education expenditure. Standard errors are clustered by scholar and institution. The KP underidentification test has $p < 0.001$; the Anderson–Rubin test has $p = 0.561$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.12 Formal Comparison by Prior Connectedness

Let $D_{j,t-1} = 1$ when scholar j had returnee or non-local international collaboration before year t, and let $I_s = \mathbf{1}[D_{j,t-1} = s]$ for $s \in \{0, 1\}$. We estimate both exposure coefficients

jointly:

$$\log AQ_{jft} = \sum_{s=0}^{1} I_s \left[\beta_s X_{kf,t-1} + \mathbf{W}'_{jkt}\boldsymbol{\gamma}_s + \mu_{jf,s} + \eta_{kf,s} + \lambda_{ft,s}\right] + \varepsilon_{jkft}. \tag{11}$$

Both $X(1-D)$ and XD are endogenous and are instrumented by $Z^{cap}(1-D)$ and $Z^{cap}D$. This is equivalent to including X and the exposure-by-connectedness interaction XD, whose coefficient is $\delta = \beta_1 - \beta_0$. All controls and fixed-effect dimensions have separate coefficients by state; the standalone status indicator is absorbed. We test $\delta = 0$ with two-way clustering by scholar and institution across both states, allowing the same scholar to contribute observations to both groups. A second specification adds the connected-sample history terms in Equation (8) and the annual-intensity controls described for Panel B; these terms are zero in the unconnected group.

The jointly estimated group coefficients reproduce the corrected separate regressions. Their estimation sample is the union of the two separately pruned risk sets: 35,363 unconnected and 47,147 connected observations. Without collaboration controls, the connected-minus-unconnected difference is -0.080 (s.e. 0.101, $p = 0.429$). With connected collaboration controls it is -0.088 (s.e. 0.099, $p = 0.376$). Neither test establishes a difference; the confidence intervals are too wide to establish equivalence either.

For sensitivity, we retain common author–field, institution–field and field–year effects, include D explicitly, and continue to allow the other control slopes to differ. On the larger pooled support this specification yields a positive difference of 0.024 ($p = 0.028$). On exactly the same observations as the separate-regression union, it yields -0.043 ($p = 0.124$). The significant pooled-support finding is consequently sensitive to the fixed-effect restrictions and sample support. It does not provide robust evidence that one group's exposure coefficient is larger.

The interacted instruments predict their corresponding exposure variables. The joint KP statistics are 13.65 and 14.13 for the separate-effects specifications. The separate-group first-stage KP statistics are 40.54 for unconnected scholars and 27.31 for connected scholars (28.25 with collaboration controls). Conditional Sanderson–Windmeijer statistics are not reported by the software in the separate-effects design because opposite-group instruments are exactly omitted from each group first stage. The reported coefficient-difference tests use conventional clustered Wald inference. Prior connectedness may itself respond to earlier institutional exposure; lagging it does not make it exogenous, and validity of the interacted instrument requires additional conditional exclusion assumptions. We therefore interpret these as conditional coefficient comparisons.

TABLE A12
Institutional Exposure: Formal Comparison by Prior Connectedness

Specification	Unconnected	Connected	Difference	p-value	Observations
Separate FE, no collaboration controls	0.122	0.042	-0.080	0.429	82,510
	(0.061)	(0.078)	(0.101)		
Separate FE, collaboration controls	0.122	0.034	-0.088	0.376	82,510
	(0.061)	(0.076)	(0.099)		
Shared FE, pooled support	0.028	0.052	0.024	0.028	111,039
	(0.054)	(0.060)	(0.011)		
Shared FE, same union sample	0.052	0.009	-0.043	0.124	82,510
	(0.043)	(0.057)	(0.028)		

Notes: The difference is connected minus unconnected and is the coefficient on exposure interacted with prior connected status under an equivalent parameterization. Both group exposure terms are instrumented. All models use corrected collaboration histories and baseline network controls and group-specific slopes for the other controls. The first two rows allow author–field, institution–field and field–year effects to differ by status, reproducing the separate regressions. Collaboration controls in row 2 are included for connected scholars only. The last two rows impose shared fixed effects and include the status indicator. Standard errors are two-way clustered by scholar and institution; p-values test equal exposure coefficients. The 95% difference intervals in rows 1 and 2 are $[-0.278, 0.119]$ and $[-0.282, 0.107]$. Joint KP statistics in row order are 13.65, 14.13, 23.37 and 19.66. These are conditional comparisons, not automatically causal heterogeneity estimates.