Research Square

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

Do Women Benefit More Than Men fromCoauthoring with a Prominent Researcher?

Daniela Aguirre-Guerrero
Universidad Autónoma Metropolitana
M. Ehekatzin Garcia-Guzman
College of Mexico
Aurora A. Ramirez-Alvarez
College of Mexico
Diana Terrazas-Santamaria
dterrazas@colmex.mx
College of Mexico

Research Article

Keywords: Gender inequality, coauthorship network, h-index, research impact

Posted Date: March 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9035498/v1

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

Additional Declarations: No competing interests reported.

Do Women Benefit More Than Men from Coauthoring with a Prominent Researcher?

Daniela Aguirre-Guerrero*

M. Ehekatzin García-Guzmán†

Aurora A. Ramírez-Álvarez‡

Diana Terrazas-Santamaría§

February 2026

Abstract: In this paper, we investigate whether the careers of junior researchers show a differentiated gender effect of collaboration with a prominent researcher within five years of their first publication. We use a sample from the Mexican National System of Researchers, and a staggered difference-in-differences (DID) imputation approach proposed by Borusyak et al. (2024) to compare similar junior collaborators and non-collaborators. We find positive effects on impact (total citations, citations per paper, and the h-index) and on

*Universidad Autónoma Metropolitana, Unidad Cuajimalpa, Department of Applied Mathematics and Systems. Current address: Avenida Vasco de Quiroga 4871, Santa Fe Cuajimalpa, Cuajimalpa, Mexico City 05348, Mexico. ORCID ID: 0000-0003-4945-7258. daguirre@cua.uam.mx.

†El Colegio de México, Center for Economic Studies. Carretera Picacho Ajusco 20, Ampliación Fuentes del Pedregal, Tlalpan, Mexico City 14110, Mexico. marcgarcia@colmex.mx.

‡El Colegio de México, Center for Economic Studies. Carretera Picacho Ajusco 20, Ampliación Fuentes del Pedregal, Tlalpan, Mexico City 14110, Mexico. ORCID ID: 0000-0001-8358-2920. aurora.ramirez@colmex.mx.

§El Colegio de México, Center for Economic Studies. Carretera Picacho Ajusco 20, Ampliación Fuentes del Pedregal, Tlalpan, Mexico City 14110, Mexico. ORCID ID: 0000-0002-7236-882X. dterrazas@colmex.mx.

the core number of the coauthorship network (a measure of the quality of connections), although we find no statistically significant differences between male and female researchers. Our results also suggest that the observed effects in impact and network measures may arise from a single "big hit" rather than a sustainable increase in the productivity and impact of treated researchers when considering only the focal paper.

Keywords: Gender inequality, coauthorship network, h-index, research impact

JEL codes: J16, J24, L14, O31, O32.

1 Introduction

Gender disparities in academia remain persistent and well-documented. Women continue to be underrepresented across most scientific disciplines (Huang et al., 2020), they face systematic gaps not only in participation but also in research productivity and in the recognition and influence of their work (Huang et al., 2020; Larivière et al., 2013; Beaudry and Larivière, 2016), as reflected in impact indicators (Huang et al., 2020; Dorantes-Gilardi et al., 2022; Beaudry and Larivière, 2016). Despite extensive evidence of these inequalities, much less is known about the mechanisms that could help narrow them. In particular, relatively few studies have examined how access to resources, mentorship, and opportunities for collaboration may shape the long-run impact of scientific careers—especially for women (Beaudry and Larivière, 2016).

Collaboration lies as a key component of everyday activity in academia, both formal and informal. In particular, early-career collaborations can provide junior researchers with exposure to high-quality networks, access to resources, and visibility within the scientific community. In this manner, coauthoring with established and influential scholars may accelerate productivity and impact. Yet, an open question remains: do such collaborations generate lasting benefits, and are these benefits distributed equally between genders?

This paper addresses two main questions regarding researchers in engineering, economics, and the health sciences in Mexico. First, we examine whether coauthorship with a prominent researcher in the early stages of a researcher's career affects subsequent productivity, research impact, and network connections. Second, we investigate whether the effect of such early collaboration differs systematically between women and men.

Our interest lies specifically in collaboration with prominent researchers. To capture prominence, we use the h-index, defined as the maximum value of h such that an author has published at least h papers that have each been cited at least h times. The h-index has become one of the most widely used measures of sustained scientific influence. As Bornmann and Daniel (2007) argue, it rewards

researchers who publish continuously throughout their careers with consistent above-average impact and is not driven solely by a small number of highly cited "big hits." For this reason, it provides a meaningful proxy for prominence and seniority within academic collaboration networks.

To address these questions, we construct a dataset that combines Scopus metadata on scientific papers and citation counts for researchers in the Mexican National System of Researchers (SNI; Sistema Nacional de Investigadores) and their coauthors. We rely on the 2021 SNI registry[1] and focus on three disciplines: engineering, economics, and health sciences. This setting allows us to examine collaboration patterns in a middle-income country context, spanning both STEM and non-STEM research disciplines.

We estimate the causal effect of collaboration with a prominent researcher on a researcher's productivity, impact, and network connections using the staggered difference-in-differences (DID) imputation approach proposed by Borusyak et al. (2024). Our treatment is defined as collaboration with a prominent researcher within the first five years of a researcher's career; specifically, one whose h-index lies in the top 2% of our sample's coauthor distribution. We compare treated researchers to the same, not-yet-treated comparison group, and allow event-time effects to vary by gender.

Over the first decade after the event, collaboration with a prominent coauthor increases citations by approximately 77% for female researchers and by 46% for male researchers. For women, we find no effect on publication output and only a transitory increase in the h-index, while for men, neither publications nor the h-index change significantly. Collaboration also increases the embeddedness of the researchers' coauthorship network: the estimated effect on the core number is about 15 units for women and roughly four units for men over the subsequent decade. Interestingly, differences across genders are not statistically significant in all outcomes.

A crucial question that stands out is whether the positive and significant effects reflect structural changes in treated researchers' trajectories or instead

[1] `https://conacyt.mx/sistema-nacional-de-investigadores/archivo-historico/`.

represent a temporary boost driven by the focal collaboration —the paper in which the treated author collaborated with a prominent researcher within the first five years of their career. To address this, we estimate effects on citations, the h-index, and the core number excluding the focal paper. The results suggest that the focal paper entirely accounts for the positive effects observed in citations, h-index, and core number.

Our study relates closely to two strands of prior work. Li et al. (2019) show that in four scientific disciplines, early coauthorship with established, highly cited researchers positively affects junior researchers' careers in terms of citations and the probability of becoming highly cited scientists themselves. Dorantes-Gilardi et al. (2023) investigate gender differences among junior economists who collaborate with highly cited authors early in their careers, finding positive effects on impact but no statistically significant differences between men and women. Nevertheless, our analysis exploits the longitudinal nature of our data to implement an event-study approach that traces the evolution of effects over time, and we provide evidence across three disciplines (engineering, economics, and health sciences), STEM and non-STEM.

We are cautious about extrapolating to disciplines with markedly different collaboration or citation norms. Although our analysis focuses on Mexican researchers and three disciplines, it sheds light on how early access to prominent networks can influence long-run inequality in scientific careers, a common concern across research communities.

The remainder of the paper is organized as follows. Section 1 introduces our hypotheses and data construction and description. Section 2 details the construction of these metrics. Section 3 describes the empirical strategy, followed by a section presenting the main results and an illustrative case study. Section 5 provides robustness checks, including estimations excluding the focal paper, and we conclude with final remarks.

2 Hypotheses and data

Our empirical analysis is guided by testable hypotheses regarding the effects of early-career collaboration with prominent researchers. Such collaborations may increase scholarly impact, reshape researchers' positions within the coauthorship network, and differ by gender. At the same time, observed effects may reflect the influence of the focal collaboration itself rather than persistent changes in researchers' broader career trajectories. We consider the following hypotheses:

- *Returns to early prominent collaboration are heterogeneous by gender.*
- *Prominent collaboration affects research impact but not research output.*
- *Prominent collaboration increases network embeddedness, as captured by the core number.*
- *Estimated effects are largely attributable to the focal paper.*

We expect gender heterogeneity because early collaboration with a prominent researcher may provide a stronger credibility signal for women than for men. If women face greater uncertainty about the perceived quality or visibility of their work, such collaborations can more effectively reduce uncertainty for editors, referees, and potential collaborators. Regarding the second hypothesis, early collaboration primarily affects a paper's attention and visibility, whereas the number of publications is constrained by time or project pipelines. Third, collaborating with a prominent scholar often implies entry into a densely connected collaboration cluster, generating additional links with already interconnected researchers and diverse groups. Finally, a large share of the estimated impact may be attributable to the collaborative paper itself, rather than to positive effects on the overall portfolio.

To test these hypotheses, we build a dataset using Scopus metadata of papers and citation counts for SNI researchers and their coauthors.[2] We apply

[2]The data was retrieved from the Scopus database via the Scopus API using the *pybliometerics* library for Python Rose and Kitchin (2019).

techniques for author name disambiguation to find researchers from the 2021 SNI registry in the Scopus database.

We use the 2021 SNI registry, and, due to data-intensity requirements, we focus on three disciplines: engineering, economics, and the health sciences.[3] The percentages of SNI researchers found in the Scopus database are 55% for engineering, 50% for economics, and 68% for health sciences. Each of these disciplines has around a thousand SNI researchers, and the standard practice is to publish their work in indexed journals likely to appear in Scopus. Engineering and health typically involve larger, more hierarchical teams, while economics features smaller teams and different authorship conventions. This mix situates our analysis at the intersection of STEM and non-STEM practices.

It is worth noting that the SNI was created in 1984 as a voluntary and merit-based system to stabilize academic careers and stimulate scientific productivity following a period of severe economic crisis in Mexico, and it is organized into disciplinary evaluation areas, each governed by specific peer-review committees and criteria (Moreno, 2019; Gil Antón and Contreras Gómez, 2017). Entry and continued membership depend on periodic peer evaluation, in which researchers must demonstrate an active scholarly trajectory; SNI membership carries both symbolic recognition and a significant non-salary stipend, making it a central determinant of prestige, hiring, and promotion in Mexican academia (Moreno, 2019).

The fact that only around 50% of our SNI members appear in Scopus reflects the broad portfolio of academic outputs recognized by the SNI in both national and international outlets —including peer-reviewed journal articles, books, book chapters, conference proceedings, patents, graduate thesis supervision, and teaching. This pattern is especially pronounced in the social sciences and humanities, where, for instance, books and book chapters remain central forms of scholarly production and can be underrepresented in international indexing systems such as Scopus. In this way, researchers in the SNI can

[3]We thus have data only for researchers active in this year and not on those who retired before 2021.

exhibit heterogeneous trajectories: some follow international indexing paths, while others maintain nationally oriented academic influence, even though both may advance similarly within the system. (Moreno, 2019; Gil Antón and Contreras Gómez, 2017).

Our data include 13,634 researchers, of whom 19% are SNI researchers and 81% are their coauthors. For each SNI researcher, we retrieve metadata for all publications through 2021,[4] as well as citation counts by year. Table 1 reports descriptive statistics for female and male SNI researchers within each discipline.

Because not all authors in the SNI registry can be matched to Scopus, the set of SNI researchers observed in our bibliometric data may differ from the full SNI population. We therefore assess whether Scopus matching is systematically related to gender, which is central to our analysis of gender gaps. Specifically, we test for imbalance in the proportion of female researchers among those found in Scopus relative to the SNI registry. The results suggest that the female share in our matched data is not statistically different from the female share in the SNI registry in any of the three fields (Table B.1).[5] This representativeness check is conducted on the full matched SNI–Scopus sample, since it is designed to evaluate potential gender-related selection into Scopus coverage. Our empirical analysis applies an additional restriction to researchers whose publishing careers start in 2005 or later to improve cohort comparability.

Female researchers in all disciplines tend to be younger, have published fewer papers, receive fewer citations, and have fewer coauthors (except in economics) than their male counterparts (Table 1). These data allow us to compute the yearly evolution of an SNI researcher's h-index and track annual publications, citations, and coauthorship outcomes.

[4]This metadata includes bibliographic, author, and funding agency information (see `https://dev.elsevier.com/sc_search_views.html`).

[5]The regression combines authors in our matched sample and in the SNI registry and regresses an indicator for being female on an indicator for being found in Scopus.

Table 1: Descriptive Statistics by Gender

	Economics			Engineering			Health Sciences		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
A) Mean									
Citations	5.39	5.94	5.57	8.97	8.56	8.88	24.89	22.54	23.53
Publucations	1.08	0.95	1.04	1.27	1.02	1.22	1.80	1.57	1.67
Coauthors	2.20	2.27	2.23	3.95	3.65	3.88	11.88	10.69	11.19
h-Index	1.42	1.39	1.41	1.97	1.93	1.96	3.74	3.46	3.58
Core number	3.63	4.51	3.93	6.74	7.58	6.93	16.42	16.23	16.31
Career length	10.01	9.54	9.85	9.33	8.12	9.06	11.45	11.07	11.23
B) Standard deviation									
Citations	15.86	15.60	15.77	17.23	13.52	16.49	53.48	53.73	53.64
Publucations	1.83	1.54	1.74	1.62	1.28	1.56	2.22	2.13	2.17
Coauthors	4.65	6.16	5.22	4.99	4.47	4.88	18.59	19.69	19.24
h-Index	1.68	1.62	1.66	1.94	1.63	1.87	3.45	3.17	3.29
Core number	4.76	7.89	6.02	3.76	4.32	3.91	14.14	15.92	15.19
Career length	3.80	3.86	3.82	4.08	3.29	3.95	3.16	3.33	3.27
C) Maximum value									
Citations	358.00	236.00	358.00	211.00	139.00	211.00	908.00	1347.00	1347.00
Publucations	28.00	18.00	28.00	19.00	14.00	19.00	29.00	26.00	29.00
Coauthors	93.00	135.00	135.00	47.00	29.00	47.00	289.00	276.00	289.00
h-Index	19.00	13.00	19.00	17.00	10.00	17.00	23.00	20.00	23.00
Core number	92.00	99.00	99.00	25.00	30.00	30.00	112.00	165.00	165.00
Career length	16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
D) Number of researchers	343.00	185.00	528.00	396.00	120.00	516.00	215.00	308.00	523.00

Note: We show the data from active researchers who started their career between 2005 and 2021. The number of SNI researchers analyzed is 1,567.

Importantly, because we have information on each coauthor, we can calculate their h-index, number of papers, and citations for each year and incorporate these into the econometric analysis. The database characteristics allow us to use the core number as a descriptive measure of a researcher's embeddedness in cohesive collaboration structures.

2.1 H-index As a Measure of Impact

Hirsch (2005) proposes the Hirsch index (h-index) to quantify scientific output by combining publications and citations into a single measure. A researcher with N_p published papers has index h if h of those papers have at least h citations each, while the remaining $(N_p - h)$ papers have no more than h citations. The h-index is widely used due to its simple computation and transparent interpretation.

By construction, the h-index is correlated with career length, as publications and citations accumulate over time (Costas and Bordons, 2007). It cannot be driven solely by a single highly cited paper or by a large number of minimally cited publications. Although the h-index reduces a researcher's career to a

single statistic and has well-known limitations (see, e.g., Costas and Bordons, 2007; Bartneck and Kokkelmans, 2011; Koltun and Hafner, 2021), we use it as a proxy for prominence, as it reflects sustained output and broad citation impact. Consistent publishing and a substantial share of moderately cited work are required, implying higher values among senior than junior researchers.

Researchers in the upper tail of the h-index distribution are typically highly visible within their disciplines. For instance, Scopus reports an h-index of 41 for both Claudia D. Goldin, the 2023 Nobel Laureate in Economics (9,000 citations; 71 papers),[6] and Avinash K. Dixit (13,771 citations; 108 papers).[7] Despite differences in citation totals and publication counts, both are widely regarded as leading contributors to the field.

We therefore define "prominent" researchers as those in the upper tail of the h-index distribution, rather than those with modestly above-average impact. Our empirical strategy focuses on coauthors in the very top of the field- and year-specific h-index distribution.

Collaboration with prominent colleagues may benefit junior researchers through several channels. Krauss et al. (2023) show that young researchers who later became prominent—defined as the top 100 h-index holders within a discipline—gained substantially from early-career collaborations with prominent coauthors: papers written with prominent researchers received more than twice as many citations as those coauthored with non-prominent scholars. Access to resources such as knowledge, information, funding, or equipment may be particularly valuable early in a career. Ebadi and Schiffauerova (2015) document that stronger funding improves scientific performance, while network position within dense and cohesive coauthorship communities increases future funding prospects (Jaramillo et al., 2023).

More broadly, Li et al. (2022) argue that coauthorship networks operate as social capital, shaping access to collaborators, information, and visibility. Differences in network position may therefore contribute to disparities in promi-

[6] https://www.scopus.com/authid/detail.uri?authorId=6701799144

[7] https://www.scopus.com/authid/detail.uri?authorId=55824766400

nence and career outcomes, including gender gaps and barriers to elite research environments.

2.2 Core Number As a Measure of Cohesive Structures

Network measures such as degree and core are outcomes that describe how the SNI early researcher becomes embedded in collaboration structures over time. Coauthorship networks have become a standard model for studying academic collaboration and connectivity dynamics in academic research (Newman, 2004). To capture these dynamics, for each research and year t, we construct a cumulative coauthorship network using all Scopus-indexed publications in that field through year t. Each network includes a SNI focal researcher and all coauthors observed in the data. Edges are weighted by the number of jointly authored papers, and these weights are used to compute the degree and core number for each focal researcher. We rely on cumulative rather than rolling-window networks to capture longer-run integration into collaboration structures.

The core number of a node is defined as the largest integer k such that the node belongs to the k-core of the network, i.e., a maximal subgraph in which every node has degree at least k. Intuitively, a small core number indicates that an author lies in the periphery of the collaboration network, whereas a large core number reflects participation in more densely connected and cohesive groups. An increase in the core number may arise when an author forms ties with collaborators who are themselves highly interconnected, thereby moving into more cohesive collaboration clusters. The construction of the coauthorship network and the computation of the core number are provided in Appendix A.

We do not interpret the core number as a proxy for "prominence" or "quality." Instead, it summarizes the extent to which an author is positioned within cohesive, highly interconnected regions of the coauthorship network. We interpret this measure within each discipline, since fields differ systematically in typical team sizes; accordingly, a core number that is high relative to an area's baseline is more informative.

Degree (the number of distinct coauthors) and the core number capture different dimensions of a coauthorship network and may therefore evolve differently over time; thus, we view them as complementary measures. At the same time, because a researcher may attain a high core number mechanically by participating in a single paper with many coauthors, we include the number of coauthors as a control variable.

Consistent with this interpretation, Li et al. (2022); Jaramillo et al. (2023) find that observed gender differences in scholarly productivity and citations are likely related to differences in the structure and composition of researchers' coauthorship networks and their k-core decomposition. Also, as noted by Malliaros et al. (2020), belonging to a deeper k-core indicates that a researcher is part of a subset of the network in which collaborators are themselves well-connected. We view this structural embeddedness as a potential source of access to information flows and repeated opportunities for collaboration, which, in turn, can shape academic careers. In this sense, increases in a researcher's core number may signal deeper integration into scientific networks, where recognition, information, and broader opportunities tend to circulate more rapidly.

3 Empirical strategy

We estimate the causal effect of collaboration with a prominent researcher on the impact, productivity, and collaboration trends of Mexican researchers, using the staggered DID imputation approach proposed by Borusyak et al. (2024). Recent literature has found that the staggered DID estimator exhibits significant biases due to group and temporal heterogeneity in treatment effects. To avoid these biases, we follow Borusyak et al. (2024) to estimate treatment effects. Their approach consists of estimating a regression of the outcome on the untreated group (i.e., never-treated or not-yet-treated) only, and then applying the predicted outputs to the treated observations to calculate the treatment effect as a weighted sum of these individual effects.

In our empirical implementation, we restrict the imputation sample to re-

searchers who have not yet been treated and exclude never-treated researchers from the estimation. This ensures that counterfactual outcomes for treated authors are constructed using researchers who will eventually be treated but have not yet had an early collaboration with a prominent coauthor, thereby reducing heterogeneity between treated and comparison groups. This restriction strengthens the research design by improving comparability: treated researchers are benchmarked against peers who are still at risk of treatment, rather than against never-treated researchers who may differ systematically.

In our setting, excluding never-treated researchers also reduces cohort composition differences (e.g., very senior researchers with mechanically larger citation stocks), so the identifying variation is driven more cleanly by differences in the timing of early coauthorship with a top-2% coauthor. When we report gender heterogeneity, it should be interpreted within this same timing-based comparison. Our gender results compare women and men among researchers who eventually reach a collaboration with a top coauthor, distinguishing those who do so within the first five years of their publishing career (treated) from those who do so only later (not-yet-treated at a given event time). They do not answer what would happen if women had the same chances as men to obtain such early collaborations in the first place.

The event is defined as a collaboration with a prominent researcher within the first five years of their career, which we refer to as the early stage of a researcher's career. This definition of the early stage is close to that of Bazeley (2003), who defines it as the five years after completing the Ph.D., and it is the same as that of Jin et al. (2020) and Krauss et al. (2023), who define it as the first five years since the first publication. We limit the event to the first five years of the academic career because, as previous literature has pointed out, the conditions in this stage correlate with researchers' future impact and prominence in academia. It is likely that the effects of collaborating with a prominent researcher are more pronounced when the collaboration occurs at an early stage of their academic career (Bazeley, 2003; Krauss et al., 2023). Additionally, we use researchers who collaborate with a prominent researcher in

a more advanced stage of their academic career as the control group, as these researchers are more likely to be similar to the treated group.

We define a prominent researcher based on the maximum coauthors' h-index distribution by year and discipline in our sample. Concretely, for each discipline and calendar year, we compute the distribution of the maximum h-index among all coauthors linked to researchers in our sample. We then define as "prominent" those coauthors whose h-index lies above the 98th percentile of this field- and year-specific distribution. A junior researcher is treated if, within the first five years of their publishing career, they coauthor at least one paper with a top-2% coauthor. In our data, this threshold choice implies, for example, that in economics in 2005, the 90th percentile of the coauthor h-index distribution is 7, whereas the 98th percentile is 26, indicating that our definition of prominence corresponds to a much more elite group of coauthors.

The number of junior researchers who experience this type of early collaboration is, therefore, small by construction (see Table B.2), and our estimates should be interpreted as speaking to this specific opportunity margin rather than to all forms of early coauthorship. Accordingly, tables report sample sizes in different units depending on the analysis (unique treated researchers versus researcher-year observations).

Since the event is defined by the first early coauthorship with a top-2% coauthor regardless of team size, estimated effects should be interpreted as averages over the range of collaboration formats observed in the data rather than as the impact of a single, uniform type of collaboration. This approach enables us to ensure that the estimated effects result from collaborating with a researcher with greater seniority and impact due to the very nature of the h-index.[8] Our baseline estimations define a prominent researcher as a coauthor in the top 2% of our sample's maximum h-index distribution for coauthors by area.[9]

[8]Figure B.1 displays the h-index of treated authors and prominent researchers. Because all observations are above the 45° line, we can be sure that all treated authors have a smaller h-index than the prominent researchers who caused the event.

[9]Hirsch (2005) explains that the h-index distribution is inherent to each discipline because of differences in productivity and citation patterns among disciplines.

The outcomes we measure are the log of the total number of citations, the log of the number of citations per paper, the log of the number of papers, the log of the h-index[10], and the core number, using the following specification:

$$Y_{i,t} = \alpha_i + \gamma_t + q'_{i,t}\psi + \sum_{m=-G}^{M} \tau_m T_{i,t-m} + \varepsilon_{i,t} \tag{1}$$

where Y represents the outcome variable of researcher i in year t, q is a vector of time-varying controls, α_i and γ_t represent individual and year fixed effects, respectively, and T is an indicator variable equal to 1 if the researcher i is treated in time t and otherwise equal to 0. G and M represent the window in which the effect is computed, G denoting the years before the event and M denoting the years after the event. In this study, we set $G = 5$ and $M = 10$. Finally, τ is the coefficient of interest and denotes the average treatment on the treated effect (ATT) in the year t after the event. Because a researcher's primary discipline is time-invariant in our data, individual fixed effects absorb baseline cross-discipline differences; discipline indicators would be collinear with researcher fixed effects. Our estimates represent pooled averages across engineering, economics, and the health sciences.

It is important to recall that the causal interpretations of the coefficients are only valid if the assumption of the treatment's exogeneity is fulfilled. To validate this assumption, we regress the timing of the treatment over the productivity and collaboration characteristics of the treated researchers in the starting year of their academic career using the following equation:[11]

$$Y_i = X'\beta + \varepsilon_i \tag{2}$$

Y is the event's timing for individual i, and X is a vector of productivity controls (including the number of citations and papers) and collaboration characteristics (including the number of coauthors and the core number) in the starting year

[10]Following Criscuolo et al. (2019) and MaCurdy and Pencavel (1986), we add one to the values of the variables before taking logs to avoid excluding observations with a value of zero, which constitute 17 to 37% of the outcome variable observations.

[11]The number and percentage of treated authors are displayed in table B.2

of the researcher's academic career. In addition, we include the first year of a researcher's career as a control.[12]

4 Results

We distinguish between two citation outcomes that capture different margins of impact. Total citations capture overall visibility and accumulated impact over time. By contrast, citations per paper serve as a proxy for average reach conditional on output, which is useful because productivity (the number of papers) may itself change after treatment. We therefore read total citations as an extensive-margin measure of cumulative visibility, and citations per paper as an intensive-margin measure of average reach.

In Table 2, we present the results from estimating equation 2. These estimations suggest that the event timing correlates with two key control variables: the initial core number and the starting year. The correlation with core number suggests that a larger network of coauthors, resulting from the first publication of young researchers, facilitates collaboration between such researchers and prominent authors (see Fig. 3). Regarding the starting year, the estimation results indicate that, in recent years, collaboration with prominent researchers has been more common than in the past. Given these results, we account for cohort differences by using researcher fixed effects and year fixed effects in equation (1).

[12] Following Huang et al. (2020), we define the beginning of a researcher's career as the year of their first publication.

Table 2: Timing of the event

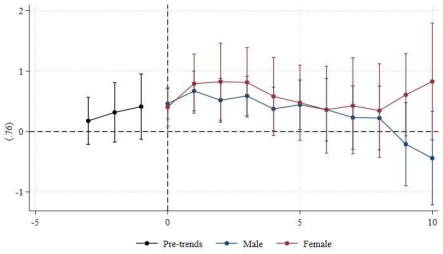
VARIABLES	(1) Timing
Initial Core Number	-0.146**
	(0.0656)
Initial Coauthors (*ln*)	0.672
	(0.907)
Initial Citations (*ln*)	-0.287
	(0.715)
Initial Publications (*ln*)	-0.215
	(1.349)
Starting Year	-0.541***
	(0.0956)
Constant	1,091***
	(192.3)
Observations	135
R-squared	0.347

Notes: All control variables are set to the starting year of each researcher's career. The regression is performed only on the treated researchers. The number of observations denotes researcher-year observations (person-years). The number of unique treated researchers is reported in Table B.2. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

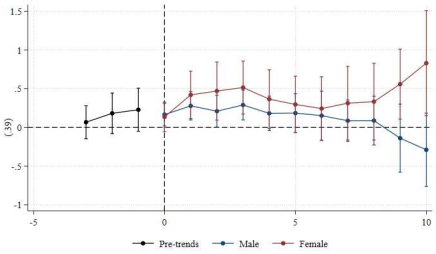
Figure 1 overlays the event-study estimates for women and men on the same scale, which facilitates a direct comparison of dynamics and uncertainty across genders. The pre-trend coefficients suggest no pre-event effects, indicating that the impact on the outcome variables is attributed solely to collaborating with a prominent researcher. Furthermore, the estimations highlight that, regardless of the researcher's gender, collaborating with prominent researchers has a significant impact on almost all impact measures, a key finding of our study.

Figure 1: Average Treatment Effect on the Treated by gender

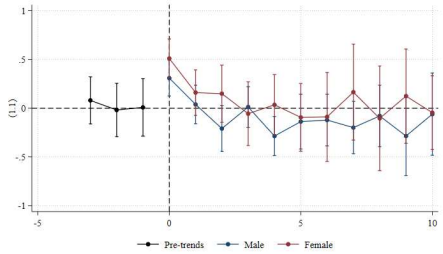
(a) Number of citations (ln)



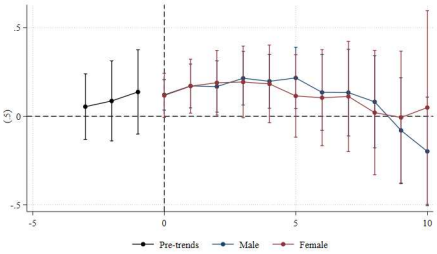
(b) Citations per paper (ln)



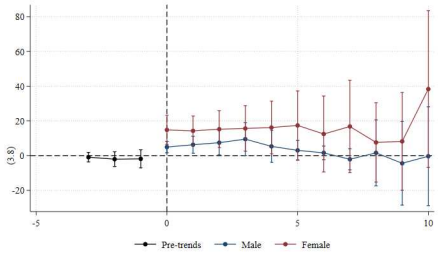
(c) Number of publications (ln)



(d) h-index (ln)



(e) Core number



Notes: The number in parentheses represents the mean of the outcome variable one year before the event. Vertical lines represent 95% CI. The event is defined as collaborating with a prominent researcher within the first five years of the authors' academic career. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution. The imputation sample excludes never-treated researchers.

According to our results, the event causes a positive effect on the number of citations for both female and male researchers. These effects are statistically significant until the third and fifth year after the event for female and male researchers, respectively. The general ATT effect on the number of citations

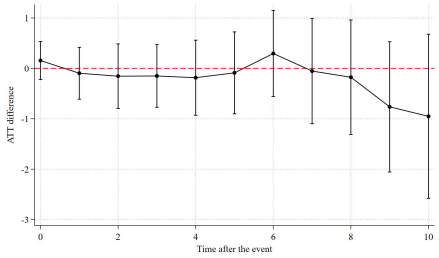
over the first 10 years after the event (see Table C.2) suggests that treated male researchers experience a 46% increase in the number of citations compared to not-yet-treated researchers. Regarding female researchers, these results suggest that treated researchers experience a substantial increase in the number of citations, approximately 77%. This indicates that, although the effect on male researchers is more prolonged over time, the effect on female researchers is more intense.

The effect on the number of citations per paper is more moderate. For women, statistically significant effects are observed one year after the event and persist until the third year after collaborating with a prominent researcher. On the other hand, for male researchers, the effects are significant between the year of the event and three years after. The general ATT over the first ten years after collaborating with a prominent researcher indicates that women experience an increase in the number of citations per paper of approximately 43%, while the ATT for male researchers is not statistically distinct from zero. Regarding the last metric of academic impact, the h-index results suggest that the effects are statistically significant until the second and third year after the event for female and male researchers, respectively. However, the results in Table C.1 indicate that the overall effect is not statistically significant for men, while the effect for female researchers (14%) is weak, as it is marginally significant.

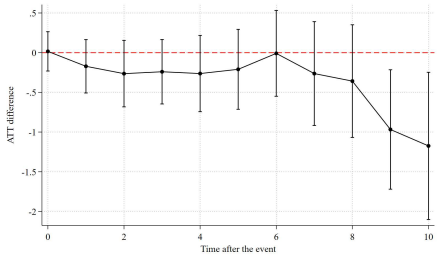
The weak effect on the h-index can be attributed to the disparity between the impact on citations and on publications. Regarding the latter, the results indicate positive, statistically significant effects for both female and male researchers on the year of collaboration with a prominent researcher. Nevertheless, these effects are not perdurable over time, as subsequent effects are not statistically distinguishable from zero. Furthermore, the results in Table C.2 suggest that the overall effect on the number of publications is not statistically significant for either male or female researchers. Because the h-index combines the number of citations and the number of publications, a lack of effect on researchers' productivity yields a null effect on the h-index.

Figure 2: ATT Effect Differences Between Female and Male Researchers

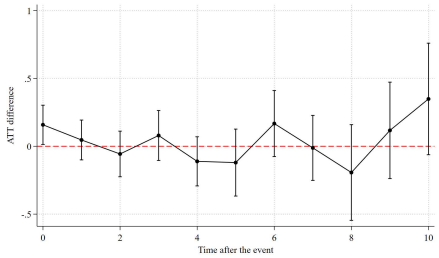
(a) Number of citations (ln)



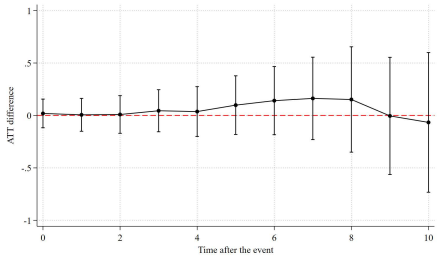
(b) Citations per paper (ln)



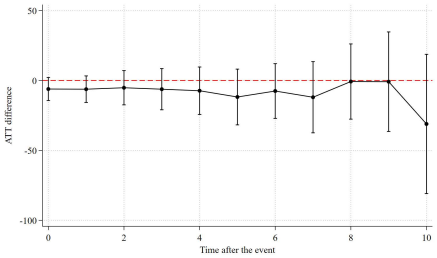
(c) Number of publications (ln)



(d) h-index (ln)



(e) Core number



Notes: To obtain the ATT difference and its confidence intervals, we estimate a linear combination of coefficients obtained by performing the regression described in equation 1. Specifically, we estimate the linear combination $\tau_{m,male} - \tau_{m,female}$. Vertical lines represent 95% CI. The imputation sample excludes never-treated researchers.

Similar to the findings on the number of citations, the effect on the authors' connectivity metric (core number) is statistically significant for both male and female researchers. For women, the effect is statistically significant from the year of the event occurrence to four years after. These effects result in an overall increase of 15.23 units in the core number relative to researchers who

have not yet been treated. For treated male researchers, the effect persists for 3 years after treatment. However, the effect on the core number is modest (4.7 units) and marginally significant.

Figure 2 shows the difference between the ATT effects for male and female researchers with confidence intervals at 95%.[13] This figure suggests that there are no important disparities between the effects on male and female researchers, suggesting that both women and men in academia obtain the same benefits from collaborating with a prominent researcher.

4.1 Illustrative case of the collaboration in the coauthorship network

To better capture how collaboration affects the core number of a junior researcher, we present an illustrative case. Figure 3 displays the coauthorship network of a real SNI researcher (green node) who collaborated with a prominent scholar (red node) within the first five years of his academic career.

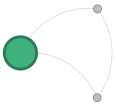
During the first four years, before the collaboration, the junior researcher had a small network with two coauthors. Specifically, one year before the event (Year -1), the SNI researcher's core number was two. At the event time (Year 0), the SNI researcher "imports" the coauthorship network of the prominent researcher, thereby expanding his own network; his number of coauthors is 5, and his core number remains 2. However, if the collaboration paper is removed, the SNI researcher would have three coauthors, and his core number remains at two. The contrast between Figure 3a and Figure 3b illustrates how the collaboration with a prominent researcher sharply increases the social capital of the junior collaborator, since he acquires social capital driven by the prominent researcher's direct coauthors and coauthors-of-coauthors.

It is not until approximately four years after the event that the SNI researcher begins to develop an independent sub-network distinct from that of

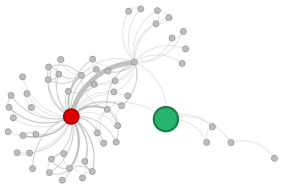
[13]The difference between the ATT effects for male and female researchers and their confidence intervals was obtained by estimating a linear combination of the form $\tau_{m,male} - \tau_{m,female}$, using the coefficients obtained by performing the regression described in equation 1.

the prominent collaborator (Figure 3c). At this point, he has ten coauthors, and his core number is four. Furthermore, eleven years after the event (Figure 3e), the SNI researcher's number of coauthors increases to 18, while his core number is five. The eventual decrease in the effect of collaboration on the core number may reflect his coauthorship network converging toward a long-term structure rather than one primarily shaped by the prominent researcher's connections.

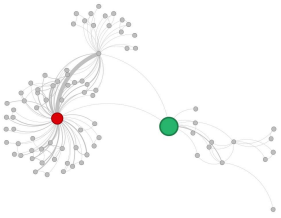
Figure 3: Effect of Collaboration with a Prominent Researcher (red node) in the Coauthorship Network of a Sample Researcher (green node).



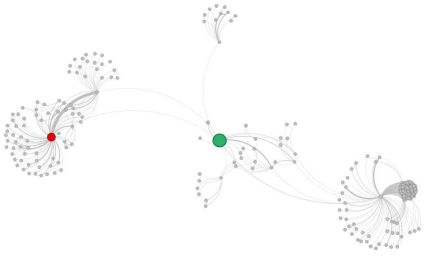
(a) Year −1.



(b) Year 0.



(c) Year 4.



(d) Year 11.

Notes: (a) Collaboration network before the collaboration with the prominent researcher. **(b)** Collaboration network during the year of the collaboration. **(c–d)** Collaboration networks in subsequent years after the collaboration.

Thus, a sharp increase in a researcher's core number signals an expansion in both the size and cohesiveness of the collaborative environment, granting him access to collaborators who themselves hold substantial social capital. Being part of a tightly knit core may reduce information frictions and create opportunities for both new and repeated collaborations. Thus, the increase in the SNI researcher core number is not only a quantitative expansion of collaborators but also a qualitative improvement in the researcher's strategic position within the scientific community.

5 Robustness

To check the robustness of our results, we estimate equation 1 using alternative definitions of the event. We use two thresholds in the maximum h-index distribution for coauthors to identify which researchers collaborated with a prominent author during the first five years of their academic careers; these thresholds correspond to the 96th and 97th percentiles of the distribution. The results of these regressions are in tables D.1 and D.2.

Our results remain qualitatively similar to the changed definition of the event. For both women and men, estimates using the 96th and 97th percentiles suggest that our results are robust since all effects are similar to those obtained in the main results (see table C.1). Furthermore, although some estimates for women and men using the alternative definitions present some statistically significant effects before the event, pre-trend p-values suggest no evidence of pre-event effects.

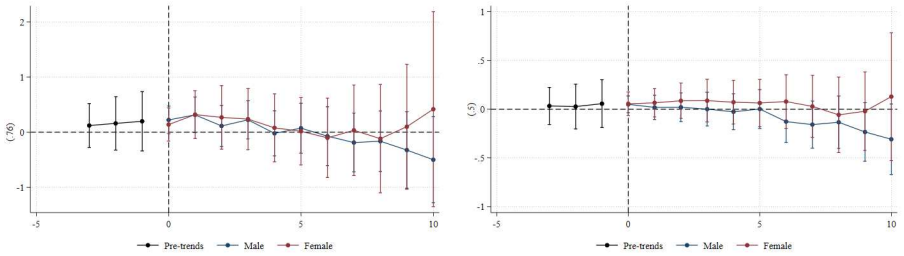
5.1 Focal Paper

A crucial question is whether the positive effects observed in productivity, impact, and connection measures are structural changes in the research activity of treated researchers or merely a temporary boost resulting from the focal paper. We define the focal paper as the paper that detonates the event, i.e., the paper

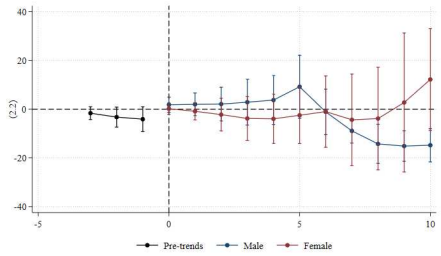
in which the treated researchers collaborate with a prominent researcher. In this sense, we identified the focal paper for all researchers who were treated. Then, we compute the number of citations and the h-index, excluding citations from the focal paper. Similarly, we compute a core number excluding the collaboration network associated with the research team included in the focal paper.[14] In Figure 4, we present the results of this exercise.

Figure 4: Average Treatment Effect on the Treated by Gender Excluding the Focal Paper

(a) Number of citations (ln) (b) h-index (ln)



(c) Core number



Notes: The number in parentheses represents the mean of the outcome variable one year before the event. Vertical lines represent 95% CI. The event is defined as collaborating with a prominent researcher within the first five years of the authors' academic career. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution. The focal paper is defined as the paper in which the treated researchers collaborate with a prominent researcher. The imputation sample excludes never-treated researchers.

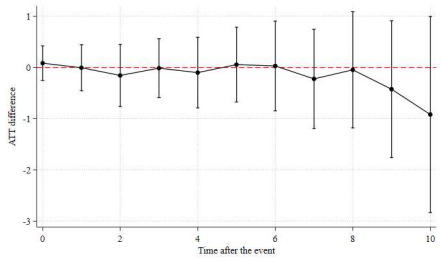
The results suggest no statistically significant effects across any of the mea-

[14]Figure B.2 in the Appendix illustrates the change in the aforementioned measures when excluding the focal paper in the year of the event, five, and ten years later.

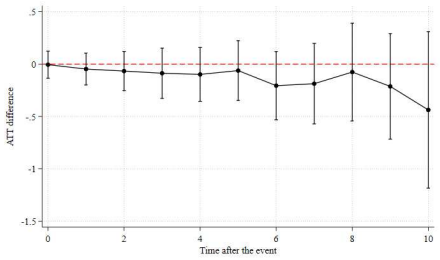
sures. These results are confirmed by the estimations presented in Table C.3, which indicate that there is no statistically significant effect on any of the measures after the event. These results suggest that the effects observed in the previous section are entirely driven by the focal paper, the resulting paper from the collaboration with a prominent researcher.

Figure 5: ATT Effect Differences Between Female and Male Researchers Excluding the Focal Paper

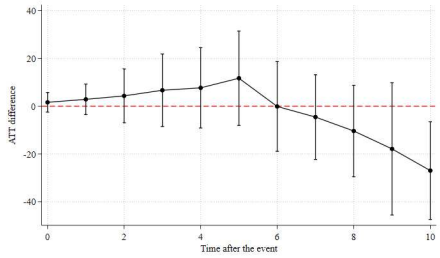
(a) Number of citations (ln)



(b) h-index (ln)



(c) Core number



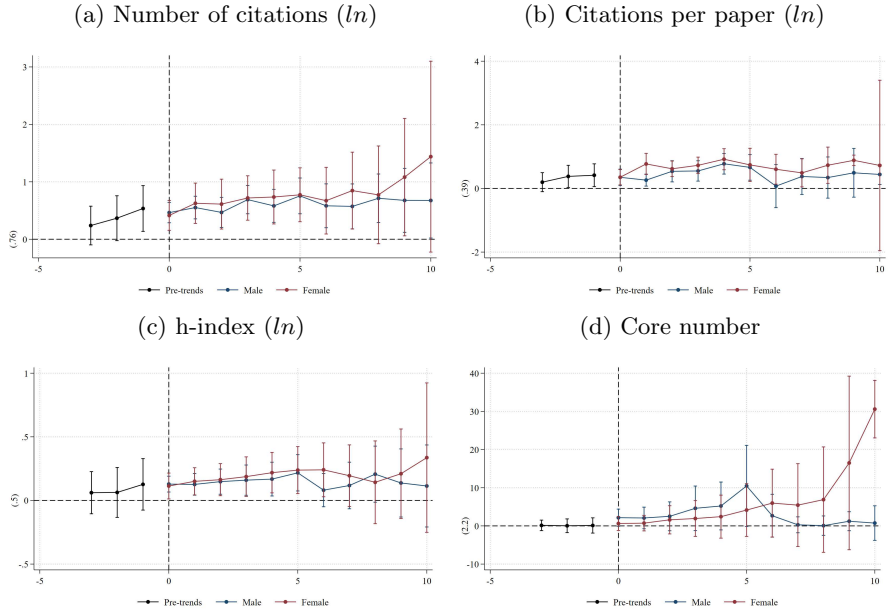
Notes: To obtain the ATT difference and its confidence intervals, we estimate a linear combination of coefficients obtained by performing the regression described in equation 1. Specifically, we estimate the linear combination $\tau_{m,male} - \tau_{m,female}$. Vertical lines represent 95% CI. The focal paper is defined as the paper in which the treated researchers collaborate with a prominent researcher. The imputation sample excludes never-treated researchers.

The results in this section should nevertheless be interpreted with caution, because our estimates may be downward-biased. For treated authors, we remove the focal collaboration paper from all outcomes, whereas for the comparison group, we keep all publications, including those stemming from collaborations with prominent researchers that occur later in their careers. As a result, treated

authors are at a mechanical disadvantage relative to the control group, which may help explain the lack of statistically significant effects.

To partially overcome the aforementioned limitation, we used an alternative control group: "never-treated" researchers—those who have never collaborated with a leading researcher in their first five years or at any other point in their academic career. The estimated effects using the never-treated researchers as a control group are depicted in Figure 6.

Figure 6: Average Treatment Effect on the Treated by Gender Excluding the Focal Paper (Alternative Control Group)



Notes: The number in parentheses represents the mean of the outcome variable one year before the event. Vertical lines represent 95% CI. The event is defined as collaborating with a prominent researcher within the first five years of the authors' academic career. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution. The focal paper is defined as the paper in which the treated researchers collaborate with a prominent researcher. The imputation sample only includes never-treated researchers.

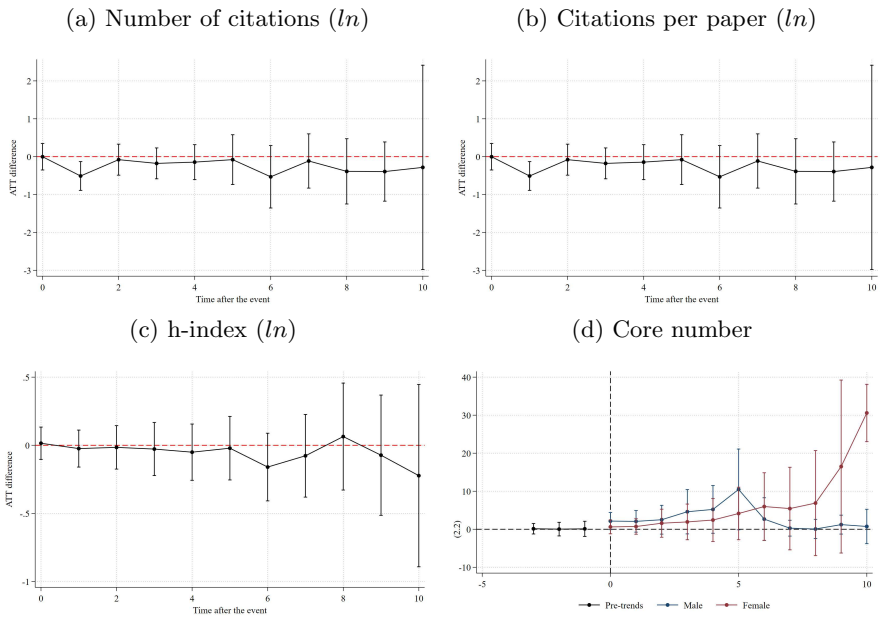
The fact that the researchers in this group have never collaborated with a leading researcher means that no highly regarded author has ever boosted their

productivity, impact, or networks. This, in turn, means that the comparison between this control group and the treated researchers is more "fair" than the previous comparison. A disadvantage of this approach is that treated authors are more likely to collaborate with prominent researcher because they are more productive than untreated authors, which could result in upward-biased estimates.

The results suggest that collaborating with a prominent researcher has positive, significant effects on citations. However, in panels (a) and (b) of Figure 6, significant pre-trends are evident; hence, the estimators lack credibility for this outcome. In contrast, results for the number of citations per paper indicate a positive effect in the first five years after the event, without pre-trends, which, along with the results on the h-index, suggests that treated authors experience an increase in the impact of their research. Finally, results for the core number indicate no significant effects in the first 9 years after the event, suggesting that treated authors do not form multiple collaborations with prominent authors or other researchers in the network. It is important to note that there are no significant differences between male and female researchers as depicted in Figure 7.

Overall, our findings provide support for the set of hypotheses in Section 2, although there are no systematic gender differences. We find sizable citation gains after early-career collaboration with prominent researchers for both women and men. Point estimates are larger for women, but the difference is not statistically significant. At the same time, we find no significant effect on the number of publications and only transitory effects on the h-index. Collaboration also places treated researchers into more cohesive regions of the coauthorship network, as indicated by an increase in the core number. Robustness estimations excluding the focal paper indicate that these effects are largely driven by the focal paper rather than by persistent structural changes in broader academic careers.

Figure 7: ATT Effect Differences Between Female and Male Researchers Excluding the Focal Paper (Alternative Control Group)



Notes: To obtain the ATT difference and its confidence intervals, we estimate a linear combination of coefficients obtained by performing the regression described in equation 1. Specifically, we estimate the linear combination $\tau_{m,male} - \tau_{m,female}$. Vertical lines represent 95% CI. The focal paper is defined as the paper in which the treated researchers collaborate with a prominent researcher. The imputation sample only includes never-treated researchers.

6 Conclusions

In this study, we investigate whether early-stage collaboration with a prominent researcher can enhance a researcher's impact, productivity, and network connections in subsequent career stages. To address this question, we constructed a dataset using Scopus metadata on papers and citation counts for researchers in the 2021 Mexican National System of Researchers (SNI) registry across three disciplines: economics, engineering, and health sciences. We then estimate the causal effect of collaboration with a prominent researcher on researchers' impact, productivity, and collaboration trends using the staggered DID imputation approach proposed by Borusyak et al. (2024).

Two features set our study apart. First, we use the authors' core number,

which captures whether an author belongs to a highly interconnected community of researchers within the broader coauthorship network, rather than relying solely on the number of their connections (i.e., coauthors). The second difference relies on our definition of a prominent researcher. Unlike most studies in the literature, we identify a prominent researcher based on their h-index rather than their total number of citations. This method ensures that the observed effects in all outcome variables result from collaborating with a researcher of greater seniority and impact in academia rather than a researcher with a single "big hit." Additionally, we partially separate the effect of the focal paper – the paper resulted from the collaboration with a prominent researcher.

Our results suggest that collaborating with a prominent researcher can amplify the impact of the treated authors' research, the total number of citations, the number of citations per paper, and increase their core number. However, there are no statistically significant effects on their annual number of publications. This null effect on productivity, measured as the number of papers, is similar to the finding of Dorantes-Gilardi et al. (2023). Furthermore, these effects do not differ according to the authors' sex.

One of our most novel findings is the collaboration metric used, the core number. The cumulative effects of women collaborating with a prominent researcher exceed those found for men, although the difference is not statistically different from zero. This disparity suggests that collaborating with a prominent researcher benefits women to a greater extent in terms of the quality of connections, for instance, gaining access to better financial resources or taking an active role in developing their discipline. However, the relationship between connection quality and testing the positive externalities require further investigation.

This study disentangles whether the positive effect is due to the paper the researcher coauthors with a prominent researcher or their entire portfolio of publications. We disentangle the effect from the focal paper by excluding the number of citations and the network resulting from the collaboration with a prominent researcher. Our results suggest that the positive effects are entirely

driven by the focal paper, rather than by the researchers' entire portfolio. Nevertheless, the results from this exercise may be biased due to the lack of an appropriate control group, which invites further exploration.

Our findings should be interpreted in light of a few scope conditions. First, by construction, only a relatively small number of junior researchers coauthor early with a top-2% prominent collaborator. Second, our data capture only Scopus-indexed outputs, leaving out nationally oriented publications that may also matter for career evaluation. Third, we cannot fully observe how junior and senior authors were matched into collaboration. In addition, a substantial share of the observed gains in impact is concentrated in the highly cited focal collaboration paper rather than being spread evenly across the rest of the publication record. Our conclusions are scoped to engineering, economics, and health sciences, and it remains an open question whether the same pattern would emerge in fields dominated by very large lab teams or predominantly single-authored traditions.

This analysis comes from a specific institutional setting—the Mexican National System of Researchers (SNI)—but one that shares key features with many national research systems: disciplinary evaluation committees, field-based hierarchies of prestige, and career rewards tied to publication and citation records. Most national systems are similarly organised along disciplinary lines; although disciplines have country-specific institutional features, they also display common norms and hierarchies that transcend national borders.

A key contribution of our analysis is that we can condition on the degree of "masculinization" of each discipline—measured by the share of women among researchers in that field—so that our comparisons are not simply capturing that women tend to work in less-cited or less-prestigious areas. Even after accounting for these differences in gender composition across disciplines, we still find sizeable gender gaps in overall impact and network position, but no evidence that women benefit differentially from early collaboration with prominent researchers.

Taken together, these features indicate that our estimates speak most di-

rectly to early collaborations with very prominent researchers in systems in which such collaborations are a salient opportunity structure, and that they should be extrapolated with caution to contexts with very different institutional arrangements or collaboration norms. Within this domain, however, the patterns we document are consistent and robust across outcomes and specifications, and provide novel evidence on how star collaborations shape early-career trajectories.

Because our definition of prominence targets the extreme upper tail of the coauthor h-index distribution, the treated group comprises junior researchers who connect early with a highly selective set of star collaborators. Productivity and impact dynamics at this tail may differ from those in the center of the distribution; therefore, our results should be interpreted as characterizing the consequences of early star collaborations rather than typical early-career coauthorships. Further investigations are needed to understand how these effects compare to those of collaborations with less extreme but still highly cited coauthors.

From a policy perspective, our results suggest that early coauthorship with a prominent researcher can facilitate highly visible "big-hit" publications for both women and men. Still, it does not automatically translate into sustained career acceleration. This pattern cautions against policies that rely exclusively on one-off star collaborations to remedy gender gaps in academic impact. Instead, our findings shift the narrative toward the importance of institutional structures that support repeated collaboration and sustained network access, particularly for underrepresented groups.

References

Bartneck, C. and S. Kokkelmans (2011). Detecting h-index manipulation through self-citation analysis. *Scientometrics 87*, 85–98.

Bazeley, P. (2003). Defining 'Early Career' in Research. *Higher Education 45*(3), 257–279.

Beaudry, C. and V. Larivière (2016). Which gender gap? factors affecting researchers' scientific impact in science and medicine. *Research Policy 45*, 1790–1817.

Bornmann, L. and H. D. Daniel (2007, 7). What do we know about the h index? *Journal of the American Society for Information Science and Technology 58*, 1381–1385.

Borusyak, K., X. Jaravel, and J. Spiess (2024, feb). Revisiting Event-Study Designs: Robust and Efficient Estimation. *The Review of Economic Studies*, rdae007.

Costas, R. and M. Bordons (2007, 7). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics 1*, 193–203.

Criscuolo, C., R. Martin, H. G. Overman, and J. Van Reenen (2019, January). Some causal effects of an industrial policy. *American Economic Review 109*(1), 48–85.

Dorantes-Gilardi, R., A. A. Ramírez-Álvarez, and D. Terrazas-Santamaría (2023). Is there a differentiated gender effect of collaboration with super-cited authors? Evidence from junior researchers in economics. *Scientometrics*.

Dorantes-Gilardi, R., A. A. Ramírez-Álvarez, and D. Terrazas-Santamaría (2022). The role of highly intercited papers on scientific impact: the mexican case. *Applied Network Science 7*, 58.

Ebadi, A. and A. Schiffauerova (2015). How to receive more funding for your research? get connected to the right people. *PLoS ONE 10*, 1–19.

Gil Antón, M. and L. E. Contreras Gómez (2017). El sistema nacional de investigadores: ¿espejo y modelo? *Revista de la Educación Superior 46*(184), 1–19.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *PNAS 102*, 16569–16572.

Huang, J., A. J. Gates, R. Sinatra, and A.-L. L. Barabási (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences of the United States of America 117*(9), 4609–4616.

Jaramillo, A. M., H. T. P. Williams, N. Perra, and R. Menezes (2023, Oct). The structure of segregation in co-authorship networks and its impact on scientific production. *EPJ Data Science 12*(1), 47.

Jin, Y., Y. Yin, S. Fortunato, and D. Wang (2020). Scientific elite revisited: patterns of productivity, collaboration, authorship and impact. *Journal of Royal Society Interface 17*(165), 1–10.

Koltun, V. and D. Hafner (2021). The h-index is no longer an effective correlate of scientific reputation. *PLoS ONE 16(6)*.

Krauss, A., L. Danús, and M. Sales-Pardo (2023, 12). Early-career factors largely determine the future impact of prominent researchers: evidence across eight scientific fields. *Scientific Reports 13*.

Larivière, V., C. Ni, Y. Gingras, B. Cronin, and C. R. Sugimoto (2013). Bibliometrics: Global gender disparities in science. *Nature 504*, 211–213.

Li, W., T. Aste, F. Caccioli, and G. Livan (2019). Early coauthorship with top scientists predicts success in academic careers. *Nature Communications 10*(1), 1–9.

Li, W., S. Zhang, Z. Zheng, S. J. Cranmer, and A. Clauset (2022, 12). Untangling the network effects of productivity and prominence among scientists. *Nature Communications 13*.

MaCurdy, T. E. and J. H. Pencavel (1986). Testing between competing models of wage and employment determination in unionized markets. *Journal of Political Economy 94*(3, Part 2), S3–S39.

Malliaros, F. D., C. Giatsidis, A. N. Papadopoulos, and M. Vazirgiannis (2020). The core decomposition of networks: theory, algorithms and applications. *The VLDB Journal 29*(1), 61–92.

Moreno, J. A. J. (2019). El sistema nacional de investigadores en méxico como mecanismo meritocrático de un estado evaluador. *Reflexión Política 21*(41), 81–90.

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences 101*(suppl_1), 5200–5205.

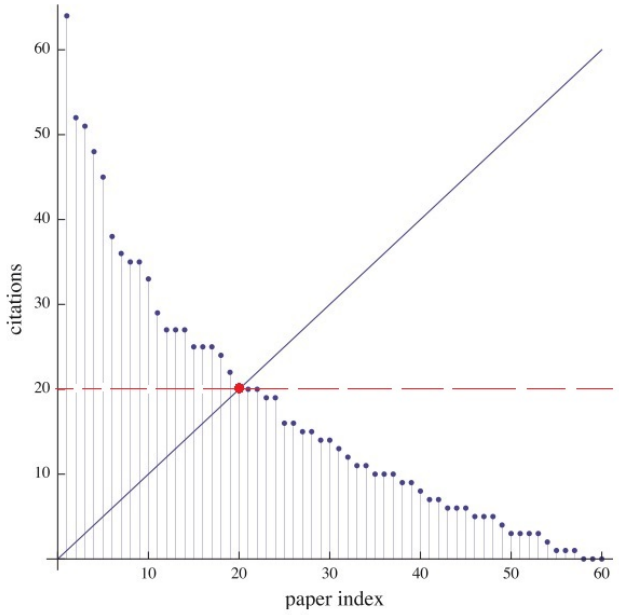
Rose, M. E. and J. R. Kitchin (2019). Pybliometrics: Scriptable bibliometrics using a python interface to scopus. *SoftwareX 10*.

Appendix

A H-index and core number

Consider Figure A.1, which provides an example of a researcher with an h-index of 20. As Bartneck and Kokkelmans (2011) explain, we can list their publications in decreasing order of the number of citations, where the papers with the most citations are at the left of the horizontal axis. The diagonal line represents the number of papers equal to the number of citations, where 20 papers have at least 20 citations each. Note that the h-index is where the diagonal line intersects with the citation profile, and that once the highly cited papers are in the top group, the number of citations they receive in excess of h is unimportant.

Figure A.1: Citation Profile of an Author with h-index of 20

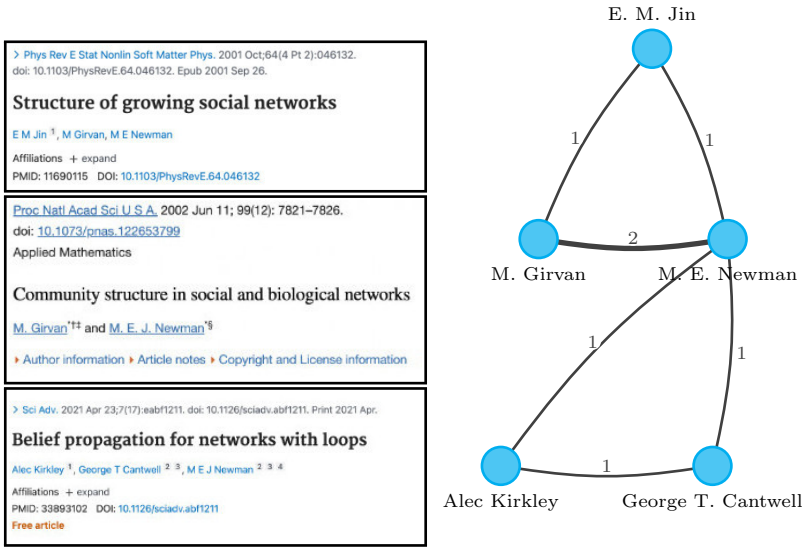


Source: Bartneck and Kokkelmans (2011).

Hirsch (2005) introduces the m index, defined as the h-index divided by the number of years of scientific activity, and argues that, for a scientist with around 20 years of activity, $m \approx 1$ (that is, an h-index of about 20) characterizes a "successful" researcher, $m \approx 2$ ($h \approx 40$) an "outstanding" one, and m ≈ 3 ($h \approx 60$) a "truly exceptional" scientist. Against this benchmark, relatively low values of the h-index, such as those around the 90th percentile of our coauthor distribution (e.g., $h \approx 7$ in economics in 2005), do not seem sufficient to justify the label "prominent".

A coauthorship network is defined by a graph $G = (V, E)$, where the set of vertices (or nodes) V represents researchers and the set of edges (or links) E represents coauthorship relationships. If two researchers coauthor a paper, they are connected through a weighted link. The weight of the edges indicates the number of coauthored papers. Figure A.2 shows an example of a coauthorship network built from three papers in which five authors are involved. While an author's degree measures the number of distinct coauthors, the core number is a descriptive measure of integration into cohesive teams within the coauthorship network.

Figure A.2: Coauthorship Network Built from Three Papers.



We construct coauthorship networks by mapping papers to their publication year and research area. For each research area and calendar year t, we build a cumulative coauthorship network using all publications in that area dated up to year t. Importantly, the network is not limited to focal SNI researchers: it includes all his coauthors and the coauthors of his couathors that appear on their Scopus-indexed papers. In this model, edges are weighted by the number of coauthored papers, and these weights are incorporated into the computation of network metrics. We then compute each focal researcher's degree and core number on this area-specific network for year t, so the measures reflect the broader structure of coauthorship ties in the area at that point in time, not only the focal author's direct links.

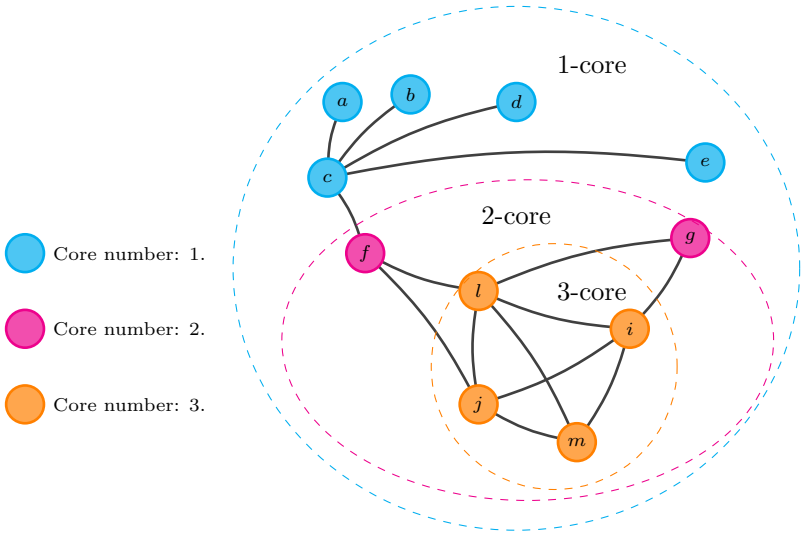
We use yearly cumulative networks rather than rolling windows. A rolling window (e.g., restricting edges to the last k years) would mechanically drop older

coauthorship ties, even when those ties may remain relevant for visibility and access in academic careers. This is particularly important in disciplines with slower publication cycles, where meaningful collaborations may not be observed repeatedly year-to-year. Using cumulative networks is therefore better suited to capture longer-run embeddedness and integration into collaboration structures.

Before defining the core number, it is necessary to define the k-core of a network. The k-core of a network G is the sub-network of G where all nodes have a degree of at least k. The k-core of a coauthorship network is thus a sub-network where every researcher has at least k coauthors. Figure A.3 shows an example of the k-core decomposition of a network, which creates three nested k-core sub-networks:

- The 3-core network is the sub-network formed by nodes with a minimum degree of 3, that is, nodes l, i, j, and m.
- The 2-core network is the sub-network formed by nodes with a minimum degree of 2, that is, nodes l, i, j, m, f and g.
- The 1-core network is the sub-network formed by nodes with a minimum degree of 1, that is, the whole network.

Figure A.3: An example of k-core decomposition.



A researcher has core number of k if they belong to a k-core but not to a $k+1$-core, which means that the researcher and their coauthors have at least k coauthors. In Figure A.3, orange nodes have a core number of 3, pink nodes have a core number of 2, and blue nodes have a core number of 1.

The core number of a researcher measures how deeply embedded they are in the coauthorship network. It is the largest value k such that the researcher belongs to a group of colleagues where each one has collaborated with at least k others within the same group. Consider researchers l and c in the coauthorship network shown in Figure A.3. Although both have a degree of 5, their core numbers are 3 and 1, respectively. This is because the coauthors of researcher l are more interconnected than those of researcher c.

Thus, the core number of a node is defined as the largest integer k such that the node belongs to the k-core of the network, i.e., a maximal subgraph in which every node has degree at least k.

B Data

Table B.1: Representativeness of the Sample

	(1)	(2)	(3)
VARIABLES	Economics	Engineering	Health Sciences
Scopus database	-0.0229	0.00813	0.0248
	(0.0258)	(0.0253)	(0.0317)
Constant	0.347***	0.219***	0.532***
	(0.0172)	(0.0184)	(0.0269)
Observations	1,364	1,085	1,229
R-squared	0.001	0.000	0.000

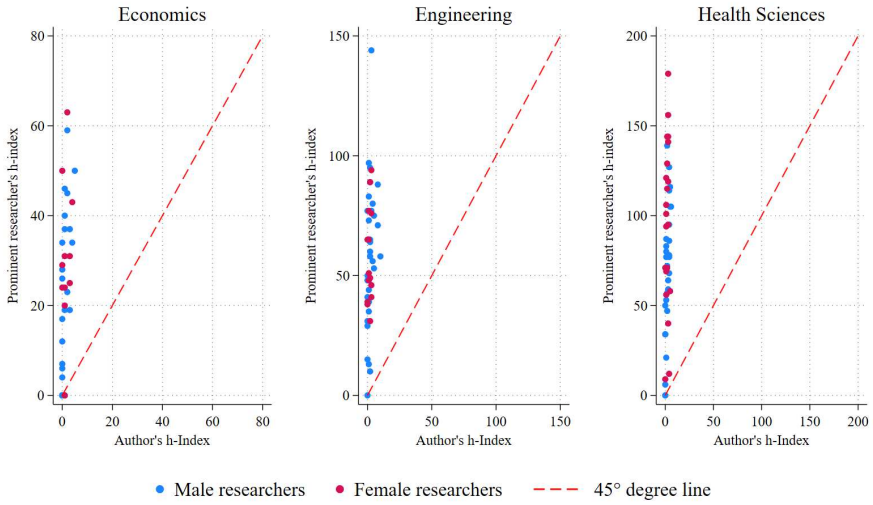
Notes: The variable Scopus database is a dummy equal to one if the author is in the Scopus database. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.2: Treated authors

	Number of treated authors			Percentage of treated authors		
	Economics	Engineering	Health Sciences	Economics	Engineering	Health Sciences
a) Event using the 96th percentile						
Female	18	21	46	9.18	16.15	9.31
Male	58	53	45	14.22	11.96	11.45
b) Event using the 97th percentile						
Female	13	17	35	6.63	13.08	7.09
Male	45	38	40	11.03	8.58	10.18
c) Event using the 98th percentile						
Female	11	14	21	5.61	10.77	4.25
Male	35	32	30	8.58	7.22	7.63

Notes: Treated authors are defined as those researchers who, within the first years of their academic career, collaborate with a prominent researcher. A prominent researcher is defined as a coauthor in the 96th, 97th, or 98th percentile of the coauthors' h-index distribution. This table reports counts of *unique treated researchers* (people), by gender (and discipline, when applicable). These counts differ from observations in regression tables, which report researcher-year observations.
Source: Authors' calculations.

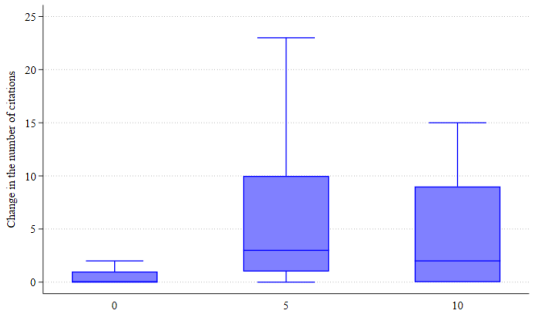
Figure B.1: Treated authors' h-index vs prominent researchers' h-index



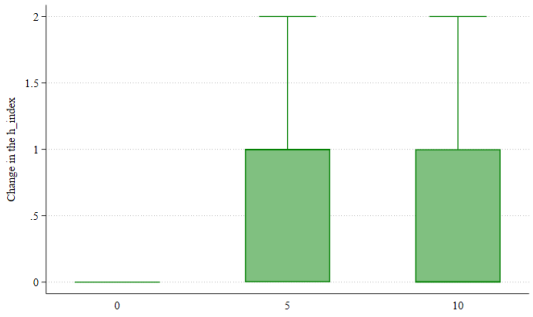
Notes: We defined prominent researchers as those in the 98th percentile of coauthors' h-index distribution in a given year and field. Treated authors are those authors who, during the first five years of their academic career, collaborated with a prominent researcher.
Source: Authors' calculations.

Figure B.2: Changes in Outcome Variables Following the Exclusion of the Focal Paper at Three Time Points: At Collaboration, and Five, and Ten Years After.

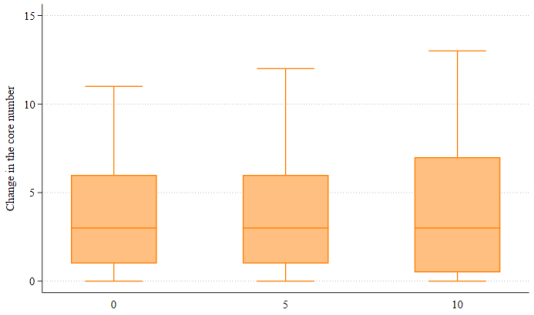
(a) Number of citations



(b) h-index (ln)



(c) Core number



Notes: Extreme values are excluded. The change in all variables presented in the figure is calculated by subtracting the outcome variable after excluding the focal paper from the original outcome variable at three different time points (horizontal axis).
Source: Authors' calculations.

C Results

Table C.1: Effect of Collaborating with a Prominent Researcher within the First Five Years of Academic Career

VARIABLES	(1) Ln(Citations + 1)	(2) Ln(Citations per paper + 1)	(3) Ln(H-Index + 1)	(4) Ln(Publications + 1)	(5) Core number
$\tau_{0,male}$	0.460***	0.165**	0.121***	0.309***	4.946***
	(0.128)	(0.0747)	(0.0440)	(0.0939)	(1.682)
$\tau_{0,female}$	0.404**	0.135	0.117*	0.510***	14.80***
	(0.164)	(0.0970)	(0.0637)	(0.102)	(4.264)
$\tau_{1,male}$	0.670***	0.278***	0.171***	0.0383	6.275**
	(0.169)	(0.0934)	(0.0630)	(0.101)	(2.523)
$\tau_{1,female}$	0.792***	0.420***	0.170**	0.160	14.27***
	(0.250)	(0.156)	(0.0776)	(0.119)	(4.350)
$\tau_{2,male}$	0.518***	0.209**	0.168**	-0.209*	7.474**
	(0.186)	(0.105)	(0.0739)	(0.120)	(3.584)
$\tau_{2,female}$	0.824**	0.469**	0.189**	0.149	15.24***
	(0.326)	(0.191)	(0.0928)	(0.150)	(5.409)
$\tau_{3,male}$	0.589***	0.289***	0.214***	0.0120	9.494**
	(0.166)	(0.0983)	(0.0772)	(0.106)	(4.819)
$\tau_{3,female}$	0.813***	0.514***	0.193*	-0.0566	15.65**
	(0.293)	(0.174)	(0.103)	(0.167)	(6.688)
$\tau_{4,male}$	0.374**	0.181	0.197**	-0.286***	5.342
	(0.185)	(0.114)	(0.0777)	(0.102)	(4.697)
$\tau_{4,female}$	0.581*	0.364*	0.182	0.0346	16.19**
	(0.330)	(0.193)	(0.112)	(0.158)	(7.735)
$\tau_{5,male}$	0.443**	0.184	0.216**	-0.138	3.050
	(0.209)	(0.127)	(0.0884)	(0.143)	(2.911)
$\tau_{5,female}$	0.476	0.296	0.115	-0.0945	17.40*
	(0.317)	(0.186)	(0.119)	(0.178)	(10.12)
$\tau_{6,male}$	0.361	0.152	0.135	-0.122	1.604
	(0.264)	(0.161)	(0.109)	(0.135)	(2.001)
$\tau_{6,female}$	0.361	0.243	0.105	-0.0893	12.46
	(0.367)	(0.210)	(0.138)	(0.233)	(11.14)
$\tau_{7,male}$	0.231	0.0861	0.133	-0.199	-2.098
	(0.268)	(0.137)	(0.125)	(0.137)	(3.112)
$\tau_{7,female}$	0.426	0.312	0.112	0.164	16.78
	(0.406)	(0.243)	(0.159)	(0.251)	(13.56)
$\tau_{8,male}$	0.222	0.0877	0.0814	-0.0801	1.608
	(0.269)	(0.161)	(0.133)	(0.161)	(9.690)
$\tau_{8,female}$	0.345	0.332	0.0206	-0.105	7.628
	(0.395)	(0.252)	(0.179)	(0.274)	(11.67)
$\tau_{9,male}$	-0.211	-0.139	-0.0796	-0.286	-4.407
	(0.351)	(0.224)	(0.152)	(0.207)	(12.29)
$\tau_{9,female}$	0.607*	0.558**	-0.00672	0.124	8.218
	(0.347)	(0.230)	(0.191)	(0.247)	(14.34)
$\tau_{10,male}$	-0.442	-0.289	-0.198	-0.0614	-0.340
	(0.394)	(0.241)	(0.156)	(0.215)	(14.51)
$\tau_{10,female}$	0.828*	0.829**	0.0490	-0.0458	38.33*
	(0.493)	(0.346)	(0.279)	(0.193)	(23.02)
τ_{-1}	0.412	0.228	0.137	0.00824	-1.861
	(0.275)	(0.142)	(0.121)	(0.150)	(2.655)
τ_{-2}	0.316	0.181	0.0863	-0.0186	-2.034
	(0.251)	(0.134)	(0.115)	(0.139)	(2.164)
τ_{-3}	0.176	0.0664	0.0542	0.0799	-0.893
	(0.199)	(0.109)	(0.0944)	(0.123)	(1.402)
Observations	1,285	1,285	1,285	1,285	1,192
Pre-trends	.9125	.3256	.5662	.7717	.4862
Individual FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Starting Year FE	Yes				
Treated ind.	84	84	84	84	84
Sample	135	135	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution in a given year and field. The imputation sample excludes never-treated researchers.

Table C.2: ATT Effect by Sex

VARIABLES	(1) Ln(Citations + 1)	(2) Ln(Citations per paper + 1)	(3) Ln(H-Index + 1)	(4) Ln(Publications + 1)	(5) Core number
$\tau_{,male}$	0.378**	0.148	0.130*	-0.0497	4.704*
	(0.161)	(0.0940)	(0.0714)	(0.0702)	(2.627)
$\tau_{,female}$	0.571**	0.357**	0.139	0.0950	15.23**
	(0.276)	(0.169)	(0.100)	(0.134)	(6.638)
τ_{-1}	0.412	0.228	0.137	0.00824	-1.861
	(0.275)	(0.142)	(0.121)	(0.150)	(2.655)
τ_{-2}	0.316	0.181	0.0863	-0.0186	-2.034
	(0.251)	(0.134)	(0.115)	(0.139)	(2.164)
τ_{-3}	0.176	0.0664	0.0542	0.0799	-0.893
	(0.199)	(0.109)	(0.0944)	(0.123)	(1.402)
Observations	1,298	1,298	1,298	1,298	1,200
Pre-trends	.9125	.3256	.5662	.7717	.4862
Individual FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Starting Year FE	Yes				
Treated ind.	84	84	84	84	84
Sample	135	135	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The coefficients represent the average ATT during the first ten years after the event. The event is defined as collaborating with a prominent researcher within the first five years of the authors' academic career. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution. The imputation sample excludes never-treated researchers.

Table C.3: Effect of Collaborating with a Prominent Researcher Excluding the Focal Paper

VARIABLES	(1) Ln(Citations + 1)	(2) Ln(h-index+1)	(3) Core number
$\tau_{0,male}$	0.221*	0.0497	1.810
	(0.128)	(0.0439)	(1.579)
$\tau_{0,female}$	0.137	0.0552	0.183
	(0.153)	(0.0617)	(1.190)
$\tau_{1,male}$	0.314*	0.0185	1.992
	(0.164)	(0.0634)	(2.378)
$\tau_{1,female}$	0.318	0.0657	-0.870
	(0.221)	(0.0743)	(1.806)
$\tau_{2,male}$	0.113	0.0199	2.094
	(0.190)	(0.0752)	(3.540)
$\tau_{2,female}$	0.268	0.0862	-2.228
	(0.295)	(0.0925)	(3.434)
$\tau_{3,male}$	0.223	0.000493	2.868
	(0.177)	(0.0885)	(4.784)
$\tau_{3,female}$	0.236	0.0876	-3.804
	(0.283)	(0.111)	(4.611)
$\tau_{4,male}$	-0.0219	-0.0267	3.726
	(0.209)	(0.0937)	(5.125)
$\tau_{4,female}$	0.0781	0.0717	-3.961
	(0.315)	(0.115)	(5.158)
$\tau_{5,male}$	0.0720	0.00151	9.202
	(0.231)	(0.102)	(6.607)
$\tau_{5,female}$	0.0155	0.0633	-2.505
	(0.313)	(0.123)	(5.942)
$\tau_{6,male}$	-0.0727	-0.128	-1.115
	(0.272)	(0.108)	(4.769)
$\tau_{6,female}$	-0.103	0.0782	-1.000
	(0.368)	(0.140)	(7.473)
$\tau_{7,male}$	-0.190	-0.158	-8.908***
	(0.272)	(0.123)	(2.545)
$\tau_{7,female}$	0.0327	0.0287	-4.353
	(0.419)	(0.162)	(9.575)
$\tau_{8,male}$	-0.164	-0.134	-14.24***
	(0.280)	(0.137)	(4.091)
$\tau_{8,female}$	-0.118	-0.0580	-3.851
	(0.502)	(0.198)	(10.75)
$\tau_{9,male}$	-0.325	-0.234	-15.14***
	(0.355)	(0.153)	(3.189)
$\tau_{9,female}$	0.0985	-0.0212	2.744
	(0.579)	(0.206)	(14.55)
$\tau_{10,male}$	-0.501	-0.308*	-14.78***
	(0.398)	(0.185)	(3.500)
$\tau_{10,female}$	0.417	0.129	12.19
	(0.902)	(0.334)	(10.67)
τ_{-1}	0.197	0.0566	-4.095
	(0.275)	(0.125)	(2.586)
τ_{-2}	0.159	0.0270	-3.245
	(0.247)	(0.117)	(2.100)
τ_{-3}	0.120	0.0330	-1.643
	(0.204)	(0.0972)	(1.338)
Observations	1,285	1,285	1,237
Pre-trends	.9125	.8882	.3926
Individual FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Starting Year FE	Yes		
Treated ind.	84	84	84
Sample	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution in a given year and field. The imputation sample excludes never-treated researchers.

Table C.4: ATT Effect by Sex Excluding the Focal Paper

VARIABLES	(1) Ln(Citations + 1)	(2) h-index	(3) Core number
τ_{male}	0.0473	-0.0406	-0.482
	(0.166)	(0.0757)	(2.813)
τ_{female}	0.122	0.0658	-2.673
	(0.269)	(0.103)	(4.994)
τ_{-1}	0.197	0.0566	-4.095
	(0.275)	(0.125)	(2.586)
τ_{-2}	0.159	0.0270	-3.245
	(0.247)	(0.117)	(2.100)
τ_{-3}	0.120	0.0330	-1.643
	(0.204)	(0.0972)	(1.338)
Observations	1,298	1,298	1,250
Pre-trends	.9125	.8882	.3926
Individual FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Starting Year FE	Yes		
Treated ind.	84	84	84
Sample	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The coefficients represent the average ATT during the first ten years after the event. The event is defined as collaborating with a prominent researcher within the first five years of the authors' academic career. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution. The focal paper is defined as the paper in which the treated researchers collaborate with a prominent researcher. The imputation sample excludes never-treated researchers.

D Robustness Checks

Table D.1: Effect of Collaborating with a Prominent Researcher within the First Five Years of Academic Career (96th percentile)

VARIABLES	(1) Ln(Citations + 1)	(2) Ln(Citations per paper + 1)	(3) Ln(H-Index + 1)	(4) Ln(Publications + 1)	(5) Core number
$\tau_{0,male}$	0.356***	0.0927*	0.102***	0.315***	3.810***
	(0.0956)	(0.0526)	(0.0377)	(0.0647)	(0.991)
$\tau_{0,female}$	0.397***	0.0783	0.0969**	0.492***	11.95***
	(0.124)	(0.0693)	(0.0481)	(0.0756)	(2.897)
$\tau_{1,male}$	0.476***	0.183***	0.150***	-0.0766	4.378***
	(0.122)	(0.0646)	(0.0469)	(0.0802)	(1.392)
$\tau_{1,female}$	0.774***	0.391***	0.182***	-0.0987	12.62***
	(0.159)	(0.0957)	(0.0578)	(0.0997)	(2.900)
$\tau_{2,male}$	0.399***	0.154**	0.144**	-0.201**	4.989***
	(0.141)	(0.0733)	(0.0559)	(0.0910)	(1.930)
$\tau_{2,female}$	0.688***	0.369***	0.181***	-0.0666	13.52***
	(0.197)	(0.122)	(0.0635)	(0.115)	(3.590)
$\tau_{3,male}$	0.353**	0.135*	0.136**	-0.0169	5.560**
	(0.153)	(0.0760)	(0.0663)	(0.0957)	(2.345)
$\tau_{3,female}$	0.613***	0.374***	0.155**	-0.204*	14.68***
	(0.196)	(0.121)	(0.0719)	(0.121)	(4.386)
$\tau_{4,male}$	0.223	0.0954	0.105	-0.260***	4.188*
	(0.172)	(0.0873)	(0.0722)	(0.0874)	(2.502)
$\tau_{4,female}$	0.532**	0.307**	0.157**	-0.0363	13.17***
	(0.217)	(0.135)	(0.0775)	(0.127)	(4.991)
$\tau_{5,male}$	0.253	0.0833	0.0940	-0.116	2.601
	(0.197)	(0.104)	(0.0802)	(0.105)	(2.127)
$\tau_{5,female}$	0.422*	0.285**	0.0799	-0.397***	14.69**
	(0.225)	(0.133)	(0.0910)	(0.128)	(6.575)
$\tau_{6,male}$	0.200	0.0549	0.0417	-0.159	3.204
	(0.213)	(0.117)	(0.0905)	(0.116)	(2.437)
$\tau_{6,female}$	0.272	0.189	0.0823	-0.369**	10.19
	(0.254)	(0.142)	(0.0957)	(0.170)	(6.764)
$\tau_{7,male}$	0.0955	0.0287	0.0972	-0.286***	3.048
	(0.246)	(0.118)	(0.101)	(0.110)	(2.959)
$\tau_{7,female}$	0.388	0.216	0.164	-0.0361	7.594
	(0.277)	(0.166)	(0.111)	(0.161)	(8.027)
$\tau_{8,male}$	0.172	0.0358	0.0610	-0.0280	5.104
	(0.237)	(0.129)	(0.109)	(0.122)	(4.075)
$\tau_{8,female}$	0.376	0.277	0.0945	-0.334	4.288
	(0.279)	(0.180)	(0.129)	(0.205)	(7.437)
$\tau_{9,male}$	-0.0836	-0.117	0.0517	-0.288**	4.649
	(0.270)	(0.156)	(0.124)	(0.146)	(6.516)
$\tau_{9,female}$	0.342	0.406**	-0.00860	-0.240	4.871
	(0.285)	(0.195)	(0.142)	(0.211)	(8.322)
$\tau_{10,male}$	-0.0307	-0.139	0.0697	-0.188	3.248
	(0.298)	(0.176)	(0.132)	(0.144)	(6.673)
$\tau_{10,female}$	0.383	0.430*	0.0887	-0.349*	25.93
	(0.345)	(0.222)	(0.168)	(0.193)	(16.44)
τ_{-1}	0.301	0.181*	0.0986	0.0939	0.140
	(0.198)	(0.109)	(0.0863)	(0.103)	(1.472)
τ_{-2}	0.313*	0.214**	0.0815	0.0294	-0.484
	(0.185)	(0.104)	(0.0810)	(0.0967)	(1.296)
τ_{-3}	0.144	0.106	0.0512	0.0137	-0.496
	(0.155)	(0.0896)	(0.0682)	(0.0878)	(0.936)
Observations	2,384	2,384	2,384	2,384	2,234
Pre-trends	.9125	.1551	.7059	.6617	.3696
Individual FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Starting Year FE	Yes				
Treated ind.	84	84	84	84	84
Sample	135	135	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 A prominent researcher is defined as a coauthor in the 96th percentile of the coauthors' h-index distribution in a given year and field. The imputation sample excludes never-treated researchers.

Table D.2: Effect of Collaborating with a Prominent Researcher within the First Five Years of Academic Career (97th percentile)

VARIABLES	(1) Ln(Citations + 1)	(2) Ln(Citations per paper + 1)	(3) Ln(H-Index + 1)	(4) Ln(Publications + 1)	(5) Core number
$\tau_{0,male}$	0.414***	0.131**	0.123***	0.311***	4.343***
	(0.114)	(0.0648)	(0.0404)	(0.0802)	(1.315)
$\tau_{0,female}$	0.467***	0.114	0.134***	0.437***	12.85***
	(0.125)	(0.0767)	(0.0479)	(0.0796)	(3.214)
$\tau_{1,male}$	0.575***	0.222***	0.188***	-0.0201	5.213***
	(0.139)	(0.0810)	(0.0510)	(0.0888)	(1.795)
$\tau_{1,female}$	0.788***	0.379***	0.175***	-0.0840	12.63***
	(0.176)	(0.112)	(0.0613)	(0.109)	(3.262)
$\tau_{2,male}$	0.473***	0.187**	0.167***	-0.195*	6.178**
	(0.157)	(0.0923)	(0.0613)	(0.101)	(2.538)
$\tau_{2,female}$	0.768***	0.432***	0.170**	-0.0955	14.46***
	(0.225)	(0.139)	(0.0707)	(0.130)	(4.045)
$\tau_{3,male}$	0.461***	0.191*	0.172**	0.0452	6.419**
	(0.169)	(0.0979)	(0.0687)	(0.0962)	(3.162)
$\tau_{3,female}$	0.711***	0.426***	0.141*	-0.202	15.88***
	(0.213)	(0.132)	(0.0785)	(0.143)	(4.772)
$\tau_{4,male}$	0.280	0.123	0.135*	-0.220**	4.405
	(0.174)	(0.107)	(0.0745)	(0.0873)	(3.301)
$\tau_{4,female}$	0.612**	0.365**	0.167*	-0.0446	14.81***
	(0.246)	(0.150)	(0.0857)	(0.133)	(5.468)
$\tau_{5,male}$	0.302	0.0934	0.128	-0.0461	3.206
	(0.198)	(0.116)	(0.0848)	(0.107)	(2.509)
$\tau_{5,female}$	0.451*	0.286*	0.0730	-0.366**	16.25**
	(0.251)	(0.148)	(0.101)	(0.152)	(7.247)
$\tau_{6,male}$	0.282	0.102	0.0714	-0.154	3.629
	(0.228)	(0.138)	(0.0957)	(0.129)	(2.885)
$\tau_{6,female}$	0.290	0.186	0.0848	-0.441**	11.78
	(0.294)	(0.164)	(0.107)	(0.196)	(7.665)
$\tau_{7,male}$	0.221	0.111	0.115	-0.202*	3.110
	(0.237)	(0.137)	(0.104)	(0.109)	(3.593)
$\tau_{7,female}$	0.383	0.235	0.141	0.0162	12.23
	(0.334)	(0.192)	(0.127)	(0.211)	(9.886)
$\tau_{8,male}$	0.354	0.160	0.106	0.200	6.869
	(0.246)	(0.154)	(0.116)	(0.135)	(5.770)
$\tau_{8,female}$	0.346	0.278	0.0730	-0.247	4.284
	(0.318)	(0.188)	(0.140)	(0.234)	(8.404)
$\tau_{9,male}$	0.0671	0.00223	0.0695	-0.224	8.644
	(0.288)	(0.188)	(0.127)	(0.159)	(10.98)
$\tau_{9,female}$	0.349	0.433**	-0.00878	-0.271	3.500
	(0.307)	(0.182)	(0.141)	(0.221)	(9.590)
$\tau_{10,male}$	-0.0305	-0.0448	-0.0119	-0.175	2.865
	(0.325)	(0.201)	(0.132)	(0.169)	(11.75)
$\tau_{10,female}$	0.255	0.495**	0.0350	-0.500***	29.82*
	(0.384)	(0.231)	(0.162)	(0.192)	(15.90)
τ_{-1}	0.552**	0.282**	0.210*	0.130	0.934
	(0.248)	(0.128)	(0.110)	(0.126)	(1.881)
τ_{-2}	0.503**	0.277**	0.159	0.0769	0.300
	(0.235)	(0.126)	(0.104)	(0.121)	(1.631)
τ_{-3}	0.244	0.133	0.0985	0.109	-0.0387
	(0.196)	(0.112)	(0.0861)	(0.107)	(1.201)
Observations	1,803	1,803	1,803	1,803	1,677
Pre-trends	.9125	.1005	.2197	.6461	.6452
Individual FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Starting Year FE	Yes				
Treated ind.	84	84	84	84	84
Sample	135	135	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. A prominent researcher is defined as a coauthor in the 97th percentile of the coauthors' h-index distribution in a given year and field. The imputation sample excludes never-treated researchers.

Table D.3: ATT Effect by Year (96th percentile)

VARIABLES	(1) Ln(Citations + 1)	(2) Ln(Citations per paper + 1)	(3) Ln(H-Index + 1)	(4) Ln(Publications + 1)	(5) Core number
τ_{male}	0.276*	0.0820	0.110*	-0.0799	4.226**
	(0.143)	(0.0755)	(0.0611)	(0.0581)	(1.647)
τ_{female}	0.502***	0.286***	0.132*	-0.0944	12.42***
	(0.180)	(0.109)	(0.0700)	(0.0953)	(4.156)
τ_{-1}	0.301	0.181*	0.0986	0.0939	0.140
	(0.198)	(0.109)	(0.0863)	(0.103)	(1.472)
τ_{-2}	0.313*	0.214**	0.0815	0.0294	-0.484
	(0.185)	(0.104)	(0.0810)	(0.0967)	(1.296)
τ_{-3}	0.144	0.106	0.0512	0.0137	-0.496
	(0.155)	(0.0896)	(0.0682)	(0.0878)	(0.936)
Observations	2,409	2,409	2,409	2,409	2,247
Pre-trends	.9125	.1551	.7059	.6617	.3696
Individual FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Starting Year FE	Yes				
Treated ind.	84	84	84	84	84
Sample	135	135	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution in a given year and field. The imputation sample excludes never-treated researchers.

Table D.4: ATT Effect by Year (97th percentile)

VARIABLES	(1) Ln(Citations + 1)	(2) Ln(Citations per paper + 1)	(3) Ln(H-Index + 1)	(4) Ln(Publications + 1)	(5) Core number
τ_{male}	0.349**	0.129	0.126**	-0.0319	4.982**
	(0.154)	(0.0924)	(0.0645)	(0.0542)	(2.149)
τ_{female}	0.527***	0.310**	0.126*	-0.109	13.44***
	(0.204)	(0.123)	(0.0752)	(0.113)	(4.670)
τ_{-1}	0.552**	0.282**	0.210*	0.130	0.934
	(0.248)	(0.128)	(0.110)	(0.126)	(1.881)
τ_{-2}	0.503**	0.277**	0.159	0.0769	0.300
	(0.235)	(0.126)	(0.104)	(0.121)	(1.631)
τ_{-3}	0.244	0.133	0.0985	0.109	-0.0387
	(0.196)	(0.112)	(0.0861)	(0.107)	(1.201)
Observations	1,822	1,822	1,822	1,822	1,686
Pre-trends	.9125	.1005	.2197	.6461	.6452
Individual FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Starting Year FE	Yes				
Treated ind.	84	84	84	84	84
Sample	135	135	135	135	135

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution in a given year and field. The imputation sample excludes never-treated researchers.

Table D.5: Effect of Collaborating with a Prominent Researcher (Alternative Control Group)

	(1)	(2)	(3)	(4)
VARIABLES	Ln(Citations + 1)	Ln(Citations per paper +1)	Ln(H-index+1)	Core
$\tau_{0,male}$	0.469***	0.332**	0.131***	1.868
	(0.0908)	(0.159)	(0.0317)	(1.196)
$\tau_{0,female}$	0.425***	0.280**	0.116**	1.156
	(0.133)	(0.132)	(0.0507)	(0.887)
$\tau_{1,male}$	0.579***	0.0275	0.137***	1.620
	(0.102)	(0.109)	(0.0438)	(1.536)
$\tau_{1,female}$	0.646***	0.631***	0.159***	1.544
	(0.179)	(0.213)	(0.0531)	(0.984)
$\tau_{2,male}$	0.475***	0.381*	0.152***	2.222
	(0.134)	(0.229)	(0.0510)	(2.093)
$\tau_{2,female}$	0.634***	0.480**	0.175***	3.101*
	(0.222)	(0.191)	(0.0621)	(1.833)
$\tau_{3,male}$	0.644***	0.172	0.148**	4.386
	(0.126)	(0.202)	(0.0624)	(3.188)
$\tau_{3,female}$	0.745***	0.0745	0.208***	3.976*
	(0.197)	(0.0896)	(0.0771)	(2.327)
$\tau_{4,male}$	0.535***	0.258	0.156**	4.705
	(0.148)	(0.217)	(0.0681)	(3.443)
$\tau_{4,female}$	0.776***	0.596***	0.244***	4.608
	(0.240)	(0.190)	(0.0797)	(2.818)
$\tau_{5,male}$	0.671***	0.175	0.186**	9.718*
	(0.159)	(0.328)	(0.0727)	(5.619)
$\tau_{5,female}$	0.841***	0.532	0.274***	6.136*
	(0.239)	(0.345)	(0.0932)	(3.484)
$\tau_{6,male}$	0.388**	-0.0397	0.00131	1.978
	(0.197)	(0.418)	(0.0678)	(3.500)
$\tau_{6,female}$	0.746**	-0.300**	0.277**	8.178*
	(0.296)	(0.121)	(0.108)	(4.534)
$\tau_{7,male}$	0.279	0.661**	-0.00164	-1.515
	(0.203)	(0.332)	(0.0953)	(1.003)
$\tau_{7,female}$	0.905***	-0.412***	0.220*	7.498
	(0.342)	(0.141)	(0.123)	(5.627)
$\tau_{8,male}$	0.463**	0.0623	0.0961	-1.709
	(0.217)	(0.403)	(0.115)	(1.194)
$\tau_{8,female}$	0.836*	-0.221	0.168	8.505
	(0.437)	(0.244)	(0.166)	(7.137)
$\tau_{9,male}$	0.390	0.392	0.0156	-1.059
	(0.292)	(0.485)	(0.141)	(1.125)
$\tau_{9,female}$	1.126**	0.261	0.224	17.90
	(0.527)	(0.453)	(0.181)	(11.75)
$\tau_{10,male}$	0.404	-0.385**	0.00400	-2.153
	(0.343)	(0.183)	(0.170)	(2.143)
$\tau_{10,female}$	1.391	-0.258	0.313	33.21***
	(0.863)	(0.517)	(0.305)	(3.575)
τ_{-1}	0.505**	-0.0109	0.114	0.217
	(0.215)	(0.150)	(0.106)	(0.999)
τ_{-2}	0.328	0.0838	0.0494	-0.0242
	(0.208)	(0.137)	(0.102)	(0.897)
τ_{-3}	0.199	0.00458	0.0550	0.0834
	(0.177)	(0.119)	(0.0859)	(0.686)
Observations	13,934	8,449	13,932	12,995
Pre-trends	0.03059	0.82248	0.31625	0.84175
Individual FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Treated ind.	135	135	135	135
Sample	1516	1516	1516	1516

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. A prominent researcher is defined as a coauthor in the 98th percentile of the coauthors' h-index distribution in a given year and field. The imputation sample only includes never-treated researchers.