

# Race to the Big Lab: Gender Disparities in Large Team Collaboration and Its Impact on Early Academic Careers

Zeyuan Chen

Department of Computer Science  
Aalto University  
Espoo, Finland

Department of Informatics and Media  
Uppsala University  
Uppsala, Sweden  
ext-ollie.chen@aalto.fi

Anh Duong Nguyen

Department of Computer Science,  
School of Science  
Aalto University  
Espoo, Finland  
anh.18.nguyen@aalto.fi

Barbara Keller

Computer Science  
Aalto University  
Espoo, Finland  
barbara.keller@aalto.fi

## Abstract

This study investigates the role of large-team collaboration in shaping early-career scholars' career development, with a focus on gender disparities. Using publication and collaboration data from SciSciNet in Computer Science, we capture the social capital accumulation process in academia with a neighborhood-based centrality metric and publication counts. Synthetic difference-in-differences (SDID) is applied to estimate the impact of early experience in large-team collaboration on subsequent research careers. Results indicate that junior scholars participating in large-team research significantly improve their network centrality, indicating more frequent collaborations with influential scholars, and produce approximately 0.75 more publications per year. Meanwhile, we document persistent gender gaps: men are 16% more likely to access large-team collaborations. These findings highlight large-team collaboration as both a source of career acceleration and a mechanism of gender inequality. We conclude with implications for equity promotion and strategies enabling more inclusive collaboration.

## CCS Concepts

• **Human-centered computing** → **Empirical studies in collaborative and social computing**; • **Social and professional topics** → *Gender*; *Computer supported cooperative work*.

## Keywords

Team Science, Large Team Collaboration, Social Capital, Early-Career Scholars, Gender Disparities, Scientific Capital, Collaboration Network, Co-Author Network, Network Analysis, Causal Inference, Scientometrics, Network Centrality, Synthetic Difference-in-Differences, Survival Analysis

## ACM Reference Format:

Zeyuan Chen, Anh Duong Nguyen, and Barbara Keller. 2026. Race to the Big Lab: Gender Disparities in Large Team Collaboration and Its Impact on Early Academic Careers. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*, April 13–17, 2026, Barcelona, Spain. ACM, New York, NY, USA, 20 pages. <https://doi.org/10.1145/3772318.3791205>



This work is licensed under a Creative Commons Attribution 4.0 International License. *CHI '26, Barcelona, Spain*

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2278-3/26/04

<https://doi.org/10.1145/3772318.3791205>

## 1 Introduction

In contemporary scientific research, collaboration in large teams has become increasingly prevalent, driving breakthroughs and shaping the trajectory of entire disciplines[1]. Large-team collaborations often entail access to substantial resources, such as extensive funding sources particularly in labor-intensive or computation-heavy fields (e.g., Computer Science), thus providing individual scholars with unique opportunities of career development. However, less attention has been paid to the consequences of large-team collaboration for the individual scholars involved, particularly those at the early stages of their careers. For early-career researchers, access to large-team collaborations may represent a pivotal opportunity for professional development, visibility, and the accumulation of social capital, and consequently bring long-term positive effects for their career development[2].

On the other hand, opportunities for large-team collaborations are not distributed equally among all individual scholars. Like other competitive resources in academia, participation in large, high-impact teams is often shaped by underlying social dynamics, resulting in disparities in who is able to benefit most from collaborative science[3]. Gender, in particular, has long been recognized as a source of inequality in academic career development, as evidenced by women's underrepresentation in senior or high-impact positions and their disadvantages in publication quantity and scholarly impact[4]. However, the role of team size as a competitive resource remains underexplored. If access to large-team collaborations is not merely a matter of individual preference but instead represents a form of scarce academic resources, then unequal access to such opportunities could be an important, yet overlooked, mechanism contributing to gender disparities in scholarly career development.

Previous studies have characterized the development of academic career using individual metrics such as publication count, citation count, H-index and possibility of receiving mentorship / professorship[5, 6]. While emphasizing direct research productivity, they overlook another implicit attribute: the position of an individual within the academic community, which may reflect their advantages or limitations in accessing information and mobilizing resources[7]. Network data can help reveal positional attributes within a given circle. As a theoretical approach to measuring the positional advantage of nodes in a network, Bonacich proposed a centrality measure that expresses an individual's power or influence as a function of the status of their neighbors — the stronger the connections to influential others, the greater one's own power

or influence[8]. While this definition is insightful, its application to collaboration networks is computationally intensive and may introduce ambiguities in distinguishing between direct and more distant connections. In other words, a scholar's centrality in a collaboration network is generally assumed to depend more heavily on direct coauthors, with indirect connections contributing less. Therefore, we draw on the weighted neighborhood centrality proposed by Liu et al.[9], which is computed based on the local network structure of each node, and has been shown to more accurately identify high-impact nodes in real-world spreading networks, to define centrality in academic collaboration networks. To this end, we construct collaboration networks based on Newman's[10, 11], using records from the Computer Science domain in SciSciNet[12] as our data source. By combining productivity and centrality, we provide a comprehensive characterization of early-career scholars' career development, allowing a thorough assessment of the sustained impact of large-team collaboration on individual researchers.

Our contributions are as follows. First, we provide empirical evidence on how joining a large team for the first time can shape a scholar's subsequent academic advancement, by incorporating network-based measure into advanced causal inference techniques to characterize a scholar's position and influence. Second, we investigate the role of gender in shaping access to large-team collaboration, thus proposing a new mechanism of gender disparities and offering a more nuanced understanding of academic career dynamics. Specifically, we investigate the following research questions (RQs):

- RQ1: Does large-team collaboration have a positive impact on the subsequent research productivity of early-career scholars?
- RQ2: Does large-team collaboration enhance the centrality of early-career scholars within academic collaboration networks?
- RQ3: Does gender influence the likelihood of early-career scholars joining large-team collaboration?

By leveraging large-scale collaboration data and network analytics, our work contributes to ongoing efforts to use computational and empirical methods to uncover patterns in human behavior and organizational dynamics[13]. Besides, our findings offer actionable insights for fostering more inclusive and sustainable academic communities, in research fields that emphasize collaborative teamwork such as HCI, resonating with the vision of designing for diversity and supporting the growth of equitable, technology-enabled communities[14].

## 2 Related Work

In this section, we review trends in team science and research collaboration, with a focus on their implications for early-career academic development. Drawing on social capital theory [15] and social network methodologies [9, 11], we highlight key knowledge gaps—particularly in understanding the mechanisms underlying unequal access to social capital within academic networks and in the operationalization of collaboration structures. These gaps will be addressed in our subsequent analysis.

### 2.1 Team Science and Early-Stage Academic Career

The size of research groups [2] and teams [16, 17] has grown rapidly across various fields since the last century. Team science allows researchers to collectively tackle more complex problems [18], offering epistemic advantages [19] through access to advanced technologies and highly specialized expertise [20] with larger funding [17]. In terms of scientific impact, collaborative work has become increasingly dominant among highly cited studies—a trend not fully attributable to self-citation effects [16]. Publications with larger teams are more likely to gain visibility, as team members promote the work within their own networks, thus compensating for the diluted credit per co-author [21]. However, team science also presents several organizational challenges in sustaining high-quality research. Collaboration requires extensive and effective coordination among members, including alignment in goals, workflow integration, incentive structures, and socio-technical infrastructure [22]. In a longitudinal study of collaboration persistence in Computer Science, Bu et al. [23] found that an initial team size of 4–5 members can be advantageous, but team performance often declines over time due to pressures and the erosion of meaningful interpersonal interaction.

Scientific collaboration has been recognized as playing a critical role in developing key career assets for junior researchers. Prior studies have primarily examined this through the lenses of human capital and social capital models, highlighting how collaboration fosters connections with influential academic figures beyond the acquisition of knowledge and research skills [24]. Interviews with scientists about their collaborative practices reveal that they strategically navigate expertise gaps and co-development focus within teams [20]—in other words, they actively mobilize academic resources to support their career progression. Empirical research also argues that participation in prestigious research groups helps to accumulate symbolic capital [25], which, as pointed out by Pierre Bourdieu [26], means a set of non-material resources including access to social networks and intellectual competencies. For early career academics, that means greater competitiveness in the academic labour market [27]. Additionally, several studies have shifted attention toward measurable outcomes such as increased publication output and visibility linked to collaboration and mentorship opportunities [2, 21]. These perspectives align closely with Lin's (2001) social capital framework [15], which conceptualizes social capital as values embedded in social ties that can be mobilized by actors in status attainment processes. Building upon this theoretical foundation, we posit that large-scale collaborations constitute a mode of mobilizing material and non-material academic resources with the aim to understand their implications for early-stage research careers related to the formation of professional networks and the attainment of career outcomes.

On the other hand, empirical research on team science and academic career development reveals a complex interplay between the scale of collaboration, modes of engagement, and individual characteristics. Empirical evidence suggests that while participation in larger research teams can enhance mentees' productivity and visibility, limited mentorship within such environments may undermine their long-term commitment to academic careers [2]. In their

analysis of academic career trajectories, Petersen and colleagues highlight the influence of collaboration radius and team efficiency on scientific productivity; notably, they find that low team efficiency can lead to diminishing marginal returns from additional co-author input [21]. Overall, prior work has largely inferred the benefits of collaboration from long-term achievements or proportional productivity gains per collaborator, often neglecting the mechanisms through which junior researchers mobilize the resources embedded in their collaborative ties. Our study addresses this gap by demonstrating how team science facilitates one's structural position or potential access to socially embedded resources, thereby supporting early-career advancement in academic settings.

## 2.2 Scientometrics

A substantial body of literature on scientific collaboration and output adopts a scientometric approach, drawing on bibliographic data. Indicators such as the number of co-authors or group members, authorship order, and frequency of collaboration are commonly used as proxies for collaboration dynamics [2, 23, 28–30], while productivity is typically measured through metrics like publication count, citation count, and H-index [2, 17, 31]. To interpret these indicators, researchers apply a combination of statistical inference [2, 30] and theoretical modeling techniques [18, 21], drawing from both econometric and computer science traditions. Among various methodological designs, social network analysis has emerged as a well-established and powerful framework to capture the conceptual structure of research fields [32] as well as collaboration [33] and citation [34] patterns. In the context of team collaborations, simulation models of team assembly structures highlight the advantage of forming teams with diverse and experienced members, especially in accessing knowledge from large, central, and well-connected researcher clusters [18]. Zhu et al. [33] further explore the dynamics among researchers, their collaborators, and second-degree connections. Their analysis offers compelling evidence of how the size of first- and second-degree collaboration neighborhoods, in conjunction with the focal author's disciplinary breadth, influences productivity and impact.

Although prior research on team science and career development offers broadly consistent insights, we identify a key methodological gap in this domain. Collaboration structures have been operationalized in diverse—and often incomparable—ways. While many studies focus on the number of collaborators or team members, this indicator is subject to varying aggregation schemes: annually [21] or within early-career time windows [2] for individual researchers, or as static snapshots of collaboration neighborhoods [33]. Moreover, narrow definitions of collaboration structures yield inconsistent assumptions in measuring collaboration gains. For example, the proportional publication opportunities per collaborator [21] assume uniform importance for all collaborators, while the number of joint publications with supervisors [2] weights mentorship quality over other collaborations. This fragmented methodological landscape underscores the need for a more flexible and systematic approach to operationalizing key collaboration structures and — one that accommodates different aggregation levels and enables fine-grained analysis of focal and peripheral collaboration ties.

## 2.3 Gender Disparities

Gender has long been recognized as a fundamental axis of inequality within academic career trajectories, with numerous studies highlighting severe under-representation of women in academic science. Two widely cited phenomena—the “leaky pipeline” [4, 35] and the “glass ceiling” [36, 37]—aptly capture this dynamic. The former describes the gradual retention of women the trajectories in academic science, while the latter refers to the invisible yet persistent barriers that prevent women from attaining positions of influence, prestige and leadership. Existing literature has identified several mechanisms contributing to these gendered outcomes. Gender stereotypes discourage girls from pursuing math-intensive majors [35]. Women researchers with children often face the so-called “motherhood penalty,” resulting in reduced productivity and a higher likelihood of exiting academia after completing their PhDs—trends that are further reinforced by self-selection [5, 35]. Women consistently receive less recognition for their research outputs compared to their men colleagues when competing for funding and academic prestige [38, 39], due in part to mechanisms such as gender role congruity [40] and homophily [41].

Several studies highlight the impact of gender disparities in mobilizing social capital through professional networks and team research [4, 24]. Women receive less interest in collaboration from colleagues—even on gender-neutral topics [40] and tend to be marginalized in collaborative networks with respect to their betweenness centrality [42] and propensity to collaborate [43], particularly in international projects that tend to yield higher publication counts and greater impact [44]. Simultaneously, gender disparities can emerge through discriminatory team processes and dynamics. Women frequently encounter subtle forms of exclusion in networking, mentorship, and promotion opportunities, often resulting from homophilic tendencies among men colleagues [45]. Scientific teams may also exhibit gendered divisions of labor that reinforce systemic inequities [46]: women researchers are often assigned or take on less visible and less valued tasks [47]. Consequently, their expertise tends to be underrecognized especially in men-dominated teams [48], and they receive less credit on scientific publications and patents compared to men [47]. Unsurprisingly, women faculties cite dissatisfaction with the work climate as a reason for leaving academia more often than their men colleagues [49]. Overall, existing evidence suggests gender disparities in collaboration and team science as a potential, yet often overlooked, enabler of the glass-ceiling and leaky pipeline phenomena. If access to such collaborations is not purely a function of merit or individual choice but instead shaped by implicit bias, institutional gatekeeping, or network effects, then participation in large-scale projects becomes a scarce and stratified resource. Distinct from prior research that has focused on the collective impact of collaboration or gender-balanced teams, this study seeks to illuminate the long-term consequences of collaborative opportunity gaps on academic career development, contributing to a deeper understanding of the root causes of gender disparities in academic science.

### 3 Data

#### 3.1 Data Source

This paper draws on the SciSciNet dataset[12], which not only provides large-scale coverage of publications and collaborations in various fields, but also contains the necessary author-level information, with records updated through December 2021. Focusing on Computer Science, SciSciNet captures 8.6 million unique authors and 7.3 million publications, offering a rich ground for exploring collaboration patterns in this fast-evolving field. As summarized in the Appendix A, Publications in Computer Science typically have a relatively large number of coauthors, indicating that collaboration is common in this field.

SciSciNet provides author-gender associations based on a first-name gender classifier, *nomquamgender* [50]. The classifier adopts mechanisms for cultural inclusivity at several stages, trained on an open-source curated multilingual dataset and accepts non-Roman inputs, with an inference process that includes diverse sources and reduces potential reliance on a few dominant ones. However, SciSciNet’s gender classification only considers the first part of the first name, resulting in a large proportion of authors being categorized as “unknown.” To improve the classification rate, we made use of an enhanced gender inference by considering all parts of an author’s first name, which reduced the number of “unknown” authors by over 2 millions (raising the detection rate from 61.12% to 67.81%) as introduced by [51]. For the sake of transparency, we report the proportion of individuals classified as “unknown” within the studied cohorts, as well as the probability distribution of probability of gendered female, in the Appendix B. While this study treats gender as a binary variable due to current technical constraints, we view gender not as an inherent dichotomy but as a spectrum of diverse identities. Our use of a name-based gender classifier does not attempt to determine anyone’s gender identity; rather, it captures how names—socially gendered markers of identity—can influence scholars’ collaboration patterns and career trajectories. Importantly, this approach acknowledges that academic institutions often adopt binary notions of gender [52, 53], within which women scholars may face systemic disadvantages [54]. Accordingly, when we categorize our sample into women and men groups in the analysis below, this should be understood as a methodological strategy rather than a claim about individuals’ gender identities.

We restrict our analysis to publications from the year 2000 onward, to capture contemporary collaboration practices and ensure adequate representation of women researchers, which is consistent with previous studies [30, 51]. Additionally, we focus only on journal and conference papers, excluding other types of entries such as datasets, books, and repositories.

#### 3.2 Sample Inclusion Criteria

Our goal is to construct a series of panel datasets of *early-career researchers* in Computer Science for statistical inference. Therefore, we identify authors who had never published any other publication before their first publication in the field of Computer Science, labeling them as fresh authors in their debut publication year. Readers should note that some individuals may have prior publications in other fields or research experience not recorded in SciSciNet. Also, this is only one of several possible ways to define the start of a

junior scholar’s career. Only authors who published more than two papers in Computer Science during the first ten years of their records were included.

Second, we exclude authors who stopped publishing within ten years of their debut publication, treating them as having exited academia. This choice follows prior work [2]. Exiting academia can reflect very different realities—for some, it may be a positive career move, while for others it may be a setback[55, 56]. Regardless of its meaning, however, an exit sets outcome indicators such as degree growth or publication count to zero. If early collaboration patterns (our independent variable) influence the likelihood of exit, then including such cases could bias our estimates. This makes our approach distinct from studies that examine how collaboration affects career duration[2, 57].

Finally, we restrict our sample to men and women fresh authors who entered academia between 2000 and 2004 (five cohorts) to focus on their development during the first ten years of their careers (Table 1). We do not include later cohorts because identifying whether an author quits academia within ten years requires an adequate post-observation window. For the latest eligible cohort (2004), our dataset provides 7 years of post-2014 publication observations (2014–2020), which we consider the minimum necessary for reliably determining career discontinuation. In addition, citation counts require sufficiently long post-publication windows to stabilize, and earlier cohorts therefore allow us to obtain more robust citation-based measures.

### 4 Methodology

This section outlines and justifies our empirical strategy. Section 4.1 details the construction of the collaboration networks. Sections 4.2 and 4.3 define and operationalize the key variables. Sections 4.4 (RQ1 and RQ2) and 4.5 (RQ3) present the identification strategy and statistical analyses used to address our research questions. Section 4.6 provides a positionality statement.

#### 4.1 Collaboration Network Construction

For each of the five cohorts from 2000 to 2004, we construct collaboration networks with a time span of ten years. We define a collaboration network from year  $Y_i$  to  $Y_j$  as a undirected multi-graph  $G = (N, E)$ , where  $N$  is the set of all authors publishing papers during this period, and  $E$  is the set of edges indicating coauthoring relationships between authors[10].

We then introduce edge weights into the collaboration network: each publication increases the edge weight between two authors by the reciprocal of the team size minus one[11]. Formally, when two authors  $u$  and  $v$  co-author a publication  $p_i$  within a team of size  $S(p_i)$  collaborators, an edge weight of  $\frac{1}{S(p_i)-1}$  is added to their collaboration tie  $(u, v)$ . The cumulative weight of the tie  $(u, v)$  is obtained by iterating this operation over all such publications  $p_i$  that both  $u$  and  $v$  have co-authored. Informatively, this definition assumes that collaboration ties are stronger in smaller teams than in larger teams. Furthermore, this approach ensures that each publication  $p_1$ , adds the same weight, namely 1, to the network, acting as a normalization that controls for team size when calculating centrality. Given that our dependent variable is highly correlated with team size, this normalization is necessary.

**Table 1: Fresh authors sample composition by cohort year between 2000-2004. Women representation is estimated at roughly 20.2%, which grows in numbers and proportionally over the years.**

Cohort Year	Total	Female-named Scholars	Female-name Percentage
2000	11,567	2,166	18.7%
2001	12,500	2,410	19.3%
2002	14,259	2,873	20.1%
2003	15,891	3,246	20.4%
2004	17,896	3,883	21.7%
<b>Total</b>	<b>72,113</b>	<b>14,578</b>	<b>20.2%</b>

## 4.2 Centrality and Productivity

We capture aspects of early research career development with relation to Lin’s social capital framework. They are researchers’ *centrality* in the collaboration network (related to the structural position element of social capital) and *productivity* or research output that are often facilitated by their collaboration ties (related to the returns upon investing in social ties).

In existing studies of academic collaboration networks, one of the most popular and straightforward metrics for assessing node centrality is the degree, defined as the count of unique collaborators connected to the focal author[30, 58]. However, this simple metric overlooks two crucial aspects of social capital and social embeddedness:

- (1) *tie strength*: collaborating with another author in a two-person team is fundamentally different from collaborating in a 100-person team in terms of mutual attention and the closeness of the relationship. Moreover, the frequency of collaborations should also correlate with the strength of the tie.
- (2) *neighbors’ influence*: given the same tie strength, connections with influential collaborators should yield greater social capital for the focal author.

Eigenvector centrality [8] is a well-established centrality measure applicable to both unweighted and weighted graphs. It accounts for the influence of a node’s neighbors by incorporating their centralities into the calculation. However, this measure has significant limitations when applied to a researcher’s structural position within a collaboration network. Because it reflects the global structure of the network, a researcher’s eigenvector centrality can be affected by distant colleagues, making it a less meaningful proxy for social capital in academic collaborations—where connections beyond a 2-hop neighborhood rarely yield substantial advantages. In addition, in academic networks, we expect values embedded in collaboration ties and the influence of neighboring nodes to exhibit a monotonic relationship, while global measures such as eigenvector centrality sometimes fail to capture this property[59]. Moreover, eigenvector centrality requires global convergence, making it computationally expensive for large-scale networks with millions of nodes. For these reasons, we prefer a metric that can be computed within ego subnetworks of small diameter, offering more intuitive insights and significantly lower computational demands.

To address these considerations, we constructed an adaption of neighborhood centrality[60], which considers both the tie strength

and the influence of 1-hop neighbors. Specifically, we first compute the weighted degree of each node as:

$$WD_i = \sum_{j \in N(i)} e_{i,j},$$

where  $N(i)$  denotes the set of neighbors of author  $i$ , and  $e_{i,j}$  represents the edge weight between authors  $i$  and  $j$ . Note that under this definition,  $WD_i$  essentially equals the total number of collaborative publications by author  $i$ . We calculate centrality as follows:

$$C_i = \sum_{j \in N(i)} e_{i,j} WD_j = \sum_{j \in N(i)} e_{i,j} \left( \sum_{k \in N(j)} e_{j,k} \right)$$

In brief,  $C_i$  denotes the weighted sum of collaborations that neighbors of a focal author have, adjusted by the strengths of their respective ties. The value of  $C_i$  increases when the focal author collaborates with more influential partners and has stronger ties with them. Conversely, in the absence of powerful collaborators or strong connections, the focal author’s  $C_i$  is lower, reflecting a marginalized position in the network. This metric captures the structural advantages available to a researcher: those who collaborate with well-connected authors, particularly through strong or frequent relationships, can gain greater visibility and resources.

For productivity, we first measure the yearly number of publications by each author, including both collaborative publications, represented by  $WD$  in the network, and solo publications. In addition, as a supplementary measure, we include a citation-weighted productivity indicator, defined as the annual sum of the Relative Citation Indicator (RCI), where the RCI of a publication is its citation count divided by the average citation count of papers in the same field and year. This normalization enables comparisons across fields and time[12, 61]. We report both measures when assessing the impact of large-team collaboration on productivity, as they respectively capture output volume and research influence, offering a more comprehensive view of scholarly productivity.

In addition, we employ two additional indicators of productivity: (a) a venue-prestige-weighted yearly publication count, where venue prestige is proxied by the importance probability provided in MAG’s ranking system. MAG defines the Rank of journals and conference series as  $-1000 \times \ln(\text{probability of being important})$ [62], and we use this underlying probability as the prestige measure. Because the procedure used by MAG to compute this probability is not publicly documented, we treat this indicator solely as

a robustness check for research output influence or quality. Nevertheless, we report in the appendix information on the 10 most prestigious journals and conferences in computer science, along with the distribution of venue prestige, all of which align with intuitive expectations; (b) the yearly average relative citation count, which isolates citation impact by removing the effect of publication volume. This measure is likewise reported as part of our robustness analysis, as our primary focus is on overall research output rather than mere citation-based outcomes, which are sensitive to factors such as self-citation loops typical of large team collaborations[16].

### 4.3 First-time Large-team Collaboration

First-time authorship in a large-team publication is selected to further operationalize young researchers' initial success event of mobilizing collaboration resources, with several technical considerations as follows. We define the event year as the first year in which an author publishes in a large team, where a large team is defined as having a team size greater than the 95% percentile of all team sizes in that year. In the robustness tests, we replace this event indicator with various alternative definitions. However, it is important to note that the number of coauthors does not equate to the total number of individuals who contributed to the publication, nor does it necessarily represent all the researchers with whom the focal author had direct contact during that collaboration. Despite these limitations, team size or the number of coauthors remains a widely used proxy for the scale of collaboration[1, 17].

Another potential limitation of this operationalization is that it does not distinguish whether the focal author became part of a large, stable research group or merely published once within a large team (for example, as a short-term research assistant). However, we do not intend to differentiate between these two scenarios. In the former, large-team collaboration may bring stable and long-term benefits for the author over several years. In the latter, the focal author might still gain potential long-term returns, such as connections with potential future collaborators, valuable collaboration experience to list on their CV, or publications in high-impact venues. Our goal is only to examine whether such first-time large-team collaboration yields significant benefits for early-career researchers in the future.

### 4.4 Synthetic Difference-in-Differences

For our first two research questions, we aim to estimate the causal effect of entering a large-team collaboration for the first time on researchers' network centrality and academic productivity. A key challenge here is *selection bias*: researchers who already have higher centrality or productivity are more likely to gain access to large-team opportunities. These researchers may already excel in many dimensions that are difficult or impossible to observe, such as research ability and social connectedness. Even without joining a large team, these individuals would likely outperform others, meaning that estimates from a traditional regression model would be confounded.

One common approach to mitigate this bias is the Difference-in-Differences (DID) method, which compares changes in outcomes between treated and control groups before and after treatment - in our study, the treated groups consists of scholars who have

participated in large-team collaborations, while the control group includes those who have never engaged in such collaborations. However, DID relies on the parallel trends assumption—the idea that treated and control groups would have followed similar trajectories in the absence of treatment[63]. In our context, this assumption is violated due to potential confounders mentioned above: as shown in Figure 9, authors who eventually join large teams already exhibit higher productivity and centrality well before treatment. This pre-existing gap makes DID unsuitable, as it would attribute part of the selection effect to the treatment itself.

To address the limitations of conventional DID, we employ the Synthetic Difference-in-Differences (SDID) method. Instead of assuming that treated and control authors would naturally follow similar trends, SDID constructs a synthetic control group that closely matches the pre-treatment trajectory of the treated group, and potentially other unobservable confounders [64]. In other words, the synthetic control balances treated and control authors on their pre-treatment outcome trajectories, and by matching these trends, it indirectly adjusts for time-invariant unobserved confounders that jointly affect both the likelihood of entering a large team and the trends of outcomes, thereby providing a more credible counterfactual for what would have happened without treatment [64]. Another theoretical study also points out that synthetic methods could be used when confounders affecting the control units are believed to exist. In this case, the trends of the control units could be used to control for these confounders, even when they are time-varying [65].

We therefore adopt the SDID method to improve causal identification by balancing pre-treatment trends [64]. Compared to matching-based DID estimators such as PSM-DID, the SDID method can construct a control group based solely on pre-treatment trends of the outcome variable without explicitly requiring observed covariates, while the former crucially rely on the Conditional Independence Assumption, requiring all the potential covariates are observable to achieve balance. As the original SDID does not support staggered adoption, we implement the SDID Event (SDIDE) extension, which accommodates staggered treatment by estimating separate SDID estimators for each treatment year and aggregating results [66]. Confidence intervals are obtained via a placebo procedure, which employs a permutation-based approach to re-estimate the effects on pseudo-treated groups generated from the control units. This method is preferable when the number of treatment units is relatively small in certain adoptive periods [64].

The treatment effects for each of the five cohorts (2000–2004) are estimated separately (All sample / Men / Women). We restrict treated individuals to those who receive treatment in their fourth career year or later to ensure there is at least a three-year window for accurate counterfactual construction. For individuals treated within the first three years (including debut large-team collaborations), we plot average outcome trends over ten years for treated and non-treated authors descriptively. Here, the non-treated group contains only those who never participated in large-team collaborations.

### 4.5 Survival Analysis

To answer the third research question, we employ survival analysis, a statistical technique designed to analyze the time until an

event occurs—in this case, the first-time large-team collaboration. Specifically, we first construct Kaplan-Meier (K-M) survival curves to visualize the cumulative probability of experiencing this event over time, stratified by gender. This visualization approach allows for a clear comparison of the event occurrence dynamics between men and women researchers throughout the early stages of their careers[67].

Subsequently, we estimate Cox proportional hazards models to statistically estimate the effect of gender on the hazard rate of experiencing a first large-team collaboration, for each of the five cohorts from 2000 to 2004:

$$h_i(t) = h_0(t) \exp(\beta \cdot \text{Gender\_women}_i)$$

where  $h_i(t)$  denotes the hazard function of individual  $i$  in year  $t$ , i.e., the  $t$ -th year of their career, and  $h_0(t)$  denotes the baseline hazard function. The Cox model assumes that the effect of gender on the hazard ( $\beta$ ) is generally constant over time (the proportional hazards assumption)[68]. To verify this key assumption, the Schoenfeld residuals test is conducted, which indicates no significant violation of proportionality.

To mitigate the concern that the severe gender imbalance in the sample (approximately 2:8) may not generate robust effects, we further conduct a placebo test. Specifically, we randomly permuted the gender labels 1,000 times while preserving the overall gender proportion, re-estimated the model for each permuted dataset, and obtained a null distribution. We then evaluated the significance of the actual coefficient relative to this empirical null distribution.

#### 4.6 Positionality Statement

Our team includes individuals with different gender identities, as well as individuals from non-Western backgrounds and a variety of educational regions. Being attentive to the ways different social positions and institutional contexts influence collaborative dynamics motivates us to investigate how underlying structural and social factors may advantage certain groups over others.

Our analysis is informed by feminist perspectives [69]: We recognize the significance of the social construction of gender, and we carefully interpret it in relation to institutional and cultural contexts, which helps us frame women’s underrepresentation in large-team collaborations not as an individual deficit, but as a consequence of systemic and structural factors embedded within scientific collaboration networks.

We acknowledge that our positionality also imposes limitations. Lack of personal experience with non-normative gender trajectories may shape the ways we classify gender and interpret related outcomes. We present our analyses with these limitations in mind and welcome critical engagement and dialogue from the broader research community to further refine understanding of gender dynamics in scientific collaboration.

### 5 Result

In this section, we present the results of our statistical analysis. Section 5.1 provides the differential trends in key variables between men and women scholars during the first 10 years of their careers in our sample. Section 5.2 examines RQ1 and RQ2, providing evidence for the impact of first-time large team collaboration on social capital

accumulation. Section 5.3 examines RQ3, demonstrating that gender differences exist in the likelihood of large team collaboration. Section 5.4 presents the robustness checks we conducted.

#### 5.1 Descriptive Statistics

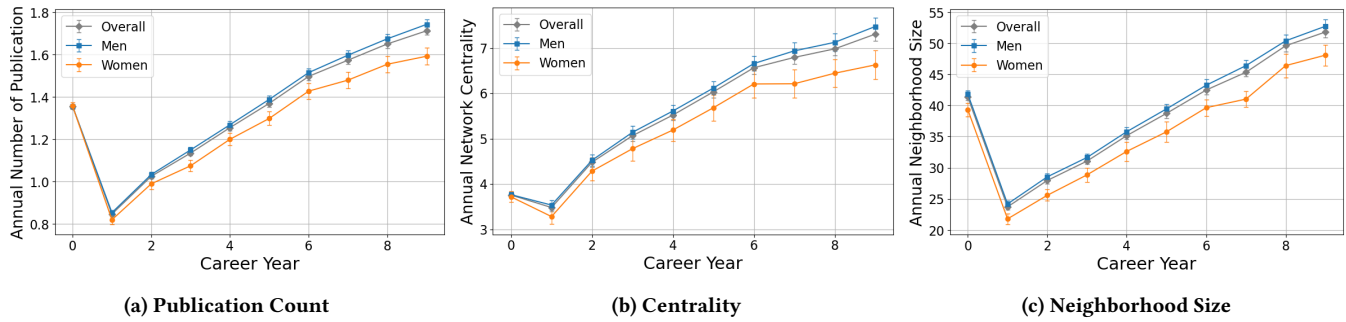
Before presenting the main analyses, we provide descriptive statistics related to collaboration team size and gender. Figure 1 displays, over the first ten years of the career, the average publication count, average centrality, and average annual neighborhood size for the overall sample (across five cohorts), as well as for men and women separately.

By definition, all individuals in the sample have at least one publication in their first career year, which leads to relatively high average values of publication count, centrality, and neighborhood size in that year. These three indicators show a steady upward trend over the following years. Meanwhile, the average trajectories for men consistently exceed those of women, with the gender gap widening over time. This suggests that early-career men scholars tend to (a) mobilize access to academic resources more rapidly, and (b) have more structural advantages and greater access to large research groups.

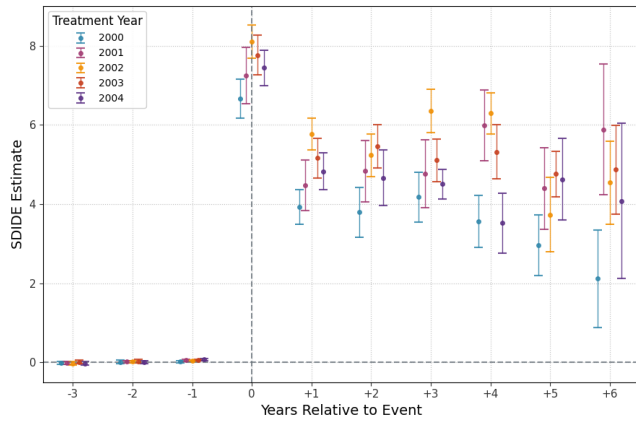
#### 5.2 The Effects of First-time Large-team Collaboration on Centrality and Productivity

Before implementing the SDID models, we conducted a traditional event study (see Appendix C.1). The findings imply that researchers who enter large-team collaborations already exhibit higher structural advantages and publication counts prior to the event, and that gender does not significantly alter the treatment effects. This supports our use of SDID, and it further reinforces our assumption that large-team participation is not merely driven by subjective collaboration preferences but also reflects structural advantages.

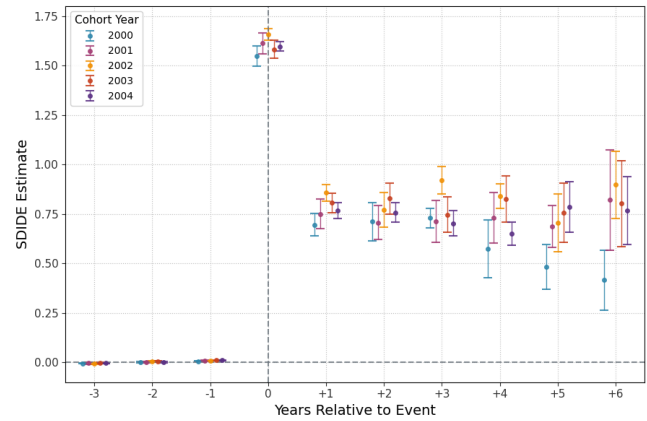
Figure 2 presents the estimation results of the SDID models for network centrality across five cohorts. We note that prior to the event, the SDID method constructs an ideal synthetic control group exhibiting nearly zero difference from the treatment group (In Appendix C.2, we present the synthetic matching results for treated groups in cohort 2004 that received the treatment in different years for demonstrative purposes. The matching results for all cohorts are available in the Supplementary file. The results show that the constructed virtual control groups closely replicate the pre-treatment trends of the treated groups.). Following the first-time large-team collaboration, individuals in the treatment groups across all five cohorts experienced a significantly enhanced growth rate in centrality compared to the control groups, which owned parallel pre-event trends. In general, after the event year, the differences between the treatment and control groups increase by about 5 compared to the pre-event period. As a comparison, the average centrality of the 2004 cohort across the ten years is 5.97, indicating that participation in large research team substantially reshapes the structural positions of junior scholars. Only cohort 2000 shows a declining treatment effect in the later period, while the treatment effects for the remaining four cohorts remain stable through their careers.



**Figure 1: Average trends of publication count, network centrality, and neighborhood size across five cohorts during the first 10 years of their careers. The three lines represent men (circle markers), women (square markers), and the overall cohort average (diamond markers). Women researchers consistently score lower than men and below the overall average, and the gender gaps generally widen as careers progress.**



**Figure 2: SDIDE Estimates of First-time Large-team Collaboration on Network Centrality for the studied five cohorts of early career Computer Scientists (2000-2004). At each time point, five markers appear from left to right in cohort-year order. First-time participation in a large-team publication enhances junior researchers' network centrality by roughly 7 points in the same year, and 4-5 points in the following years.**

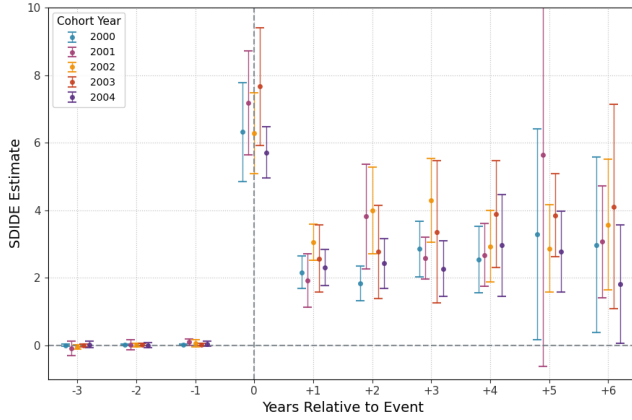


**Figure 3: SDIDE Estimates of First-time Large-team Collaboration on Publication Counts for the studied five cohorts of early career Computer Scientists (2000-2004). At each time point, five markers appear from left to right in cohort-year order. First-time participation in a large-team publication enhances junior researchers' publications by roughly 1.5 additional article(s) in the same year, and 0.5 article more annually in the following years.**

Figure 3 displays the estimation results of the SDID models for publication counts. Similarly, all cohorts demonstrate highly significant treatment effects, with the first-time large-team collaboration bringing nearly 1 additional publications per year for fresh scholars. In average, after the event year, the differences between the treatment and control groups increase by about 0.75 compared to the pre-event period, meaning that treatment units are likely to publish 0.75 more papers per year. Except for cohort 2000, the dynamic treatment effects remain stable across the other four cohorts. It is worth noting that by definition, scholars must have collaborations and publications in the year the event occurs, so the treatment effect in the event year is expected to be higher than in subsequent years.

To account for publication impact when addressing productivity, Figure 4 reports the estimation results of the SDID models for publication counts weighted by RCI. Overall, the results continue to support our hypothesis that first-time large-team collaboration significantly improves early-career scholars' productivity. The notably wider error bars for the 2001 cohort (extending to the x-axis) are primarily attributable to two factors: diminished statistical power due to fewer observations after the third post-event year, and the inherent volatility of the RCI-weighted variable. The latter arises because RCI, following a wide, multiplicative, and approximately log-normal distribution on a log-log scale coordinate, tend to inflate variance of estimates. [61]. Nevertheless, the average estimates remain stable (around 3) over the six-year post-event period, and the overall post-event treatment effect is significant. Intuitively, treated

authors produce papers that are cited as if they had three extra "average" papers' worth of citations compared with the control group.



**Figure 4: SDIDE Estimates of First-time Large-team Collaboration on Publication Counts weighted by RCI for the studied five cohorts of early career Computer Scientists (2000-2004).** At each time point, five markers appear from left to right in cohort-year order. First-time participation in a large-team publication increases junior researchers' citation-weighted output by roughly 7 RCI-weighted articles in the event year and by about 3 additional RCI-weighted articles per year thereafter.

Table 2 presents the estimated pooled coefficients from the SDID models described above, with t-statistics calculated using 10 placebo replications. All coefficients are statistically significant at the 0.001 level. Additionally, we estimated separate SDID models for men and women samples. The results are presented in Table 2. The effect sizes do not differ significantly between the two groups, consistent with the findings from the traditional event study models presented earlier. This suggests that large-team collaboration offers comparable benefits to researchers of both genders.

Finally, we plot average trends in productivity (both measures) and network centrality over ten years for early-treated (treated in the first three years of career life) and control groups. From the Figure 5, we observe that individuals who engaged in large-team collaborations within the first three years consistently show higher average annual productivity and centrality than the control group. Moreover, this difference is similar for both men and women researchers. These findings align with previous results, suggesting that first-time large-team collaboration facilitates the mobilization of academic resources, the accumulation of structural advantages, and that the magnitude of this benefit is comparable across genders.

### 5.3 Gender Disparities in First-time Large-team Collaboration

While descriptive statistics in Section 5.1 may suggest men researchers accumulate more capital via large-team collaborations, we see no systematical difference in terms of treatment effects on men and women groups (see in Section 5.2). However, when we

turn to the timing or the likelihood of participation in large-team collaborations, a gender difference clearly emerges. Figure 6 shows that men researchers are more likely to experience their first large-team collaboration earlier than women researchers, as indicated by the blue survival curves lying above the red ones. Overall, the gender gap in the proportion of researchers having their first-time large-team collaboration tends to widen as their career progresses.

Cox proportional hazards model is used to statistically test whether men are more likely than women to enter large-team collaborations. As shown in Figure 6, the model based on the pooled data from all five cohorts indicates that, on average, men researchers are 16% more likely than women colleagues to experience their first-time large-team collaboration, and this effect is statistically significant. The gender coefficients in the Cox models fitted separately for each of the five cohorts are also significantly positive. These results suggest that gender is indeed an important factor influencing early-career access to large-team collaborations.

The 1,000 placebo estimates of the Cox proportional hazards model also yield significant results ( $p = 0.0010$ ; Figure 7), indicating that the estimated coefficients are unlikely to be driven by chance due to imbalanced gender ratios, thus supporting the robustness of the observed gender differences.

Because the validity of our survival analysis relies on the accuracy of the gender classifier, we restrict the sample to authors whose predicted probability of belonging to a given gender category exceeded 80%, even though such a restriction does not align with established best practices (see Appendix B). Re-estimating on this subsample yields coefficients nearly identical to those in the full-sample analysis (with overall hazard ratio = 1.16), indicating that the gender differences are dominated by authors whose gender classifications can be inferred with higher certainty.

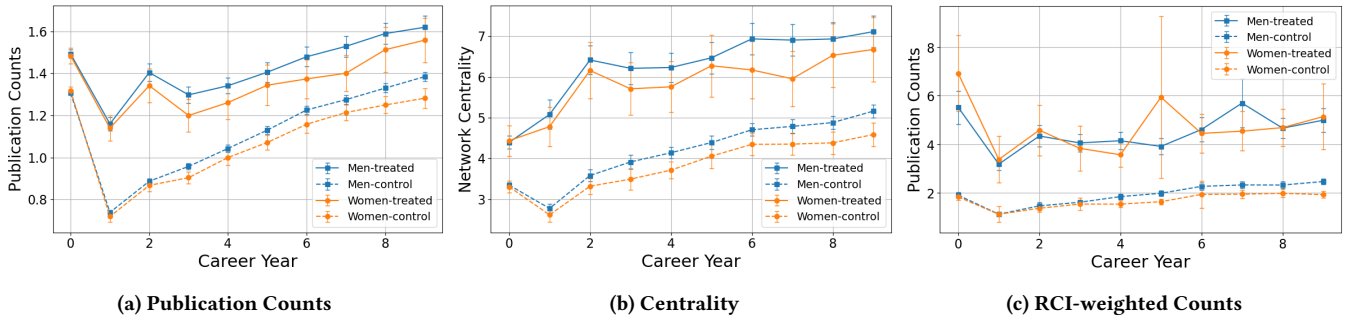
### 5.4 Robustness Tests

We further perform a series of robustness checks to verify the effects of first-time large-team collaboration on early-career scholars' social capital reported above. First, we test the robustness of SDID estimation by changing the definition of event indicator. Specifically, we replace the event indicator with three alternative definitions: those having a team size greater than (a) the 90% percentile of team sizes within the year, (b) the 90% percentile of the largest team sizes across the first 10 years of the fresh authors, and (c) a fixed threshold of 8 coauthors. Second, we change the measure of centrality and publication count, from the annual network centrality computed on yearly snapshots (collaboration network in year  $Y_i$ ) to the cumulative network centrality computed on cumulative collaboration networks (from year  $Y_i$  to  $Y_j$ ); Also, we incorporate two additional measures of productivity: (1) the publication count weighted by venue prestige, and (2) the average RCI. Third, we exclude the top 5% of sample scholars ranked by the total number of publication produced during the observation period.

Finally, we conduct a subfield analysis. SciSciNet has only annotated a subset of publications with subfields, making it difficult to conduct comprehensive discipline-sensitive analysis. Based on a rough author-level classification (see Appendix D), we select three subfields with distinct collaboration patterns and gender ratios, as

**Table 2: SDID Estimated Effects of Large-Team Collaboration by Cohort and Gender. The effects of large-team collaboration are not significantly different between men and women researchers in any cohort.**

	Centrality			Publication Count			RCI-weighted Count		
	All	Men	Women	All	Men	Women	All	Men	Women
Cohort 2000									
Treatment Effect	4.389	4.291	4.704	0.857	0.858	0.845	3.288	3.275	3.247
Std. Err	(0.24)	(0.25)	(0.40)	(0.04)	(0.04)	(0.10)	(0.17)	(0.19)	(0.23)
Cohort 2001									
Treatment Effect	5.413	5.196	5.900	0.924	0.915	0.906	3.939	3.950	3.739
Std. Err	(0.23)	(0.28)	(0.48)	(0.02)	(0.04)	(0.07)	(0.21)	(0.25)	0.55
Cohort 2002									
Treatment Effect	6.142	6.573	4.783	1.016	1.037	0.937	4.120	4.229	3.642
Std. Err	(0.20)	(0.21)	(0.42)	(0.03)	(0.03)	(0.07)	(0.15)	0.20	0.20
Cohort 2003									
Treatment Effect	5.756	5.638	5.729	0.969	0.954	1.018	4.147	4.670	2.076
Std. Err	(0.25)	(0.29)	(0.60)	(0.03)	(0.03)	(0.07)	(0.17)	0.14	0.53
Cohort 2004									
Treatment Effect	5.144	4.930	5.560	0.928	0.912	0.957	3.169	3.059	3.546
Std. Err	(0.21)	(0.26)	(0.40)	(0.02)	(0.03)	(0.05)	(0.12)	(0.17)	(0.21)

**Figure 5: Trends in Publication Counts (left) and Centrality (right): Early-treated vs. Control Groups, by Gender. Early-treated groups (solid line) show significantly more annual publications and network centrality compared to control groups (dashed line). While treated men (square markers) and women (circle markers) researchers are comparable in both values, the control groups exhibit a larger gender gap that widens along with their career progress, particularly in publication counts.**

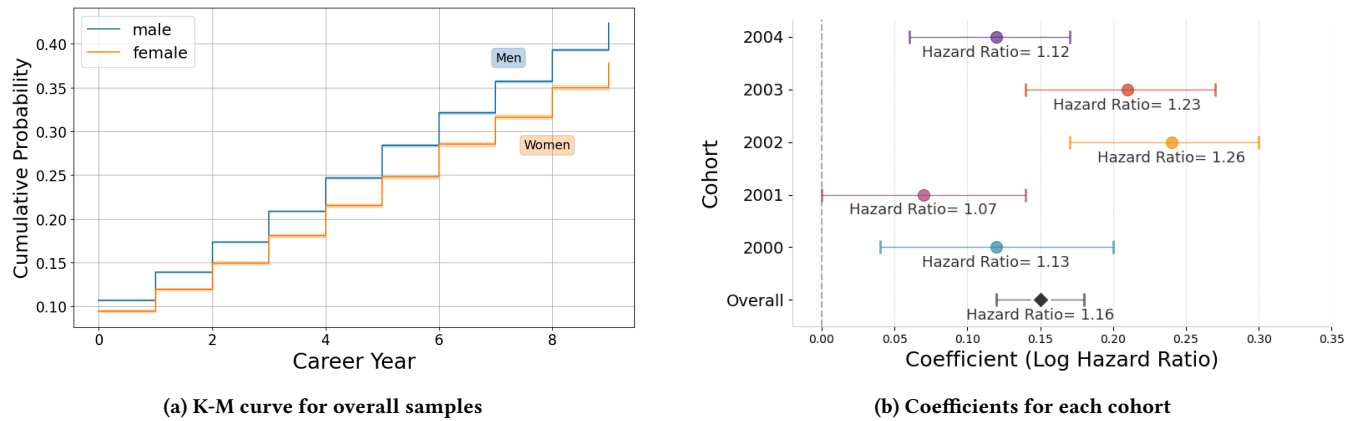
well as sufficiently large sample sizes (Human-Computer Interaction, Data Science, and Computer Security) and estimate separate SDID models for each of them. This allows us to assess whether the observed effects of first-time large-team collaboration are consistent across different academic communities and collaboration norms, and to examine potential systematic differences in treatment effects across subfields. We further conducted separate survival analyses for each subfield to ensure that, across research environments with different gender compositions and collaboration norms, women consistently face more limited opportunities to enter large-team collaborations.

The results in the Appendix D indicate that the effects on the productivity and centrality estimated by our robustness testing models are all valid. Overall, the above results suggest that the effects of

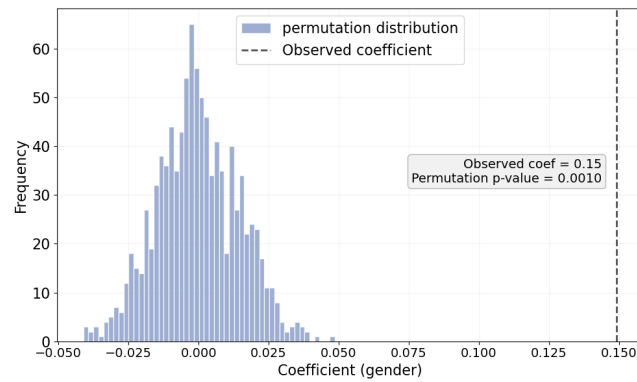
first-time large-team collaboration on early-career scholars' career development are robust.

## 6 Discussion

In this section, we contextualize our findings and discuss their broader implications for the career development of early-career scholars. Section 6.1 examines the benefits that early-career scholars may gain from large-team collaboration, as well as possible mechanisms. Section 6.2 addresses gender disparities in access to large-team collaboration and their implications for the accumulation of social capital. Section 6.3 builds on these findings to explore potential measures for mitigating inequality in academia under current research conditions.



**Figure 6: Survival Analysis for Men and Women Group on First-time Large-team Collaboration with the studied five cohorts of early career Computer Scientists (2000–2004).** The cumulative probability of participating first-time in a large-team publication (left) is consistently higher for men researchers compared to their women peers. Results from Cox proportional hazard models (right) indicates that men researchers are on average 1.16 times more likely than their women peers to co-author at least one large-team publication, and this gender gap persists across all cohorts.



**Figure 7: Placebo distribution of the gender coefficient from 1,000 permutation tests of the Cox proportional hazards model.** The vertical dashed line indicates the observed coefficient, while the histogram shows the distribution of coefficients obtained under random permutation of gender labels (gender composition is kept).

## 6.1 Advantages of Large-team Collaboration

Our empirical findings support the positive impact of large-team collaboration on scholars' subsequent structural position and professional returns in their early-stage career. This aligns with a prior study on team science, which found that being in a large research group is positively associated with the academic performance of junior scholars who remain in academia[2].

In particular, our findings demonstrate that large-team collaborations strengthen early-career scholars' structural positions (network centrality) and academic productivity (publication counts). This advances prior understanding of how large-team collaborations contribute to the accumulation of social capital. On the one

hand, participation in large teams increases the likelihood and intensity of collaborating with influential senior scholars during the early career stage. In other words, after joining large-team collaborations, junior scholars have greater opportunities to establish close ties with high-capital scholars, thereby expanding their access to information, reputation, and future opportunities[18]. On the other hand, participation in large-team collaboration can increase publication frequency and scholarly influence. In the 'publish or perish' academic environment, this provides a strong advantage, not only by increasing production and visibility but also by demonstrating the ability of scholars to effectively mobilize the surrounding social capital into measurable academic performance. Taken together, the accumulation of centrality and productivity are interrelated and mutually reinforcing processes in academic career development with significant values. Theoretically, our findings highlight the dynamic nature of social capital in the academia: mobilization of resources, including socially embedded ones, yields substantial returns and generates structural advantages, while one's structural position simultaneously facilitates access to resources and contributes to the success of future mobilization. From the empirical perspective, we show that the social capital framework offers a powerful analytical lens for understanding the mechanisms of academic network development in large-team collaborations. Importantly, these benefits are not contingent on the duration of collaboration; even short-term engagements, such as internships or project-based work, can yield structural opportunities and tangible returns.

The advantages of large-team collaboration can arise simultaneously from several sources:

The first is the inherent advantages of large teams. For scholars who remain embedded in large teams or research groups over a sustained period, they are continuously more likely to gain access to substantial funding, large-scale datasets, and state-of-the-art facilities[17, 19, 20]. Meanwhile, large teams also have clear advantages in carrying out labor-intensive tasks, which is particularly important in certain areas of computer science, such as HCI[70],

that emphasize experimentation and data processing. These benefits accelerate the process from idea to publication for early-career scholars.

Second, large-team collaboration helps to gain higher academic visibility. Large-team projects are more likely to be published in high-impact venues and to attract attention from a wider audience [16, 17]. This is also supported by our robustness tests (see Appendix D). For early-career scholars, being associated with such publications provides important symbolic capital and signals competence to the broader community.

Third, large research teams are able to provide intensive and diverse research training and resources. Participation in large teams exposes early-career scholars to diverse methodological expertise and a breadth of disciplinary perspectives [20, 22]. This offers opportunities for intensive on-the-job training, mentorship from senior collaborators, and early immersion in complex, systematic scientific practices.

Finally, Large teams embed early-career scholars in dense and diverse collaboration networks. Importantly, our empirical evidence (see Appendix E) shows that such participation not only creates immediate ties within the team but also facilitates the extension of collaboration beyond it. Scholars engaged in large teams are more likely to establish new and enduring partnerships with a broader set of researchers, rather than repeatedly collaborating with a small circle of colleagues or engaging in one-off projects. In this sense, large-team collaboration serves as a platform for building durable academic relationships and for positioning early-career scholars in more central and influential roles within the collaboration network.

Our findings align with prior organizational research suggesting that scientific collaboration is not gender-neutral. Unequal access to mentoring [45], less brokerage positions [42] and low-visibility tasks [46] may continuously disadvantage women in scientific teams. The persistent gender gap in large-team participation we observe may reflect and further reinforce these broader processes, which keep reproducing gender inequality in the academia.

These findings also enter into dialogue with another line of team science research, which has shown that large-team collaboration does not consistently improve team-level performance, and in some cases may even negatively affect the frequency and impact of publications [2, 21, 23]. This discrepancy suggests that the effects of large-team collaboration are subject to *multi-level effects*. At the **team level**, prior studies have found that as team size increases, coordination and communication costs also rise, task division becomes more complex, and conflicts more likely, which may reduce innovation efficiency and overall output. At the **individual level**, and particularly for early-career scholars, the situation differs. Even when large teams are not more efficient as collective entities, junior researchers can still gain valuable opportunities through the mechanisms we have discussed. Moreover, even in cases where team-level performance is not optimal, individual authorship experience within large teams may itself function as symbolic capital, signaling value in subsequent career development. We further concur with prior research emphasizing that diversity in academic teams may matter more than sheer scale [23]. As we have argued, for individual early-career scholars, experience in large-team publications enhances their likelihood of encountering and collaborating

with a wider range of researchers, which in turn facilitates the establishment of broader and more enduring collaborative ties.

## 6.2 Gender Disparity

Our study also reveals the disadvantaged position of women scholars in accessing large-team collaboration during their early-stage career. To our knowledge, no prior work has quantitatively identified the causal link between gender and the likelihood of joining large research teams at the early-career stage. Nevertheless, our findings align with prior studies, which point to the marginalization of women in academic collaboration networks [40, 43].

Given the significant advantages of large-team collaboration, this gender gap represents a potential driver of inequality in academia. Women's restricted opportunities to participate in large teams may translate into limited network centrality and reduced productivity, thereby hindering their career development in both the short and long run. This partly explains why we observe increasing gender gaps as careers progress (Figure 1), and may also provide insight into the glass ceiling and leaky pipeline phenomena—that is, “inherent systemic disadvantages for women in academic fields, which contribute to the leaks during each stage of the academic pipeline.” [71] More broadly, our findings offer a new perspective on gender dynamics in academic collaboration networks. While previous research has examined mechanisms such as gender ratios, homophily, and preferential attachment models to explain structural gender differences [41, 51], few studies have considered the role of team size. Specifically, prior studies on gender inequality within the HCI community and CHI publications have primarily examined disparities in overall gender representation [72–74], while overlooking the less visible mechanisms through which gender inequality is reproduced, particularly unequal access to advantageous collaboration patterns. For HCI, as a subfield that places strong emphasis on interdisciplinarity, experimentation, and empirical research, large-team collaboration represents an important and advantageous mode of collaboration [13]. Our findings suggest that despite the steady rise in women representation in academia, the gender gap in large-team participation has not narrowed, and should therefore be integrated into gender-equity conversations within the HCI community.

Finally, another gender-related finding contributes to a new understanding of large-team collaboration. Unlike prior studies that interpret team size as a matter of collaborative preference [40], we view large-team collaboration as a competitive resource for capital accumulation, in which women are structurally disadvantaged. This perspective urges us to reconsider the complex mechanisms underlying academic collaboration and to explore how more inclusive forms of team collaboration could help mitigate systemic discrimination.

## 6.3 Equity and Inclusion in Academic Collaboration

Our findings highlight that large-team collaborations can significantly boost early-career researchers' network centrality and productivity. However, our findings also indicate that large-team collaboration is a type of highly competitive resource, and we observe notable gender disparities in access to such resource. In fields like

Computer Science, particularly in labor-intensive subfields such as HCI that rely on experimental work and public fundings, this unequal access can limit the accumulation of social capital for disadvantaged groups. Over time, these differences may reinforce systemic inequities, as early visibility and productivity advantages compound throughout a career. Promoting equity and inclusion is therefore not only a moral imperative but also a practical approach to mitigating systemic disadvantages and preventing the perpetuation of weak positions within academic networks.

Addressing these disparities requires both institutional and individual strategies in design for policy implementations and collaborative technologies. Institutions could promote more equitable participation by adopting transparent and open allocation of collaboration opportunities, mitigating potential discrimination towards underrepresented groups in organizational processes, and valuing qualitative contributions such as research quality or intellectual originality rather than relying solely on quantitative metrics like publication counts. In addition, academic journals and publishers could place greater emphasis on work produced by small research teams while demonstrating more original thoughts. At the personal level, disadvantaged early-career scholars may benefit from actively leveraging network strategies to seek diverse collaboration opportunities. Currently, several supporting mechanism to promote inclusive interactions in collaborative systems have been explored: diverse identity expressions [75], emphasis on group cohesion over identity [76] and gender-aware interactive designs [77, 78]. Particularly, avatars have been shown to reduce gender bias in virtual interactions with designs such as gender-neutral avatars, or gender-matching [75, 76]. Furthermore, to address potential bias in organisational processes, designers and policymakers can attend to different phases in coordinating collaboration. Collaboration recommendation system [79] can adopt fairness-aware algorithms, while multidimensional mechanisms and retrospective usage of visualization tool can support fairness in team collaboration settings [80]. Our findings encourage further examination of design implications with respect to not only accessibility but also resource/visibility distribution and long-term equity outcomes. For example, one might ask whether shifting from majority-minority dynamics (by leveraging more diverse or cohesive identities) helps improving visibility and opportunity gap between privileged and marginalized groups, and if so, to what extent and how it plays out over time. We envision this approach helpful in gaining comprehensive knowledge on benefits and challenges different design solutions introduce.

## 7 Limitations

First, we focus exclusively on early-career researchers in the field of Computer Science. Prior research suggests that collaboration patterns and their consequences vary across disciplines [10, 17]. The effects we document may therefore not generalize to fields where large-team collaboration is more or less prevalent, such as Psychology and Social Science. Replicating our analyses in other domains would help assess potential heterogeneity and explain the nuanced mechanism behind the association between large-team collaboration and social capital in academia.

Second, the gender variable reflects only inferred binary gender based on their names, which inevitably simplifies the complex

realities of cross-gender, non-binary, and otherwise diverse gender identities, and thus may not fully correspond to scholars' real experience and self-identification. Besides, the degree of gender ambiguity is not evenly distributed across linguistic or cultural contexts. For instance, as documented in prior research, inferring gender from romanized Chinese names is often more challenging, which may introduce regionally patterned misclassification and representational imbalance in our analysis [50, 81]. Although the proportion of ambiguous names in our sample is relatively small (see Appendix B), future work would benefit from datasets in which gender information can be observed rather than inferred.

Third, the ways of constructing event indicators in this study are imperfect. Although various robustness tests have been conducted, we consistently treat large-team collaboration as a binary condition based simply on coauthor counts, which overlooks important heterogeneity, for example, sustained participation in a large laboratory versus occasional involvement, collaborations within a single group versus across groups, variation in team size, or variation in actual access to team resources. Future research should leverage more detailed data, and potentially qualitative evidence, to unpack more nuanced mechanisms.

Forth, we rely on network centrality and publication counts as proxies for early-career scholars' structural advantages and success, which also has shortcomings. First, our measure of centrality assumes that structural advantages are positively associated with tie strength and the influence of collaborators, without accounting for the quality or impact of the coauthored work. Second, while we have accounted for productivity using citation-based and venue-weighted metrics, these proxies still cannot fully capture the multifaceted nature of scholarly success, such as breakthrough contributions or long-term intellectual significance that may not be immediately reflected in citations or venue landings. Future research could address these issues by incorporating altmetrics to better capture scholarly influence. Additionally, extending network measures to account for publication quality, for example, weighting ties by the impact of joint work, may provide a more nuanced understanding of how large-team collaborations contribute to social capital of academia.

Fifth, our definition of "fresh" authors is also not without limitations. We include all authors with no prior publications recorded in SciSciNet. However, some of these authors may have published in journals or conferences not indexed by SciSciNet, such as non-English journals or monographs, potentially introducing bias into our analysis.

Finally, due to data accessibility constraints, our statistical models do not directly control for covariates, leaving the findings susceptible to unobserved confounding. While we employ SDID to indirectly address potential confounding, and conduct subfield analyses across three fields that differ substantially in their collaboration patterns to account for field-level disparities, future research could focus on subsamples where richer institutional and biographical information is available.

## References

- [1] Vincent Larivière, Yves Gingras, Cassidy R Sugimoto, and Andrew Tsou. 2015. Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology* 66, 7 (2015), 1323–1332.

- [2] Yanmeng Xing, Yifang Ma, Ying Fan, Roberta Sinatra, and An Zeng. 2025. Academic mentees thrive in big groups, but survive in small groups. 902–916 pages.
- [3] Noriko Hara, Paul Solomon, Seung-Lye Kim, and Diane H Sonnenwald. 2003. An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology* 54, 10 (2003), 952–965.
- [4] Alice N Pell. 1996. Fixing the leaky pipeline: women scientists in academia. *Journal of animal science* 74, 11 (1996), 2843–2848.
- [5] Mark Lutter and Martin Schröder. 2020. Is there a motherhood penalty in academia? The gendered effect of children on academic publications in German sociology. *European Sociological Review* 36, 3 (2020), 442–459.
- [6] Christiane Gross, Monika Jungbauer-Gans, and Natascha Nisic. 2017. Cooperation and career chances in science. *Social dilemmas, institutions, and the evolution of cooperation* (2017), 165–188.
- [7] Ronald S Burt. 2018. Structural holes. In *Social stratification*. Routledge, 659–663.
- [8] Phillip Bonacich. 1987. Power and centrality: A family of measures. *American journal of sociology* 92, 5 (1987), 1170–1182.
- [9] Ying Liu, Ming Tang, Tao Zhou, and Younghae Do. 2016. Identify influential spreaders in complex networks, the role of neighborhood. *Physica A: Statistical Mechanics and its Applications* 452 (2016), 289–298.
- [10] Mark EJ Newman. 2001. Scientific collaboration networks. I. Network construction and fundamental results. *Physical review E* 64, 1 (2001), 016131.
- [11] Mark EJ Newman. 2001. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical review E* 64, 1 (2001), 016132.
- [12] Zihang Lin, Yian Yin, Lu Liu, and Dashun Wang. 2023. SciSciNet: A large-scale open data lake for the science of science research. *Scientific Data* 10, 1 (2023), 315.
- [13] Jonas Oppenlaender. 2025. Past, Present, and Future of Citation Practices in HCI. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 453, 16 pages. doi:10.1145/3706598.3713556
- [14] Dana McKay, Huiwen Zhang, and George Buchanan. 2022. Who am I, and who are you, and who are we? A Scientometric Analysis of Gender and Geography in HCI. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 396, 19 pages.
- [15] Nan Lin. 2001. *Social Capital: A Theory of Social Structure and Action*. Cambridge University Press.
- [16] Stefan Wuchty, Benjamin F. Jones, and Brian Uzzi. 2007. The Increasing Dominance of Teams in Production of Knowledge. *Science* 316, 5827 (May 2007), 1036–1039.
- [17] James D Adams, Grant C Black, J Roger Clemmons, and Paula E Stephan. 2005. Scientific teams and institutional collaborations: Evidence from US universities, 1981–1999. *Research policy* 34, 3 (2005), 259–285.
- [18] Roger Guimerà, Brian Uzzi, Jarrett Spiro, and Luís A. Nunes Amaral. 2005. Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance. *Science* 308, 5722 (April 2005), 697–702.
- [19] K. Brad Wray. 2002. The Epistemic Significance of Collaborative Research. *Philosophy of Science* 69, 1 (March 2002), 150–168.
- [20] Beatrice D'Ippolito and Charles-Clemens Rüling. 2019. Research collaboration in Large Scale Research Infrastructures: Collaboration types and policy implications. *Research Policy* 48, 5 (June 2019), 1282–1296.
- [21] Alexander M. Petersen, Massimo Riccaboni, H. Eugene Stanley, and Fabio Pam-molli. 2012. Persistence and uncertainty in the academic career. *Proceedings of the National Academy of Sciences* 109, 14 (March 2012), 5213–5218.
- [22] Noriko Hara, Paul Solomon, Seung-Lye Kim, and Diane H. Sonnenwald. 2003. An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology* 54, 10 (June 2003), 952–965.
- [23] Yi Bu, Ying Ding, Xingkun Liang, and Dakota S. Murray. 2017. Understanding persistent scientific collaboration. *Journal of the Association for Information Science and Technology* 69, 3 (Oct. 2017), 438–448.
- [24] Barry Bozeman and Elizabeth Corley. 2004. Scientists' collaboration strategies: implications for scientific and technical human capital. *Research policy* 33, 4 (2004), 599–616.
- [25] Harald Bauder, Charity-Ann Hannan, and Omar Lujan. 2017. International experience in the academic field: Knowledge production, symbolic capital, and mobility fetishism. *Population, Space and Place* 23, 6 (2017), e2040.
- [26] Pierre Bourdieu. 1986. *âœThe Forms of Capitalâœ*, dalam J. Richardson, ed. *Handbook of Theory and Research for the Sociology of Education*.
- [27] Maggi WH Leung. 2013. 'Read ten thousand books, walk ten thousand miles': Geographical mobility and capital accumulation among Chinese scholars. *Transactions of the Institute of British Geographers* 38, 2 (2013), 311–324.
- [28] Chao Lu, Chenwei Zhang, Chengrui Xiao, and Ying Ding. 2022. Contributorship in scientific collaborations: The perspective of contribution-based byline orders. *Information Processing & Management* 59, 3 (May 2022), 102944.
- [29] Huimin Xu, Yi Bu, Meijun Liu, Chenwei Zhang, Mengyi Sun, Yi Zhang, Eric Meyer, Eduardo Salas, and Ying Ding. 2022. Team power dynamics and team impact: New perspectives on scientific collaboration using career age as a proxy for team power. *Journal of the Association for Information Science and Technology* 73, 10 (May 2022), 1489–1505.
- [30] Alireza Abbasi. 2016. A longitudinal analysis of link formation on collaboration networks. *Journal of Informetrics* 10, 3 (2016), 685–692.
- [31] Luigi Aldieri, Maxim Kotsemir, and Concetto Paolo Vinci. 2018. The impact of research collaboration on academic performance: An empirical analysis for some European countries. *Socio-Economic Planning Sciences* 62 (2018), 13–30.
- [32] Yong Liu, Jorge Goncalves, Denzil Ferreira, Bei Xiao, Simo Hosio, and Vassilis Kostakos. 2014. CHI 1994–2013: mapping two decades of intellectual progress through co-word analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, 3553–3562.
- [33] Nibing Zhu, Chang Liu, and Zhilin Yang. 2021. Team Size, Research Variety, and Research Performance: Do Coauthors' Coauthors Matter? *Journal of Informetrics* 15, 4 (Nov. 2021), 101205.
- [34] Clement Lee, Andrew Garbett, Junyan Wang, Bingzhang Hu, and Dan Jackson. 2019. Weaving the Topics of CHI: Using Citation Network Analysis to Explore Emerging Trends. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, 1–6.
- [35] Stephen J. Ceci, Donna K. Ginther, Shulamit Kahn, and Wendy M. Williams. 2014. Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest* 15, 3 (Nov. 2014), 75–141.
- [36] Chen Avin, Barbara Keller, Zvi Lotker, Claire Mathieu, David Peleg, and Yvonne-Anne Pignolet. 2015. Homophily and the glass ceiling effect in social networks. In *Proceedings of the 2015 conference on innovations in theoretical computer science*. 41–50.
- [37] Olga Bain and William Cummings. 2000. Academe's Glass Ceiling: Societal, Professional-Organizational, and Institutional Barriers to the Career Advancement of Academic Women. *Comparative Education Review* 44, 4 (Nov. 2000), 493–514.
- [38] Marc J. Lerchenmueller and Olav Sorenson. 2018. The gender gap in early career transitions in the life sciences. *Research Policy* 47, 6 (July 2018), 1007–1017.
- [39] Anne E. Lincoln, Stephanie Pincus, Janet Bandows Koster, and Phoebe S. Leboy. 2012. The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science* 42, 2 (Feb. 2012), 307–320.
- [40] Silvia Knobloch-Westerwick, Carroll J. Glynn, and Michael Huges. 2013. The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest. *Science Communication* 35, 5 (Feb. 2013), 603–625.
- [41] Matthias Potthoff and Fabian Zimmermann. 2017. Is there a gender-based fragmentation of communication science? An investigation of the reasons for the apparent gender homophily in citations. *Scientometrics* 112, 2 (April 2017), 1047–1063.
- [42] Madalina Vlasceanu, Miroslav Dudik, and Ida Momennejad. 2022. Interdisciplinary, Gender Diversity, and Network Structure Predict the Centrality of AI Organizations. In *2022 ACM Conference on Fairness Accountability and Transparency (FAccT '22)*. ACM, 1–10.
- [43] Marek Kwiec and Wojciech Roszka. 2021. Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics* 15, 3 (Aug. 2021), 101171.
- [44] Giovanni Abramo, Ciriaco Andrea D'Angelo, and Flavia Di Costa. 2019. A gender analysis of top scientists' collaboration behavior: evidence from Italy. *Scientometrics* 120, 2 (May 2019), 405–418.
- [45] Elena Greguletz, Marjo-Riitta Diehl, and Karin Kreutzer. 2018. Why women build less effective networks than men: The role of structural exclusion and personal hesitation. *Human Relations* 72, 7 (Nov. 2018), 1234–1261.
- [46] Katherine N. Quinn, Michelle M. Kelley, Kathryn L. McGill, Emily M. Smith, Zachary Whipples, and N. G. Holmes. 2020. Group roles in unstructured labs show inequitable gender divide. *Physical Review Physics Education Research* 16, 1 (May 2020).
- [47] Matthew B. Ross, Britta M. Glennon, Raviv Murciano-Goroff, Enrico G. Berkes, Bruce A. Weinberg, and Julia I. Lane. 2022. Women are credited less in science than men. *Nature* 608, 7921 (June 2022), 135–145.
- [48] Aparna Joshi. 2014. By Whom and When Is Women's Expertise Recognized? The Interactive Effects of Gender and Education in Science and Engineering Teams. *Administrative Science Quarterly* 59, 2 (March 2014), 202–239.
- [49] Katie Spoon, Nicholas LaBerge, K. Hunter Wapman, Sam Zhang, Allison C. Morgan, Mirta Galesic, Bailey K. Fosdick, Daniel B. Larremore, and Aaron Clauset. 2023. Gender and retention patterns among U.S. faculty. *Science Advances* 9, 42 (Oct. 2023).
- [50] Ian Van Buskirk, Aaron Clauset, and Daniel B Larremore. 2023. An Open-Source Cultural Consensus Approach to Name-Based Gender Classification. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 866–877. <https://github.com/ianvanbuskirk/nbgc>.
- [51] Anh Duong Nguyen, Jiaming Jiang, and Barbara Keller. 2025. Mind the Gap: Gender, Homophily and the Glass Ceiling in Academic Networks. *Proceedings of the ACM on Human-Computer Interaction* 9, CSCW (2025).

- [52] Donna Haraway. 2013. Situated knowledges: The science question in feminism and the privilege of partial perspective 1. In *Women, science, and technology*. Routledge, 455–472.
- [53] Marieke Van den Brink and Yvonne Benschop. 2014. Gender in academic networking: The role of gatekeepers in professorial recruitment. *Journal of management studies* 51, 3 (2014), 460–492.
- [54] Mathias W Nielsen. 2016. Limits to meritocracy? Gender in academic recruitment and promotion processes. *Science and Public Policy* 43, 3 (2016), 386–399.
- [55] Aldo Geuna and Sotaro Shibayama. 2015. Moving out of academic research: Why do scientists stop doing research? In *Global mobility of research scientists*. Elsevier, 271–303.
- [56] Andrea Kis, Elena Mas Tur, Daniël Lakens, Krist Vaesen, and Wybo Houkes. 2022. Leaving academia: PhD attrition and unhealthy research environments. *Plos one* 17, 10 (2022), e0274976.
- [57] Xiao Han T Zeng, Jordi Duch, Marta Sales-Pardo, João AG Moreira, Filippo Radicchi, Haroldo V Ribeiro, Teresa K Woodruff, and Luís A Nunes Amaral. 2016. Differences in collaboration patterns across discipline, career stage, and gender. *PLoS biology* 14, 11 (2016), e1002573.
- [58] Chi-Lan Yang, Chien Wen Yuan, and Hao-Chuan Wang. 2019. When knowledge network is social network: Understanding collaborative knowledge transfer in workplace. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23.
- [59] Andrea Landherr, Bettina Friedl, and Julia Heidemann. 2010. A critical review of centrality measures in social networks. *Business & Information Systems Engineering* 2, 6 (2010), 371–385.
- [60] Junyi Wang, Xiaoni Hou, Kezan Li, and Yong Ding. 2017. A novel weight neighborhood centrality algorithm for identifying influential spreaders in complex networks. *Physica A: Statistical Mechanics and its Applications* 475 (2017), 88–105.
- [61] Filippo Radicchi, Santo Fortunato, and Claudio Castellano. 2008. Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences* 105, 45 (2008), 17268–17272.
- [62] Microsoft Corporation. 2022. *Microsoft Academic Services Graph – Resources & FAQ*. <https://docs.microsoft.com/en-us/academic-services/graph/resources-faq> Accessed: 2025-11-14.
- [63] Andrew M Ryan, Evangelos Kontopantelis, Ariel Linden, and James F Burgess Jr. 2019. Now trending: Coping with non-parallel trends in difference-in-differences analysis. *Statistical methods in medical research* 28, 12 (2019), 3697–3711.
- [64] Damian Clarke, Daniel Pailaño, Susan Athey, and Guido W Imbens. 2023. Synthetic difference-in-differences estimation. (2023).
- [65] Bruno Ferman and Cristine Pinto. 2021. Synthetic controls with imperfect pretreatment fit. *Quantitative Economics* 12, 4 (2021), 1197–1221.
- [66] Zachary Porreca. 2022. Synthetic difference-in-differences estimation with staggered treatment timing. *Economics Letters* 220 (2022), 110874.
- [67] Siddhartha Asthana, Pushpendra Singh, and Parul Gupta. 2015. Survival analysis: objective assessment of wait time in HCI. In *Proceedings of the 33rd annual acm conference on human factors in computing systems*. 367–376.
- [68] S Andrew Sheppard, Julian Turner, Jacob Thebault-Spieker, Haiyi Zhu, and Loren Terveen. 2017. Never too old, cold or dry to watch the sky: A survival analysis of citizen science volunteerism. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–21.
- [69] Barbara J Risman. 2018. Gender as a social structure. In *Handbook of the Sociology of Gender*. Springer, 19–43.
- [70] Elena Agapie, Shefali Halder, and Sharmaine G Poblete. 2022. Using HCI in cross-disciplinary teams: a case study of academic collaboration in HCI-health teams in the US using a team science perspective. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–35.
- [71] Courtney E Gasser and Katharine S Shaffer. 2014. Career development of women in academia: Traversing the leaky pipeline. *Professional Counselor* 4, 4 (2014), 332–352.
- [72] Dana McKay and George Buchanan. 2020. Feed the tree: Representation of australia-based academic women at HCI conferences. In *Proceedings of the 32nd Australian Conference on Human-Computer Interaction*. 263–269.
- [73] Joseph 'Jofish' Kaye. 2009. Some statistical analyses of CHI. In *CHI'09 extended abstracts on human factors in computing systems*. 2585–2594.
- [74] Divy Thakkar, Nithya Sambasivan, Purva Kulkarni, Pratap Kalenahalli Sudarshan, and Kentaro Toyama. 2018. The unexpected entry and exodus of women in computing and HCI in India. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [75] Kaiwen Li, Shuyan Chen, Yifei Cai, Yifei Chen, and Ningning Xu. 2025. LivehouseVR: Creating an Immersive and Inclusive Social VR Experience for DJs and Audiences towards Gender Equality. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. ACM, 1–8. doi:10.1145/3706599.3720299
- [76] Alim Yassien, ElHassan B. Makled, Passant Elagroudy, Nouran Sadek, and Samir Abdennadher. 2021. Give-Me-A-Hand: The Effect of Partner's Gender on Collaboration Quality in Virtual Reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, 1–6. doi:10.1145/3411763.3451601
- [77] Carolin Wienrich, Marc Erich Latoschik, and David Obremski. 2024. Gender Differences and Social Design in Human-AI Collaboration: Insights from Virtual Cobot Interactions Under Varying Task Loads. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. ACM, 1–8. doi:10.1145/3613905.3650827
- [78] Ana O. Henriques, Sónia Rafael, Victor M Almeida, and José Gomes Pinto. 2023. The problem with gender-blind design and how we might begin to address it: A model for intersectional feminist ethical deliberation. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, 1–12. doi:10.1145/3544549.3582750
- [79] David W McDonald. 2003. Recommending collaboration with social networks: a comparative evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 593–600.
- [80] Ryan Kelly, Leon Watts, and Stephen J Payne. 2016. Can visualization of contributions support fairness in collaboration?. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco California USA). ACM, New York, NY, USA.
- [81] Junming Huang, Alexander J Gates, Roberta Sinatra, and Albert-László Barabási. 2020. Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the national academy of sciences* 117, 9 (2020), 4609–4616.
- [82] David Donoho. 2017. 50 years of data science. *Journal of Computational and Graphical Statistics* 26, 4 (2017), 745–766.

## A SciSciNet Fields Statistics

The field-specific descriptive statistics are reported in Table 3. Computer Science is a particularly suitable field for our analysis because it features a large number of authors and publications, with both the mean and median team sizes being relatively high and the variance at a moderate level. This indicates that the field exhibits both widespread small-team collaborations and frequent large-team collaborations, providing a balanced setting for estimating the effects of participation in large-team projects.

## B Gender Classification Coverage

To assess the classifier's coverage, we analyzed the proportion of names that could not be classified across the studied cohorts, which fall into the "unknown" leaf of the taxonomy. As detailed in Table 4, the overall coverage for the total sample of 77,516 scholars is stably high, with an average of 6.97% of individuals having unknown names.

When we broaden the threshold for ambiguous authors to 0.3–0.7, the share of authors classified as unknown increases only to 12.7%. However, within this unknown group, the women-to-men ratio rises to 0.72 (approximately 3:4). When we further expand the threshold to 0.2–0.8, the unknown share grows slightly to 18.2%, while the women-to-men ratio remains significantly higher than the population ratio (approximately 1:2). These results suggest that adopting stricter thresholds offers limited gains in gender-classification accuracy, yet disproportionately excludes women authors. Moreover, because name-based gender inference is harder in certain linguistic or cultural contexts (e.g., Chinese names), stricter thresholds may also reduce the representativeness of authors from these groups. For these reasons, and following best practices recommended in prior research [51, 53], in the main analyses, we classify authors as unknown only when their prediction score equals 0.5 or when no prediction score is available. That said, we also perform a robustness check by excluding observations where the authors' classification certainty is lower than 80%, and then replicate the estimation of the Cox model (Section 5.3).

Next, a histogram of the probability distribution for the probability of being classified as 'gendered female' is plotted in Figure 8.

**Table 3: Descriptive Statistics of Academic Fields in SciSciNet**

Field	# authors	# publications	Team size (med)	Team size (m)	Team size (SD)
Physics	4,157,559	3,818,350	2	5.776	61.203
Geology	1,955,881	1,607,498	2	2.936	2.623
Engineering	4,834,252	3,690,898	2	2.272	1.933
Philosophy	761,352	917,958	1	1.186	1.233
Mathematics	3,528,933	3,724,147	2	2.170	3.106
Material Science	9,698,995	7,607,713	3	3.961	2.967
<b>Computer Science</b>	8,571,815	7,334,319	3	2.903	2.550
Economics	1,221,498	1,397,750	1	1.722	1.821
Environmental Sci.	2,890,840	1,790,789	3	3.457	3.197
History	1,007,958	1,230,142	1	1.210	1.287
Geography	1,847,217	1,208,279	2	2.557	4.138
Chemistry	14,406,174	10,819,704	3	3.785	2.609
Business	2,388,906	1,919,524	1	1.984	2.021
Political Science	1,934,629	2,005,642	1	1.482	2.101
Art	940,510	1,064,796	1	1.228	2.113
Medicine	31,516,175	23,048,415	3	4.145	5.971
Biology	13,586,695	10,252,991	3	4.258	4.463
Sociology	1,989,353	2,262,412	1	1.439	1.293
Psychology	4,761,400	4,439,462	2	2.642	2.508

**Table 4: Gender Classifier Coverage Across Studied Cohorts, Defined by the Percentage of Names Falling into the "Unknown" Category. The average proportion of "unknown" category is estimated at roughly 6.97%**

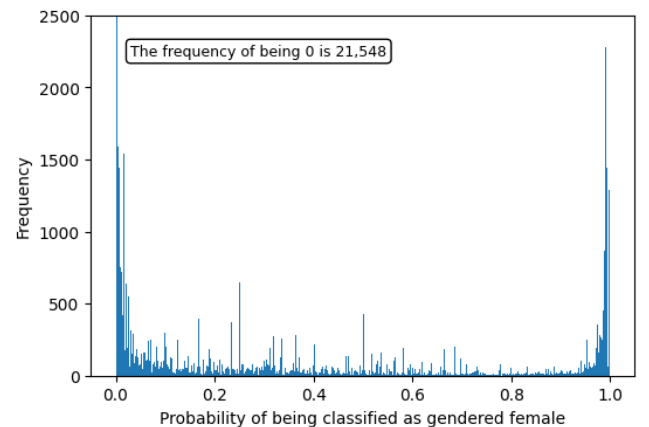
Cohort Year	Total	Unknown Scholars	Unknown Percentage
2000	12,499	932	7.46%
2001	13,384	884	6.60%
2002	15,339	1,080	7.04%
2003	17,069	1,178	6.90%
2004	19,228	1,332	6.93%
<b>Total</b>	<b>77,516</b>	<b>5,406</b>	<b>6.97%</b>

The figure exhibits a strong bimodal distribution, with peaks heavily concentrated at 0.0 and 1.0, while the sample size around 0.5 is notably small. This signifies that the vast majority of names in the sample possess a strong gender association according to the cultural consensus [50], which in part supports the efficacy and robustness of our gender classifier.

## C Detailed Results of Event Study and SDID

### C.1 Event Study

Before estimating our Synthetic Difference-in-Differences Event (SDIDE) models, we conduct a conventional event-study on the subsample of scholars who first enter a large-team collaboration in career year five, compared to a never-treated control group. This design serves two purposes. First, it provides a transparent diagnostic of the parallel-trends assumption: by inspecting pre-event coefficients, we can assess whether treated and control units follow comparable trajectories in the absence of treatment, given that no additional covariates are included in this study. Second, by augmenting the event-time indicators with gender interactions,

**Figure 8: Distribution of Consensus Gender Probability Estimates for all Classified Names in the Studied Cohort**

we obtain an initial view of whether treatment dynamics differ between men and women researchers. We emphasize that this step is descriptive and diagnostic; the SDID analysis that follows leverages optimized weighting to improve balance in pre-event paths and yields our main causal estimates.

Let  $O_{it}$  denote the outcome (e.g., annual publication counts) for scholar  $i$  in year  $t$ . Define the event time relative to the first large-team collaboration as  $k = t - T_i$ , where  $T_i$  is the treatment year ( $T_i$  = cohort year+5 for treated units,  $T_i$  = 9999 for never-treated units).

We create event-time dummies for  $k \in \{-5, \dots, 4\} \subset \mathbb{Z}$ . Our estimating equation is:

$$O_{it} = \alpha + \sum_k \left( \beta_k D_{it}^k + \gamma_k D_{it}^k \times \text{Female}_i \right) + \delta \text{Female}_i + \lambda_t + \varepsilon_{it},$$

where  $D_{it}^k = 1\{t - T_i = k\}$  indicates event time,  $\text{Female}_i$  is a gender dummy, and  $\lambda_t$  are year fixed effects.

Results are reported in Figure 9, which show that although the gap between the treated and control groups in publication counts and centrality widens and remains at a higher level after the event, significant pre-treatment differences persist between the two groups, breaking the parallel trends assumption of DID model.

Table 5 present the estimates of effects of large-team collaboration on network centrality and publication counts, together with interaction with gender (the coefficients of yearly fixed effects and gender are omitted in the table to save space).

## C.2 SDIDE by Treatment Year for Cohort 2004

For each treatment group by treatment year, we plot the trend differences between the SDID-synthesized treated and control groups. Figures 10 report the results for cohort 2004 (plots for the other four cohorts can be found in the Supplementary file). As shown, the SDID procedure successfully constructs a credible counterfactual control group for each treated group, with pronounced upward divergences emerging after treatment.

## D Robustness Checks

We conducted four sets of robustness checks. First, we varied the definition of the event (first-time large-team collaboration) in three ways: (a) the 90th percentile of team sizes within a given year (event1), (b) the 90th percentile of the largest team size across the first ten years of an early-career author (event2), and (c) a fixed threshold of eight coauthors (event3). We then re-estimated the SDIDE model on centrality and publication count separately for the five cohorts. The results, reported in Table 6, show that all estimates remain positive and statistically significant, supporting our main finding that participation in large-team collaborations substantially enhances both network centrality and research productivity for early-career scholars. It is worth noting that in some years the alternative event definitions may coincide, in which case SDIDE produces identical point estimates. However, due to the uncertainty introduced by the bootstrapping procedure, the standard errors may still differ.

Second, we exclude the top 5% of sample scholars ranked by the total number of publications produced during the observation period (the first ten years of their careers), and re-estimated the

SDIDE model on centrality and publication count. The results are reported in Table 6 (Column "Restricted"), showing that after excluding extremely high-influential samples, large-team collaboration is still able to improve structural position and academic productivity.

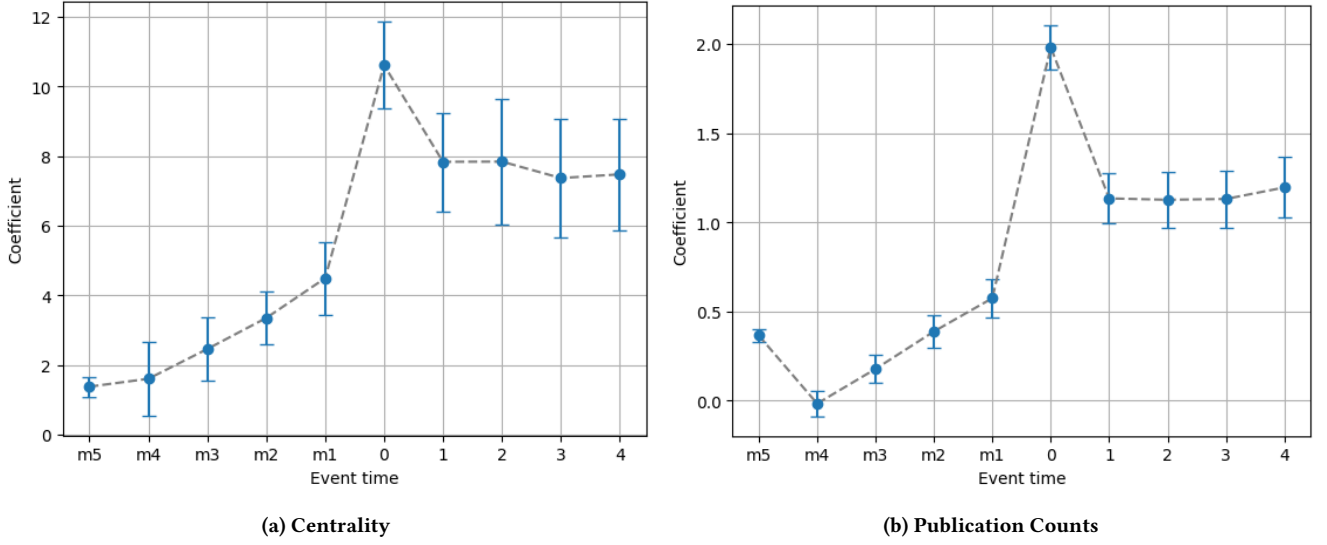
Third, we replaced the original outcome variables, which were based on annual collaboration networks, with centrality and publication counts in cumulative collaboration networks. We do not calculate the cumulative sum of RCI, as the citation count itself is already accumulated since publication. The intuition behind this replacement is that, in academia, a junior scholar is typically evaluated by their overall profile rather than their collaborations in just the past year. Thus, cumulative measures provide a more accurate reflection of the benefits of large-team collaborations for early-career researchers. We further tested the robustness of our results by using two alternative productivity measures: yearly publication counts weighted by venue prestige, and the yearly average RCI, which abstracts from output volume. The results of this robustness check are reported in the supplementary file.

Fourth, we estimate SDID models for centrality, publication count, and RCI-weighted publication count within three computer science subfields: Human-Computer Interaction (HCI), Data Science (DS), and Computer Security (CSec). These subfields were selected based on their distinct collaboration patterns. It is important to note that SciSciNet does not provide a complete hierarchical field-subfield structure or consistent subfield labels for all publications. As a result, we classify an author into a subfield only if they have a sufficiently large share of publications in that subfield (defined as above the average share among all authors with at least one publication in that subfield within the cohort). This procedure substantially reduces the sample size available for subfield analyses.

Figure 11 illustrates the differences in team-size dynamics across the three subfields between 2000 and 2019. HCI exhibits relatively large teams on average, whereas CSec has consistently smaller teams. DS shows moderate average team sizes but experiences a rapid increase, along with high variance, likely due to the stronger involvement of industry-affiliated research groups, which typically form larger teams. They also have different gender compositions: the ratio of women researchers is 25.8% for HCI, 24.1% for DS and 18.9% for CSec. Despite these differences, all three subfields display a rising trend in collaboration size, suggesting a broader shift within computer science toward larger-team collaborations.

The subfield-specific results are presented in the supplementary file. All SDID models return significant post-event treatment effects, indicating that the positive impact of first-time large-team collaboration persists across heterogeneous collaboration cultures. HCI exhibits strong and stable treatment effects across all three outcomes. In DS, the effects on centrality and raw publication count are comparatively weaker, but the effect on RCI-weighted publication count is the strongest, possibly reflecting the field's closer ties to industry and its emphasis on high-impact, rapidly disseminated research [82].

The survival analyses show that in all three subfields, women are consistently less likely than men to enter large-team collaborations. The effect is strongest in CSec, where women's probability of entering a large team is 26% lower than that of men.



**Figure 9: Dynamic Effects of First-time Large-team Collaboration on Centrality and Publication Counts Among Individuals Treated at Career Year 6**

**Table 5: Event-study Coefficients with Gender Interactions on Two Outcomes, with coefficients significant at  $p < 0.001$  in bold**

Variable	Centrality			Publication Counts		
	Coef	Std. Err.	p-value	Coef	Std. Err.	p-value
event_time_m5	<b>1.3754</b>	0.145	0.000	<b>0.3656</b>	0.018	0.000
event_time_m4	<b>1.6015</b>	0.542	0.003	-0.0161	0.037	0.661
event_time_m3	<b>2.4569</b>	0.462	0.000	<b>0.1779</b>	0.041	0.000
event_time_m2	<b>3.3477</b>	0.387	0.000	<b>0.3881</b>	0.046	0.000
event_time_m1	<b>4.4901</b>	0.536	0.000	<b>0.5744</b>	0.054	0.000
event_time_0	<b>10.6097</b>	0.641	0.000	<b>1.9807</b>	0.063	0.000
event_time_1	<b>7.8339</b>	0.723	0.000	<b>1.1336</b>	0.072	0.000
event_time_2	<b>7.8390</b>	0.917	0.000	<b>1.1261</b>	0.080	0.000
event_time_3	<b>7.3667</b>	0.868	0.000	<b>1.1306</b>	0.081	0.000
event_time_4	<b>7.4735</b>	0.816	0.000	<b>1.1950</b>	0.086	0.000
event_time_m5_female	0.3979	0.341	0.243	0.0855	0.044	0.054
event_time_m4_female	-0.0397	0.804	0.961	0.0670	0.076	0.375
event_time_m3_female	-1.3017	0.635	0.040	-0.0607	0.082	0.459
event_time_m2_female	-0.6297	0.805	0.434	-0.0846	0.098	0.388
event_time_m1_female	-0.0938	0.949	0.921	0.0496	0.118	0.674
event_time_0_female	1.6849	1.824	0.356	0.0497	0.159	0.755
event_time_1_female	1.1555	1.967	0.557	0.1104	0.198	0.578
event_time_2_female	2.1181	2.089	0.311	0.2601	0.207	0.208
event_time_3_female	1.7811	1.890	0.346	0.2623	0.204	0.198
event_time_4_female	1.5162	2.173	0.485	0.0036	0.197	0.985

## E Evidence on Network Expansion through Large-team Collaborations

In order to capture the dynamics of scholars' network expansion, we adopt the Jaccard Similarity metric, which measures the extent to which current collaborators overlap with the cumulative two-hop neighborhood from previous years:

$$J_t = \frac{|N_t^{(1)} \cap (\bigcup_{\tau < t} N_t^{(2)})|}{|N_t^{(1)}|}$$

where  $N_t^{(k)}$  denotes the  $k$ -hop neighborhood in year  $t$ . We focus on two-hop neighborhoods because they provide a proxy for a scholar's research group environment, allowing us to assess

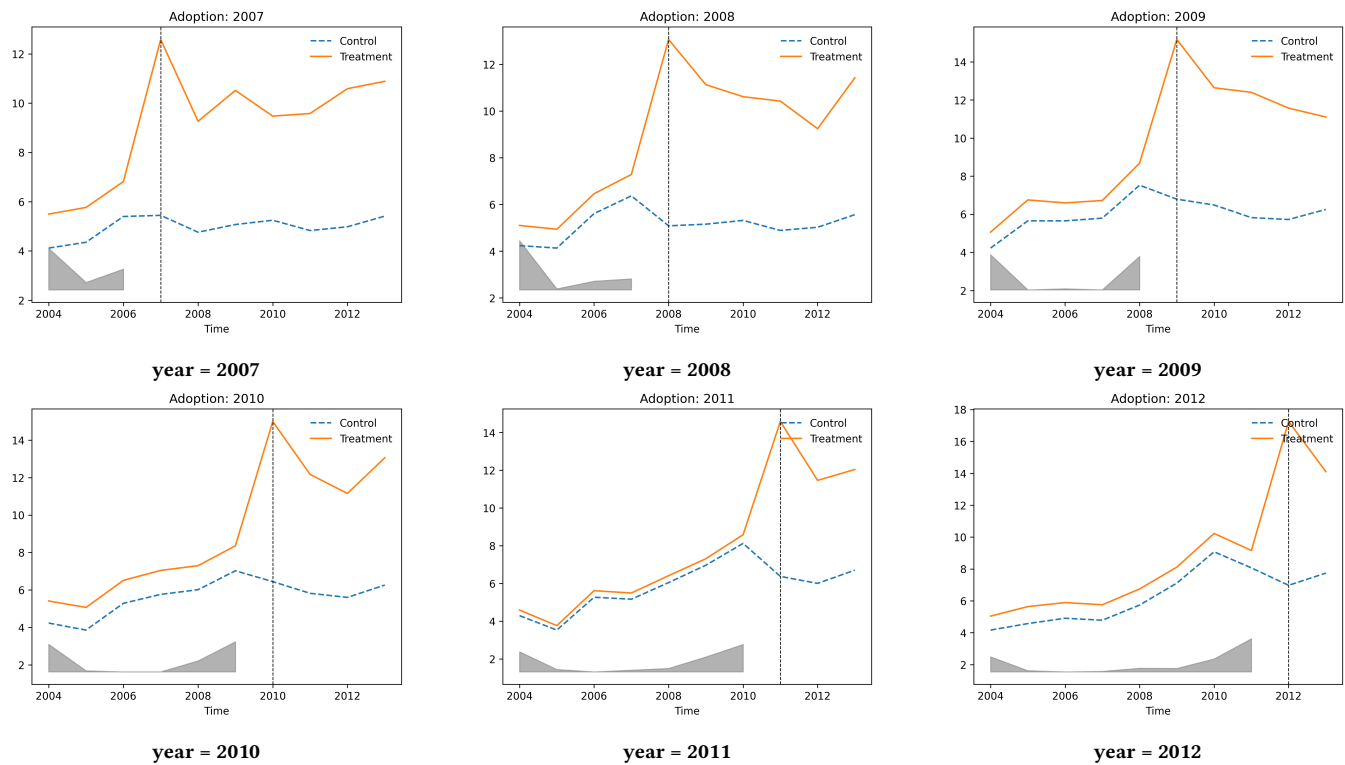


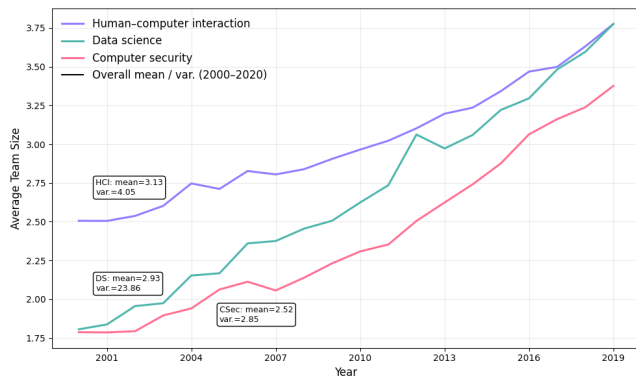
Figure 10: SDIDE of Network Centrality by Treatment Year of Cohort 2004

Table 6: SDID Estimated Effects for Alternative Event Definitions

	Centrality				Publication Counts			
	Event1	Event2	Event3	Restricted	Event1	Event2	Event3	Restricted
<b>Cohort 2000</b>								
Treatment Effect	5.632	3.849	3.706	2.779	1.043	0.756	0.684	0.619
Std. Err	(0.179)	(0.313)	(0.324)	(0.118)	(0.028)	(0.045)	(0.036)	(0.030)
<b>Cohort 2001</b>								
Treatment Effect	6.299	4.782	4.722	2.665	1.088	0.824	0.775	0.611
Std. Err	(0.161)	(0.215)	(0.273)	(0.168)	(0.027)	(0.036)	(0.047)	(0.023)
<b>Cohort 2002</b>								
Treatment Effect	6.544	5.232	4.616	3.132	1.070	0.852	0.749	0.695
Std. Err	(0.158)	(0.219)	(0.281)	(0.141)	(0.021)	(0.032)	(0.050)	(0.021)
<b>Cohort 2003</b>								
Treatment Effect	7.114	3.659	3.659	3.034	1.165	0.745	0.745	0.666
Std. Err	(0.171)	(0.358)	(0.231)	(0.179)	(0.026)	(0.034)	(0.037)	(0.022)
<b>Cohort 2004</b>								
Treatment Effect	5.782	3.989	3.989	3.090	1.075	0.757	0.757	0.684
Std. Err	(0.220)	(0.294)	(0.236)	(0.168)	(0.026)	(0.034)	(0.034)	(0.020)

whether the current group facilitates the formation of new collaborations, or how much the scholars depend on their current research environment.

The results of treatment group are reported in Figure 12, showing that both men and women scholars exhibit a sharp drop in similarity in the event year, followed by a rebound and stabilization at higher levels. In the year when first-time large-team collaboration occurs,

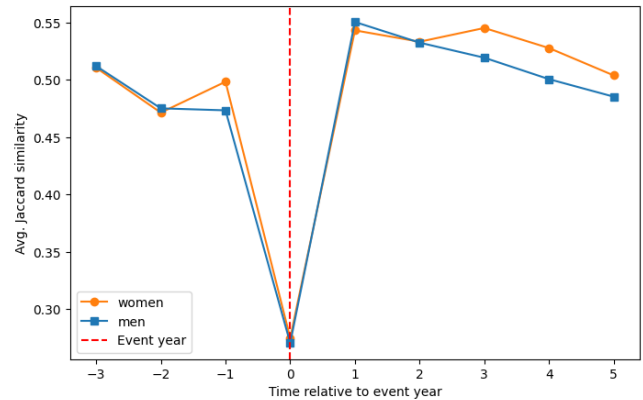


**Figure 11: Trends in average team size from 2000 to 2019 across three subfields: Human-Computer Interaction (HCI), Data Science (DS), and Computer Security (CSec). HCI exhibits the largest teams throughout the period, CSec maintains consistently smaller teams, and DS shows moderate team sizes with a sharp rise and large variability in the later years. All three subfields demonstrate a clear upward trend.**

the average similarity between the set of new collaborators and the set of previous 2-hop/1-hop neighbors is 0.2745 for women, and 0.2664 for men, and the medians are 0.2 and 0.1818, respectively, indicating a low degree of overlapping. In fact, there are 26% men and women with a similarity of 0 in that event year, which means that the authors are likely to find a new research group when

participating in large teams. For comparison, when they publish in the years after the event year, there are only 10.97% women-year and 10.41% men-year have the similarity of 0.

For the control group, the average proportion of similarity of 0 is 30.68% throughout all the years, indicating that their collaboration relationship is not as stable as the treatment units.



**Figure 12: Jaccard Similarity Dynamics around First-time Large-team Collaboration for Treatment Group. The two lines represent men (square markers) and women (circle markers).**