Gender Homophily and Unequal Access to Elite Collaborators in Economics

Patricia Cortés Randall P. Ellis Linh T. Tô

April 2026*

Abstract

We document substantial gender homophily in economics coauthorship using data on over 2,800 faculty in the top 100 U.S. economics departments and their publications from 2010–2023. Female economists collaborate with men in only 70 percent of their coauthorship pairs—10 percentage points below random-matching benchmarks—and these patterns persist after controlling for field, cohort, and training. Homophily is amplified in high-stakes settings: men exhibit stronger same-gender sorting in top-five journal publications. Homophily generates gender gaps in access to elite collaborators, defined here as authors in the top 10 ranked departments or top decile of publications. Moreover, women's elite same-gender network is structurally fragile: removing just ten highly connected elite women reduces women's access to elite female coauthors by nearly 70 percent, whereas the analogous exercise for men has little effect. A simple model shows that symmetric same-gender preferences can produce asymmetric outcomes when combined with a male-dominated elite; no discrimination is required. A bounding exercise suggests that differential access to elite coauthors can account for up to 28 percent of the gender gap in top-five publications.

*Acknowledgments: The authors are grateful for financial support for this research from the Social Science Research Council (SSRC) and Committee on the Status of Women in the Economics Profession (CSWEP). We would also like to thank research assistants and data entry assistance from Camilo Rios Bernal, Aaliyah Lee, Changwoo Lee, Liqiang Liu, Maryam Rahamian, Yilin Wang, and Lixiang Zheng.

1 Introduction

Research in economics is increasingly produced through collaboration. Over the last several decades, the share of coauthored papers has risen sharply (Card and DellaVigna, 2013), and the profession has developed a highly interconnected collaboration network (Goyal et al., 2006). As a result, publication success increasingly depends on the ability to form productive research teams. At the same time, the economics profession remains highly gender imbalanced. Women represent roughly one quarter of academic economists and are substantially underrepresented among senior faculty and top publishers. A large literature documents gender gaps in productivity, promotion, and representation (e.g., Sarsons et al., 2021; Card and DellaVigna, 2020), yet we know considerably less about how the structure of collaboration networks themselves may contribute to these disparities.

This paper examines whether gender homophily in coauthorship—systematic differences in collaboration patterns by gender—operates as a structural mechanism that can disadvantage women in economics. In a profession where men represent a large majority, even modest same-gender affinity in collaboration choices can mechanically restrict the opportunities available to women. Importantly, these disparities do not require discrimination or asymmetric preferences. Even when men and women share identical homophilous tendencies, homophily interacting with a male-dominated elite can generate unequal access to elite collaborators, which we define here as authors in the top decile of publishing productivity or working at top-ranked departments. We further show that women's collaborations with elite female economists are concentrated among a small number of individuals, making this channel of elite access thin and potentially fragile.

We assemble a new dataset covering more than 2,800 faculty members employed in the top 100 U.S. economics departments and link them to their complete publication histories between 2010 and 2023. We further construct an extended universe of coauthors that includes collaborators of our main sample in business schools, policy institutions, and non-U.S. universities. This allows us to map collaboration networks surrounding economists in top U.S. economics departments and to measure gender mixing at the level of author–coauthor–paper pairs, capturing both one-off and repeated collaborations.

Our first contribution is to provide comprehensive descriptive evidence on the magnitude of gender homophily in economics, benchmarked against transparent random-matching counterfactuals that account for gender composition. Female economists collaborate with men in only about 70 percent of their coauthorship pairs—about 10 percentage points below what would be expected by chance. These patterns are remarkably stable after conditioning on cohort, field specialization, doctoral training, and institutional rank, and they persist across cohorts, fields, and department rankings.

Second, we show that gender homophily is particularly consequential in high-stakes environments. Collaboration patterns in top-five journals—the primary currency for hiring, tenure, and promotion in economics—are more gender segregated than collaborations in other outlets. Using within-author comparisons, we document that men, but not women, exhibit stronger same-gender sorting in their top-five publications, suggesting that homophily is amplified precisely where career returns are highest. This need not reflect stronger preferences per se; it may also arise if collaborations involving male researchers are, on average, more likely to culminate in top-five publications.

Third, we examine the implications of homophily for access to elite collaborators, which we define as either in the top decile of research productivity or at a top 5 (or top 10) economics department. Because these elite collaborators are disproportionately male, homophily can mechanically limit women's exposure to high-impact collaborators. We document that female economists are significantly less likely to collaborate with top-department faculty or with scholars in the top decile of the productivity distribution. These gaps arise primarily on the extensive margin—whether an author ever collaborates with an elite scholar. Decomposing by the gender of the elite coauthor shows that the gap is driven almost entirely by differential access to male elites: women are significantly less likely to collaborate with elite men, while their access to elite women is similar or slightly higher. Men's elite collaborations are overwhelmingly with male elites, whereas women draw relatively more on the much smaller pool of elite female economists.

Fourth, we show that women's access to elite female collaborators is not only limited but structurally thin. Collaborations between female economists and elite women are highly concentrated among a small number of individuals, whereas men's elite same-gender collaborations are spread across a much larger pool. A counterfactual hub-removal exercise illustrates this difference in network structure: removing the ten most-connected elite women reduces women's probability of having any elite female coauthor by nearly 70 percent, while the analogous exercise for men reduces their elite same-gender access by only 11 percent. More importantly, removing elite women reduces women's overall access to elite collaborators while leaving men largely unaffected, whereas removing elite men reduces access for both genders. Women's networks are thus exposed to shocks from both sides of the elite network, whereas men's are largely insulated from the loss of elite women. This fragility suggests that the gender gap in elite access is not simply a matter of numbers, but reflects a different network architecture in which women's connections to elite collaborators depend more heavily on a small set of individuals.

To interpret these patterns, we develop a simple model of collaboration choice with homophily. The model illustrates how symmetric same-gender affinity, combined with a male-dominated elite, can generate asymmetric outcomes in access to elite collaborators and expected research output. The model is intentionally stylized and is not meant to identify causal mechanisms; rather, it provides a transparent framework that clarifies how network structure alone can amplify inequality.

Finally, we use the empirical relationship between elite collaboration and productivity to conduct a bounding exercise that quantifies how much of the observed gender gap in top-five publications can plausibly be attributed to homophily in elite coauthorship. Interpreted as an upper bound, the exercise suggests that homophily can account for a meaningful—but not complete—share of the gender gap in high-impact publications, indicating that collaboration networks are an important, though not exclusive, contributor to persistent disparities.

Taken together, our results highlight the role of collaboration networks as a structural feature shaping inequality in scientific careers. A central finding is that women's access to elite collaborators depends disproportionately on a small set of highly connected individuals. Because connections to elite women are highly concentrated, the departure of even a few such individuals could substantially reduce access for other women. Conversely, policies that expand the pool of elite women or facilitate new collaborations may have disproportionate effects by broadening these pathways.

This paper contributes to the literatures on gender inequality, networks, and the science of science by documenting gender homophily in economics and its implications for access to elite collaborators and top-journal publications.

Early work examined whether economists tend to collaborate with coauthors of the same gender. Boschini and Sjögren (2007) test the hypothesis of gender-neutral collaboration and reject it, showing that coauthorship patterns differ systematically by gender. Ferber and Teiman (2006) emphasize the role of "old boy" networks in shaping women's career outcomes. More recent work documents gender differences in economists' collaboration networks and relates these differences to research productivity (Ductor et al., 2023), while other work provides evidence that gender homophily is present in economics coauthorships (Ductor and Prummer, 2024).

More broadly, our findings relate to the growing "science of science" literature showing that assortative collaboration patterns shape scientific impact (Freeman and Huang, 2015), to theoretical work demonstrating how homophily and minority status can jointly generate segregation in networks (Currarini et al., 2009), and to evidence that network proximity influences the formation of research collaborations (Fafchamps et al., 2010). Finally, by documenting unequal access to elite collaborators and linking these gaps to high-stakes publication outcomes, our paper complements work on team formation and productivity in economics (Anderson and Richards-Shubik, 2022) and evidence that the loss of a superstar collaborator can have lasting productivity consequences (Azoulay et al., 2010).

2 Data and Variable Construction

2.1 Faculty Sample Construction

Our primary sample consists of more than 3,200 faculty members employed in the top 100 U.S. economics departments as ranked by RePEc in May 2023. We assembled faculty lists by scraping department websites in 2023, supplemented by the 2020–2021 Economics Faculty Directory (Hasselback). We excluded lecturers, visitors, emeritus faculty, and courtesy appointments. For each scholar, we collected: (i) name and institutional affiliation, (ii) current academic rank, (iii) PhD-granting institution and PhD year, and (iv) faculty photo.

We match each faculty member in our sample to all publications in economics journals indexed in Web of Science (WoS) between 2010 and 2023. Approximately 575 faculty initially could not be matched to any WoS-indexed publication, largely due to name spelling discrepancies between department rosters and bibliographic records. For these cases, we supplemented the matching with OpenAlex, a comprehensive open-access bibliographic database, recovering publication records for roughly 250 authors. After this augmentation, 325 faculty remain without any matched publication and are excluded from all analyses. Of these, approximately 50 are recent graduates (PhD 2019 or later) who may not yet have published; the remaining 275 are likely name-matching failures. The unmatched faculty are disproportionately female (31 percent, compared with 20 percent among matched faculty), a gap that is especially pronounced among recent graduates (50 percent female). One natural explanation for the higher female non-match rate is surname changes after marriage, which complicate bibliographic matching. However, the gender gap is largest among younger cohorts, where name changes are less common, suggesting that this explanation is incomplete. This pattern, if anything, would lead our estimates to understate the true extent of gender sorting. Our final analysis sample consists of 2,812 faculty members with at least one matched publication with a coauthor.

For every publication, we record the names and institutions of the coauthors. Only 17% of the distinct coauthors of our core faculty appear in the top-100 universe. The remaining 83% originate from other institutions (business schools, non-U.S. economics departments, government, private firms, and multilateral institutions).

2.2 Gender Assignment

We assign gender to faculty in the core sample using profile photographs reviewed by humans, supplementing with web pronoun references when uncertain. For coauthors outside the core faculty sample, we assign gender using the Namesor algorithm, classifying a name as female when the predicted probability of being female exceeds 0.5, and analogously for male. Most name–gender probabilities are well above this threshold.

2.3 Field Assignment

We assign our core sample scholars to fields to explore the significance of women and men being unevenly distributed across fields. Because field assignment is non-trivial, we rely on several complementary sources: the journals in which authors publish, NBER working-paper field codes, an AI-based field classifier, and dissertation field information, and examine both single- and multiple-field assignments. The full procedure is described in the Appendix.

2.4 Construction of Main Variables

Because papers often involve multiple authors and repeated collaborations with the same coauthor, we measure gender mixing at the level of *author–coauthor–paper* pairs. For each focal author i, every paper p on which i appears generates one coauthorship pair for each distinct coauthor on that paper. Thus, if author i writes two papers with the same coauthor j, this contributes two coauthorship pairs to i's portfolio, not one. This construction ensures that repeated collaborations receive more weight than one-off interactions.[1]

Formally, let $\mathcal{P}_i$ denote the set of papers authored by i, and let C_{ip} denote the set of

[1]As a robustness check, we also measure gender mixing at the level of unique coauthors rather than author–coauthor–paper pairs; results are very similar.

coauthors of i on paper $p \in \mathcal{P}_i$. The total number of coauthorship pairs for author i is then

$$N_i = \sum_{p \in \mathcal{P}_i} |C_{ip}|.$$

Our main statistic summarizing the gender composition in i's collaborations is the share of these coauthorship pairs that involve a male coauthor:

$$\text{ShareMale}_i = \frac{\sum_{p \in \mathcal{P}_i} \sum_{j \in C_{ip}} \mathbf{1}\{j \text{ is male}\}}{\sum_{p \in \mathcal{P}_i} |C_{ip}|}. \tag{1}$$

This statistic is convenient for two reasons. First, it offers a transparent characterization of each author's collaboration network, independent of the number of papers or team size. Second, it maps cleanly into a random-matching benchmark, which allows us to quantify the extent to which the observed gender composition of collaborations departs from what would arise mechanically given the population of potential coauthors.

In our data, 58% of coauthorship pairs involve a coauthor who is also in the sample, meaning that both members of the ordered pair are faculty in top-100 economics departments. However, only 17% of all distinct coauthors of the core sample are themselves in the top 100. This discrepancy is partly mechanical: collaborations between two scholars in our sample are recorded twice in our dataset. But even after accounting for this double counting, collaborations among top-100 faculty are substantially more likely to persist across multiple papers than collaborations with coauthors outside the top-100 universe.

2.5 Descriptive Statistics

Table 1 presents summary statistics for the faculty sample. Women represent 20% of the 2,812 faculty in our sample (572 women, 2,240 men). Because the share of women in economics has risen over time, female faculty have, on average, more recent PhD graduation years than male faculty—an average gap of about 6 years. This cohort difference is partly reflected in rank composition: 35% of women are assistant professors compared to 22% of men, while only 39% of women are full professors compared to 57% of men.

Turning to academic background, women are employed at lower-ranked departments on average: 22% of women work in top-20 departments compared to 30% of men.

Women and men also differ in their field distribution. Women are overrepresented in Labor, Health, Education, and Public Economics and underrepresented in Theory, Macro, Econometrics, Industrial Organization, and Finance. These differences motivate the inclusion of field fixed effects in subsequent analyses.

In terms of research output, women publish fewer papers on average (4.8 vs. 6.0) and fewer top-five journal articles (0.7 vs. 1.1). They also have fewer unique coauthors (8.7 vs. 11.1), reflecting both lower publication counts and smaller collaboration networks.[2] Importantly, these productivity differences are not solely driven by cohort, field, or training: as shown in Table A.1, a significant gender gap in publications persists after controlling for PhD year, PhD institution rank, and field fixed effects.

Finally, the unconditional share of male coauthorship pairs is substantially lower for women than for men. To interpret these descriptive patterns, we next compare the observed gender composition of collaborations to the benchmarks implied by random matching.

3 Evidence on Gender Homophily in Coauthorship

To quantify the extent of gender sorting in collaborations, we compare the observed ShareMale values to those implied by random matching. Random-matching benchmarks provide the relevant counterfactual because they describe the gender composition of collaborations that would arise purely from the population of potential coauthors, absent any systematic gender preferences or structural sorting.

Let s_m denote the share of male scholars in a relevant population of potential collaborators. Under random matching, each coauthorship pair is formed independently of gender, so the expected male share of an author's coauthorship pairs equals s_m, regardless of the author's gender.

In our core sample, 20% of scholars are women, so under random matching the expected share of male coauthors would be 80% for authors of both genders.[3]

The observed gender composition of coauthorship pairs differs from this random-matching benchmark. Female economists collaborate with men in only about 70% of their

[2] Solo-authored papers account for a similar share of output among women (9.5%) and men (8.3%; $p = 0.34$), so differential selection into coauthorship is unlikely to drive our results.

[3] Because men generate more coauthorship pairs on average than women, the aggregate population of pair-level observations is more male-intensive than the population of scholars. If one weights potential coauthors by their productivity, the expected male share rises to approximately 84%. We use the simpler 80% benchmark throughout, which treats each scholar as equally likely to be a potential collaborator.

coauthorship pairs—10 percentage points below the 80% implied by random assignment. Male economists display less homophily, but note that their deviation from the benchmark is mechanically constrained by the fact that the male share cannot exceed 100%.

Figure 1 shows the full distributions of share male coauthorship pairs separately for men and women, along with vertical lines indicating the gender-specific means and the 80% random-matching benchmark. The two distributions exhibit limited overlap: women's collaboration portfolios are more female-intensive than would be expected under random matching.

Figure A.2 shows this pattern across fields. The bars report the male share of scholars in each field—the benchmark implied by random matching. If collaborations were random within fields, both men and women would have the same share of male coauthors as this benchmark. Instead, women's share of male coauthors (circles) predominantly falls below the benchmark, while men's (squares) is at or above it. The gender coauthorship gap is larger in three fields with the highest female representation (Health, Education, Labor), where opportunities for same-gender collaboration are greater, and in Public. The gap is smallest in International, Econometrics, and Theory.

3.1 Regression Evidence Controlling for Composition

The random-matching benchmark accounts for aggregate gender shares and productivity but does not address compositional differences across fields, cohorts, or training backgrounds. To assess whether these factors explain the observed homophily, we estimate

$$\text{ShareMale}_i = \alpha + \beta \cdot \text{Female}_i + \gamma_f + \delta_c + \theta_p + \varepsilon_i,$$

where γ_f are field fixed effects, δ_c are PhD-cohort fixed effects, and θ_p controls for PhD-institution rank.[4]

Across all specifications, the coefficient on Female_i remains large, negative, and remarkably stable (Table 2). Adding controls for current department rank or professorial rank produces similarly stable results, though these controls are potentially endogenous. Results

[4] We classify authors into 15 fields based on their publication record: Macro, International, Finance, Labor, Theory, Econometrics, Industrial Organization, Public Economics, Development, Health, Education, Environmental, Urban, Political Economy, and Economic History. Field is assigned hierarchically using journal-based classification, NBER program codes, dissertation field, and an LLM-based classification of publications (see Appendix A.5). PhD-year fixed effects are individual PhD-year dummies (spanning 1955–2022). PhD-institution rank has five categories: top 5, top 6–10, top 11–30, rank 31+, and foreign PhD.

are similar when measuring gender composition at the level of unique coauthors rather than coauthorship pairs (Table A.2, Column 1). Moreover, homophily is present both in the formation of new collaborations and in repeated collaborations, and is in fact stronger for repeated ties (Table A.2, Columns 3–4).

A natural concern is that men and women may specialize in different topics *within* fields, and that topic-level sorting could generate apparent homophily even absent same-gender preferences. To assess whether such unobservables could explain our results, we apply Oster (2019) bounds. As shown in Table A.3, unobservables would need to be more than twice as important as observables to eliminate the gender gap in share male coauthors.

3.2 Market-Adjusted Benchmark

The additive controls in Table 2 do not fully absorb the gender composition of each author's collaboration "market." To address this, we construct an author-specific benchmark: for author i in market m, the leave-one-out expected share of male coauthors is

$$\text{ExpectedShareMale}_i = \frac{\text{Males in } m - \mathbf{1}[\text{male}_i]}{N_m - 1},$$

where the leave-one-out construction ensures that author i does not contribute to her own benchmark. We then define Excess Share Male as the difference between observed and expected: $\text{ExcessShareMale}_i = \text{ShareMale}_i - \text{ExpectedShareMale}_i$. A negative female coefficient on this outcome indicates that women coauthor with fewer men than expected *given the gender composition of their market*.

Table A.4 reports the results across three market definitions: field only as a broad definition (15 cells), field × PhD institution rank (top 10 vs. rest) × cohort (pre- vs. post-2003) as the primary definition (~60 cells), and field × cohort (4 groups) × PhD rank (5 groups) as a narrow definition (~300 cells). The female coefficient remains negative and highly significant in all specifications, confirming that women collaborate with fewer men than expected given the gender composition of their market cell. Results are similar when the expected share is computed weighting each author by their number of unique coauthors rather than equally.

Figure 2 provides a visual counterpart to these regressions. Using the primary market definition, we plot each author's observed share of male coauthors against the leave-one-out expected share implied by their market cell. Under random matching within markets, all points would lie along the 45-degree line. Instead, men cluster above the line (more male coauthors than expected) and women cluster below it (fewer male coauthors than expected), with a clear and consistent gap across the full range of market gender compositions.

A potential concern with the regression-based inference above is that standard errors assume independent and identically distributed errors across authors. Because authors within the same market cell share a common pool of potential coauthors, this assumption may be violated. To address this, we implement a conditional permutation test that directly evaluates whether the observed gender sorting exceeds what would arise by chance within markets, without relying on parametric distributional assumptions. Figure A.1 shows the permutation distribution for the primary market definition. The observed coefficient lies far in the left tail of the distribution, well beyond the 2.5th percentile, confirming that the within-market gender gap is not an artifact of market composition.

3.3 Heterogeneity in Homophily Across Cohorts, Fields, and Department Ranks

Gender homophily is present across most segments of the profession (Tables A.5 and A.6). The gender gap is significant for all PhD cohorts from 1990 onward, but is small and not statistically significant for the oldest cohort (PhD 1955–1989).[5] Homophily varies somewhat across department ranks. The gender gap is largest at mid-ranked departments (ranks 21–45) and smaller at lowest ranked institutions. However, a joint test of Female $\times$ rank interactions fails to reject equality across rank groups. Field-level heterogeneity is more substantial: the female coefficient ranges from -0.16 in Health/Education/Public economics to -0.04 in Theory. Homophily is strongest in fields with higher female representation, consistent with same-gender sorting being more pronounced when there are more opportunities for it (Figure A.3).

3.4 Homophily in High-Stakes Collaborations: Top-Five Publications

Because publication in top-five journals is central to tenure and promotion in economics, collaboration patterns on these publications provide insight into whether homophily is especially strong when returns are highest. To assess this, we restrict the sample to authors who have at least one top-five and one non–top-five publication and compute ShareMale

[5] Because cohort and age are mechanically related in our cross-sectional data, we cannot distinguish cohort effects from age or career-stage effects.

separately for each subset. We report both cross-sectional comparisons across authors and within-author estimates using the author–fixed-effects model

$$\text{ShareMale}_{it} = \alpha + \beta \cdot \text{Top5}_t + \beta_1(\text{Top5}_t \times \text{Female}_i) + \mu_i + \varepsilon_{it},$$

where Top5_t indicates whether publication type t is a top-five journal and μ_i absorbs all time-invariant characteristics of author i.

Table 3 confirms that homophily is amplified in high-stakes collaborations. In the cross-sectional comparison (Columns 1–2), the gender gap in share male is larger in top-five publications (−0.155) than in non-top-five publications (−0.101), with the difference statistically significant at the 10 percent level ($p = 0.084$). The within-author estimates (Columns 3–4) reinforce this pattern. For male authors, the share of male coauthors increases by 3.4 percentage points in top-five papers relative to non–top-five papers, a statistically significant effect. For female authors, however, the interaction term is negative but not statistically significant, and the implied total effect of top-five publication on share male is close to zero. Thus, men exhibit amplified homophily in their highest-stakes collaborations, while women do not.

We have established that gender homophily is pervasive and, for men, especially pronounced in high-stakes collaborations. In the next section we explore whether these patterns translate into systematic gender differences in access to the most influential economists in the profession.

4 Access to Elite Collaborators

In economics, collaborations with highly productive researchers and with faculty in top-ranked departments are particularly valuable: they shape project visibility and increase the likelihood of publication in elite journals. At the same time, the upper tail of the productivity and prestige distribution is disproportionately male. Women constitute only 15% of faculty in top-5 departments and 15% of top-10 department faculty—below their 20% share in the full sample ($p < 0.01$)—and just 13% of scholars in the top productivity decile ($p < 0.01$). In such an environment, even modest homophily may mechanically restrict women's access to elite collaborators.

4.1 Overall Access to Elite Collaborators

We examine whether women are less likely to collaborate with elite economists, defined by institutional prestige (faculty in top-5 or top-10 departments) or research productivity (top decile of top-5 publications). For each author, we construct both an extensive-margin indicator (any elite coauthor) and a count of unique elite coauthors, then regress these on Female with field, cohort, and PhD-rank controls.

Table 4 Panel A shows substantial gender gaps. Women are 4.5 percentage points less likely to have any elite coauthor (top-5 definition), representing a reduction of about 20% relative to the sample mean. Unconditional counts tell a similar story: women have significantly fewer unique elite coauthors across all definitions. When we condition on having at least one elite coauthor, the coefficients become smaller and mostly insignificant, suggesting that the gender gap operates primarily on the extensive margin—whether an author connects with any elite collaborator—rather than the intensive margin of accumulating additional elite connections.

4.2 Access by Gender of the Elite Collaborator

Decomposing by the gender of the elite coauthor reveals the mechanism. The gap is driven entirely by differential access to male elites (Panel B): women are 6–7 percentage points less likely to have any male elite coauthor. In contrast, women are equally or slightly more likely to collaborate with female elites (Panel C). Men's elite collaborations are overwhelmingly with other men, while women partially compensate by drawing on the scarce pool of elite women—but this compensation is incomplete, leaving women with systematically lower overall access.

4.3 Concentration of Elite Same-Gender Collaborations

Panel C of Table 4 shows that women partially compensate for reduced access to elite men by collaborating with elite women at comparable or higher rates. However, because the pool of elite women is small, this compensation may rely on a limited number of individuals. If so, women's access to elite same-gender collaborators may be structurally thin—concentrated among a small set of elite women and therefore vulnerable to the presence or absence of these individuals.

Table 5 examines the structure of these collaborations. For each definition of elite status, we identify every elite coauthor by name, count the number of unique partnerships each elite individual forms with authors of the relevant gender, and compute the Herfindahl index (HHI) of concentration across elite coauthors. A higher HHI indicates that collaborations are concentrated among fewer individuals.

Panel A shows that women's collaborations with elite female coauthors are highly concentrated. Under the top-5 department definition, only 20 elite women form partnerships with female authors in our sample, generating 37 unique author–coauthor pairs. The HHI is 0.081, implying an equivalent number of equally connected collaborators of just 12. The 10 most connected elite women account for 73% of all partnerships. Similar patterns emerge under the top-10 department and top-productivity definitions.

Panel B provides a benchmark. Male authors draw on 162 elite men in top-5 departments, forming 829 unique partnerships. The HHI is 0.010—an order of magnitude smaller—with an equivalent number of 103 collaborators. The top 10 elite men account for only 19% of partnerships. Men's elite same-gender collaborations are therefore spread across a much larger and deeper pool of individuals.

Panel C shows that the high concentration is specific to the elite female pool. When women collaborate with elite *male* coauthors, concentration drops sharply (HHI = 0.017 under the top-5 definition), much closer to the pattern observed for men.

These patterns imply that women's elite same-gender collaboration network is thin. A small number of elite women serve as hubs connecting many female collaborators. Figure 3 provides a visual summary. The unconditional Lorenz curve (Panel A), which includes all authors regardless of whether they have any elite coauthor, shows extreme compression toward the lower right. The conditional Lorenz curve (Panel B), restricted to authors with at least one elite coauthor, reveals that among participants, women's elite female partnerships are distributed nearly equally (Gini = 0.121), while men's elite male partnerships are more concentrated (Gini = 0.326). The bottleneck for women is not unequal access *among* those who reach elite women, but rather that so few women reach them at all.

Table 6 quantifies the structural dependence through a hub-removal counterfactual. We rank elite coauthors by the number of unique partnerships they maintain and sequentially remove the most connected.

Under the top-5 definition, removing the 3 most-connected elite female coauthors causes the probability that a female author has any elite female coauthor to fall by 41%. Removing the top 10 eliminates 69% of women's elite female access. Yet this dramatic collapse in same-gender elite ties has a smaller effect on women's *overall* elite access, which falls by

9%—because most of women's elite access runs through men.

The critical finding emerges from comparing who is affected when each gender's hubs are removed. When top elite women are dropped, women's overall elite access declines by 9%, but men's access is virtually unchanged (−1%). When top elite men are dropped, women's overall access again declines—by 8%—and men's access falls by 9%. Women's elite networks are thus exposed to shocks from *both* sides: losing top elite women or top elite men reduces their access by roughly comparable magnitudes. Men's networks, by contrast, are insulated from the loss of elite women entirely.

This asymmetric dependence is a structural inequality in itself. It implies that women cannot build resilient elite networks through same-gender ties alone—the elite female pool is simply too small. Their access to top collaborators is fundamentally contingent on cross-gender ties to elite men, ties over which homophily works against them.

4.4 Market-Adjusted Elite Access

The raw gender gaps could partly reflect women being in markets with fewer available elites. To address this, we apply two market-adjusted benchmarks (Table A.7). The first uses the same leave-one-out (LOO) adjustment as for Share Male: for each author, we subtract the market-cell average elite access rate. The second computes a random-matching probability, $1 - (1 - q_m)^{k_i}$, where q_m is the fraction of unique coauthors who are elite in market m and k_i is author i's number of unique coauthors. Because k_i is itself an outcome of the collaboration process, this benchmark likely overcontrols—it absorbs the channel through which homophily reduces elite access by limiting collaboration opportunities.

Under the LOO benchmark, the gender gaps are in fact *stronger* after market adjustment: the female coefficient remains negative and significant across all three elite definitions in Panels A and B. Even under the more conservative random-matching benchmark, the gender gap in access to elite male coauthors remains significant across all three definitions (Panel B), with coefficients roughly two-thirds the magnitude of the LOO benchmark.

4.5 Heterogeneity by Male Concentration in the Elite

If homophily drives the gender gap in elite access, the gap should be larger in markets where the elite is more male-dominated. We test this prediction by interacting the female indicator with the male share of the elite in each market, defined as field × PhD institution rank (top 10 vs. rest) × cohort (pre- vs. post-2003). Table A.8 reports the results.

The interaction term is negative and statistically significant across all three definitions of elite status, with the strongest effects for the top productivity decile (-0.244, $p < 0.01$) and top-10 department (-0.173, $p < 0.05$) definitions. These patterns are consistent with the homophily mechanism: when the pool of elite researchers is more male-dominated, same-gender sorting more severely limits women's access.

4.6 Robustness

Table A.9 shows that the gender gap in elite access survives several robustness checks. Column 2 controls for publication volume and the number of unique coauthors, addressing the concern that women may have fewer opportunities to connect with elite scholars due to smaller research output or collaboration networks. Column 3 restricts the sample to assistant professors to abstract from lifecycle differences in collaboration patterns, such as those arising during family formation years. Column 4 excludes authors at top-10 departments to ensure the results are not driven by within-department collaborations with elite colleagues.

Oster (2019) bounds provide further reassurance: the estimated δ for elite access is -1.92 (Table A.3), indicating that adding observables moves the coefficient *away* from zero—so unobservables would need to operate in the opposite direction to fully explain the gap.

5 A Simple Model of Homophily and Elite Collaborators

The empirical patterns documented above raise the question of how homophily interacts with a disproportionately male elite to generate systematic gender differences in collaboration opportunities. To illustrate how symmetric homophily interacts with a male-dominated elite to generate unequal access to elite collaborators, we develop a simple model that isolates the key forces at work, abstracting from strategic behavior and richer dynamics to highlight the minimal ingredients needed to reproduce the observed patterns.

5.1 Setup

The population consists of two groups, men and women, with shares s_M and s_F, where $s_M > s_F$. A fraction q_g of each gender $g \in \{M, F\}$ are elite collaborators, with $q_M > q_F$, reflecting the overrepresentation of men among top economists. Each economist forms k collaborations drawn from a large pool of potential collaborators. In the baseline model we assume that k is the same for all economists. In the calibration below, however, we allow for gender differences in the number of collaborations.

5.2 Homophilous Collaboration Choice

Collaboration choices are governed by a simple homophily parameter $h \in [0, 1]$. For an economist of gender g, each of the k collaborators is chosen independently according to the following rule:

1. With probability h, the collaborator is drawn uniformly from authors of the same gender g;
2. With probability $(1 - h)$, the collaborator is drawn from the overall population in proportion to group sizes.

This reduced-form device captures a taste or affinity for same-gender collaboration without specifying its micro-foundations. It nests random matching as a special case: when $h = 0$, collaborators are chosen entirely in proportion to population shares.

Let $p_{gg'}$ denote the probability that an economist of gender g chooses a collaborator of gender g'. For a male economist ($g = M$), the probability of collaborating with a male is

$$p_{MM} = h + (1 - h)s_M, \tag{2}$$

and the probability of collaborating with a woman is

$$p_{MF} = (1 - h)s_F. \tag{3}$$

Analogously, for a female economist ($g = F$),

$$p_{FF} = h + (1 - h)s_F, \qquad p_{FM} = (1 - h)s_M. \tag{4}$$

When $h > 0$, both groups collaborate with same-gender economists more frequently than under random matching. Because men constitute the majority, however, the scope for observing homophily among men is mechanically limited.

5.3 Access to Elite Collaborators

We next characterize access to elite collaborators. Let E denote the event that a given collaborator is elite. The probability that a randomly chosen collaborator is elite depends on both the gender of the focal author and the gender composition of the elite.

For a male economist, the probability that a given collaborator is elite is

$$\Pr(E \mid M) = p_{MM}q_M + p_{MF}q_F,$$

Similarly, for a female economist,

$$\Pr(E \mid F) = p_{FM}q_M + p_{FF}q_F.$$

Using the expressions for $p_{gg'}$, these probabilities can be written as

$$\Pr(E \mid M) = [h + (1-h)s_M]q_M + (1-h)s_Fq_F,$$

$$\Pr(E \mid F) = (1-h)s_Mq_M + [h + (1-h)s_F]q_F.$$

The expected number of elite collaborators for an economist of gender g is then

$$E[\#E \mid g] = k \cdot \Pr(E \mid g).$$

In the benchmark of random matching ($h = 0$), the probability of collaborating with an elite economist is simply

$$\Pr(E \mid M; h = 0) = \Pr(E \mid F; h = 0) = s_Mq_M + s_Fq_F,$$

so that men and women have the same expected number of elite collaborators.

Proposition 1. *Suppose $q_M > q_F$. Then for any $h > 0$,*

$$\Pr(E \mid M) - \Pr(E \mid F) = h(q_M - q_F) > 0,$$

so that male economists have a higher expected number of elite collaborators than female economists. Moreover, the gap is linearly increasing in h.

Intuition. Homophily shifts collaboration choices toward same-gender partners, altering the gender composition of collaborators. Because the male pool is more elite-intensive

($q_M > q_F$), this shift increases men's exposure to elite collaborators while reducing women's exposure. Notably, the population shares s_M and s_F cancel: the gap depends only on the difference in elite intensity across genders and the strength of homophily. Thus, even symmetric homophily—where men and women share the same homophily parameter h—generates asymmetric access to elite collaborators when the elite pool is male dominated. We corroborate this prediction empirically in Table A.8: the gender gap in elite access is significantly larger in markets where the elite is more male dominated.

Because collaborators are drawn independently, the model's predictions for $\Pr(E \mid g)$ also translate directly to the extensive-margin outcome used in the empirical analysis. The probability that an author has at least one elite collaborator is

$$\Pr(\#E \geq 1 \mid g) = 1 - (1 - \Pr(E \mid g))^k,$$

which is strictly increasing in $\Pr(E \mid g)$ (see Appendix Lemma 1).

5.4 Implications for Productivity and Top Publications

To connect access to elite collaborators with research output, consider a simple reduced-form production function in which an economist's expected output depends on the number of elite collaborators:

$$Y_i = \alpha + \beta \cdot \#E_i + \varepsilon_i,$$

where $\beta > 0$ captures the productivity return to collaborating with elite economists and ε_i is an idiosyncratic shock. This provides a simple channel through which unequal access to elite collaborators can translate into differences in research output.

Taking expectations conditional on gender,

$$E[Y \mid g] = \alpha + \beta k \Pr(E \mid g).$$

By Proposition 1, when $h > 0$ and $q_M > q_F$ we have

$$\Pr(E \mid M) > \Pr(E \mid F),$$

which implies

$$E[Y \mid M] > E[Y \mid F].$$

Thus, even if men and women have identical underlying ability, face the same number of collaboration opportunities k, and share the same homophily parameter h, gender homophily combined with a male-dominated elite can generate differences in expected research output.

Because publication in top journals and cumulative citation counts are convex in researcher quality and collaboration networks, such differences in expected output may translate into disparities in top publications and career advancement.

The framework is also consistent with the empirical finding that high-stakes collaborations exhibit stronger same-gender sorting among men. If the probability that a project achieves a top-five publication is increasing in the number or average quality of elite collaborators, then projects that succeed at top journals will tend to be those formed around the most elite (and therefore most male) collaboration networks.

5.5 Calibrating the Model to Data

We can move beyond qualitative predictions by calibrating the model's key parameter directly from the data. The homophily parameter h equals the within-market gender gap in the share of male coauthors: in the model, male economists choose male collaborators with probability $h + (1 - h)s_M$, while female economists choose male collaborators with probability $(1 - h)s_M$, yielding a difference of exactly h. The regression coefficient on *female* in Column 3 of Table 2—which conditions on field, PhD cohort, PhD institution rank, current department rank, and academic rank—implies $h = 0.088$.

Combined with the observed elite shares ($q_M = 0.084$, $q_F = 0.061$ for top-5 departments) and gender-specific collaboration counts ($k_M = 11.0$, $k_F = 10.0$ conditional on market fixed effects), the model predicts a gender gap in the probability of having at least one elite coauthor of 4.4 percentage points—closely matching the 4.5 pp estimate in Column 1 of Table 4. The model thus replicates a non-targeted moment (the elite access gap) using a parameter calibrated from a different moment (coauthor gender composition).

Decomposing the predicted gap reveals two channels through which homophily operates. Approximately one-fifth of the gap arises from the *sorting channel*: holding the number of collaborations fixed, homophily steers women's coauthorships toward the less-elite female pool. The remaining four-fifths operates through the *extensive margin*: women form fewer collaborations overall. This extensive-margin gap could itself reflect a downstream consequence of homophily in a male-dominated field—if same-gender preferences constrain women to draw from a smaller pool of potential collaborators, they will form fewer collaborations in equilibrium (as in the friendship formation model of Currarini et al., 2009).

Under this interpretation, the full elite access gap could be attributable to homophily; if instead the extensive-margin gap fully reflects other factors, homophily's direct contribution is approximately one-fifth.

The model's qualitative predictions are also consistent with descriptive patterns in the data. In particular, women exhibit lower average publication counts and fewer top-five publications. These differences remain after conditioning on field, cohort, and PhD department rank, as we show in Table A.1.

6 How Much Can Homophily Explain? A Bounding Exercise

The descriptive evidence and the simple model in Section 5 suggest that gender differences in collaboration patterns—particularly in access to elite coauthors—may contribute to gender gaps in research output. In this section, we provide a transparent *upper bound* on the contribution of this mechanism. The exercise attributes all gender differences in elite coauthorship to homophilous matching and interprets the reduced-form relationship between elite collaborators and productivity as causal. Consequently, the estimates below should be viewed as the *maximum* effect implied by our data and framework.

We focus on the extensive margin of elite collaboration: whether an author has at least one coauthor employed in a Top-5 department. As shown in Table 4, women are 4.5 percentage points less likely than men to have such a coauthor, conditional on field, PhD cohort, and PhD institution rank. Meanwhile, conditional on the same controls, men publish 0.31 more Top-5 articles than women (Table A.1). To estimate the productivity return to elite collaboration, we regress the number of Top-5 publications on the elite-coauthor indicator, controlling for gender, field, PhD cohort, and PhD institution rank. Including gender ensures that the coefficient captures the within-gender association between elite coauthorship and productivity, rather than conflating elite access with other gender-related determinants of output. This regression yields a coefficient of 1.93, indicating that authors with at least one Top-5 department coauthor publish nearly two additional Top-5 articles on average.

Under the assumption that the entire gender gap in elite coauthorship is driven by homophily, the implied contribution to the Top-5 publication gap is:

$$\Delta Y^{\text{homophily}} = 1.93 \times 0.045 \approx 0.09.$$

Relative to the residualized gender gap of 0.31 Top-5 publications, homophily in elite coauthorship can account for:

$$\frac{0.09}{0.31} \approx 0.28,$$

or at most 28 percent of the gender gap in Top-5 counts. The model calibration in Section 5 suggests that this upper bound requires attributing both the direct sorting effect of homophily and the gender gap in the number of collaborations—which may itself be a downstream consequence of homophily—to the homophily mechanism. If only the direct sorting channel operates—holding collaboration intensity fixed—the implied contribution falls to approximately 6 percent, which should be interpreted as a conservative lower bound.

A key assumption of this bounding exercise is that the association between elite coauthorship and productivity reflects a causal relationship rather than selection on unobservables. More able economists may both attract elite collaborators and publish more in top journals, which would bias our estimate upward. To assess this concern, we apply Oster (2019) bounds. As shown in Table A.3 Column (3), $\delta = 4.11$ for the productivity regression: unobservables would need to be more than four times as important as observables to fully explain the association between elite coauthorship and top-5 publications. This provides confidence that the productivity return to elite collaboration—the key parameter in our bounding calculation—is not driven by selection. If we take a more conservative approach and use the $\delta = 1$ bound for the effect of elite access on productivity, the implied contribution of homophily to the gender gap in top-5 publications falls to approximately 22 percent.

Taken together, these results indicate that even modest same-gender preferences can generate meaningful disparities in access to elite collaborators and top publications in a male-dominated field, while leaving substantial scope for other forces to contribute to observed gender gaps.

7 Conclusion

This paper documents gender homophily in economics coauthorship and shows how it translates into unequal access to elite collaborators. In a profession where the elite is disproportionately male, even modest same gender sorting generates systematic differences in exposure to high impact collaborators, operating primarily on the extensive margin. A simple model illustrates that these disparities can arise even with symmetric collaboration preferences, and a bounding exercise suggests that this mechanism can account for a meaningful share of the gender gap in top publications.

A key finding is that women's access to elite female collaborators is structurally fragile. Because the number of elite women is small, collaborations with them rely on a limited set of individuals, so removing a few of them sharply reduces women's same gender access. In contrast, men's elite same gender network is deep and diversified. As a result, women cannot build resilient elite networks through same gender ties alone and must rely more heavily on cross gender collaborations, where homophily works against them.

While several mechanisms could generate the patterns we document, including same gender preferences, differences in meeting opportunities, or bias in collaboration decisions, these forces produce similar reduced form outcomes and are difficult to disentangle empirically. Our contribution is to show that asymmetric outcomes do not require asymmetric preferences. Even in the absence of differential treatment across genders, symmetric homophily generates unequal access in a male dominated elite. At the same time, this does not rule out an important role for other forces. Existing evidence suggests that differential credit allocation in mixed gender teams (Sarsons et al., 2021) and changes in social norms affecting cross gender interactions (Gertsberg, 2025) can affect who collaborates with whom and may increase same gender sorting, thereby amplifying the disparities we document.

Our analysis has limitations. First, we cannot cleanly identify a causal effect of homophily on elite access, as we do not exploit exogenous variation in collaboration opportunities, and the rarity of top publications limits what we can learn from dynamic analyses. Second, our field classification captures broad subfields rather than narrower research niches, so some of what we measure as homophily may reflect within field topic specialization. While our robustness checks, including bounds analyses and alternative specifications, mitigate some of these concerns, stronger causal evidence would require data that track the evolution of collaboration opportunities over time, ideally with exogenous variation in access to potential collaborators.

Taken together, our findings have implications for the dynamics of inequality in the profession. Collaboration ties persist, and access to elite collaborators shapes the production of high impact research. If homophily limits women's access today, fewer women will reach elite status tomorrow, further shrinking the pool of potential collaborators for future cohorts. Because women's access depends on a small number of individuals, changes in the composition of the elite may generate disproportionately large spillovers.

These findings point to specific policy levers. Increasing the representation of women among elite economists may have outsized effects by expanding access to high impact collaborations and reducing fragility in collaboration networks. Interventions that facilitate new connections, particularly across gender lines and at early career stages, may help close the extensive margin gap we document. These interventions may help broaden access to elite collaborators and reduce the persistence of inequality in the profession.

References

Anderson, Katharine A. and Seth Richards-Shubik, "Collaborative Production in Science: An Empirical Analysis of Coauthorships in Economics," *Review of Economics and Statistics*, 2022, *104* (6), 1241–1255.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang, "Superstar Extinction," *Quarterly Journal of Economics*, 2010, *125* (2), 549–589.

Boschini, Anne and Anna Sjögren, "Is Team Formation Gender Neutral? Evidence from Coauthorship Patterns," *Journal of Labor Economics*, 2007, *25* (2), 325–365.

Card, David and Stefano DellaVigna, "Nine Facts about Top Journals in Economics," *Journal of Economic Literature*, 2013, *51* (1), 144–161.

_ and _ , "What Do Editors Maximize? Evidence from Four Economics Journals," *Review of Economics and Statistics*, 2020, *102* (1), 195–217.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin, "An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, 2009, *77* (4), 1003–1045.

Ductor, Lorenzo and Anja Prummer, "Gender Homophily, Collaboration, and Output," *Journal of Economic Behavior & Organization*, 2024, *221*, 477–492.

_ , Sanjeev Goyal, and Anja Prummer, "Gender and Collaboration," *Review of Economics and Statistics*, 2023, *105* (6), 1366–1378.

Fafchamps, Marcel, Marco J. van der Leij, and Sanjeev Goyal, "Matching and Network Effects," *Journal of the European Economic Association*, 2010, *8* (1), 203–231.

Ferber, Marianne A. and Michelle Teiman, "Two to Tango? The Importance of Networks for the Careers of Academic Economists," *Economic Inquiry*, 2006, *44* (3), 413–430.

Freeman, Richard B. and Wei Huang, "Collaborating with People Like Me: Ethnic Coauthorship within the United States," *Journal of Labor Economics*, 2015, *33* (S1), S289–S318.

Gertsberg, Marina, "The Unintended Consequences of #MeToo: Evidence from Research Collaborations in Economics and Finance," 2025. Forthcoming Journal of Finance.

Goyal, Sanjeev, Marco J. van der Leij, and José L. Moraga-González, "Economics: An Emerging Small World," *Journal of Political Economy*, 2006, *114* (2), 403–412.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Sarsons, Heather, Klarita Gërxhani, Ernesto Reuben, and Arthur Schram, "Gender Differences in Recognition for Group Work," *Journal of Political Economy*, 2021, *129* (1), 101–147.

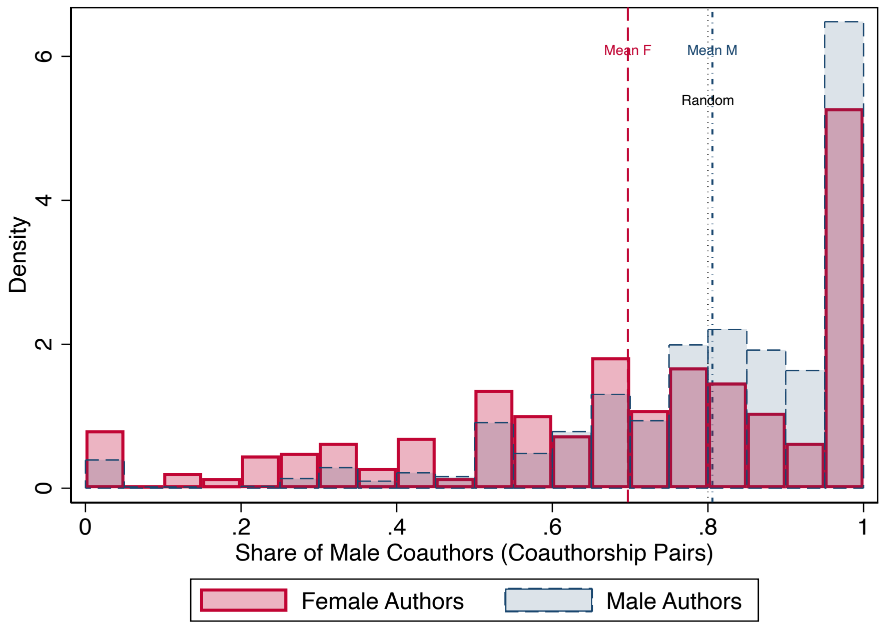

Figure 1: Distribution of Share of Male Coauthorship Pairs by Gender

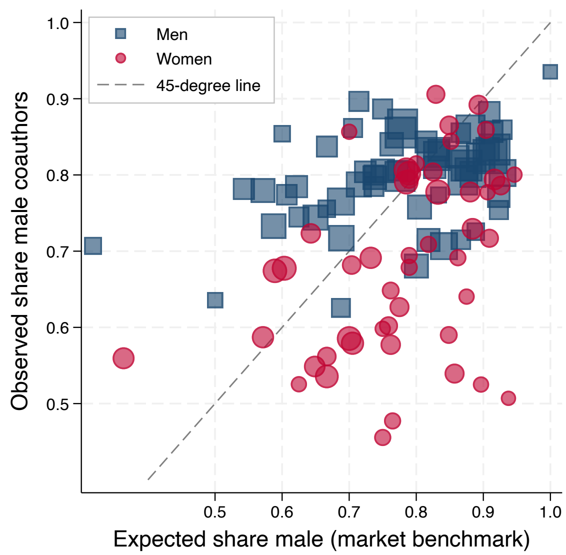

Figure 2: Observed vs. Expected Share of Male Coauthors by Market

Notes: Each point is a market cell defined by field × PhD institution rank (top 10 vs. rest) × cohort (pre- vs. post-2003), with a minimum of 3 authors of each gender per cell. Squares show the average observed share of male coauthors for male economists; circles show the same for female economists. Point size is proportional to cell size. The dashed line is the 45-degree line where observed equals expected. Under random matching within markets, both series would lie on this line.

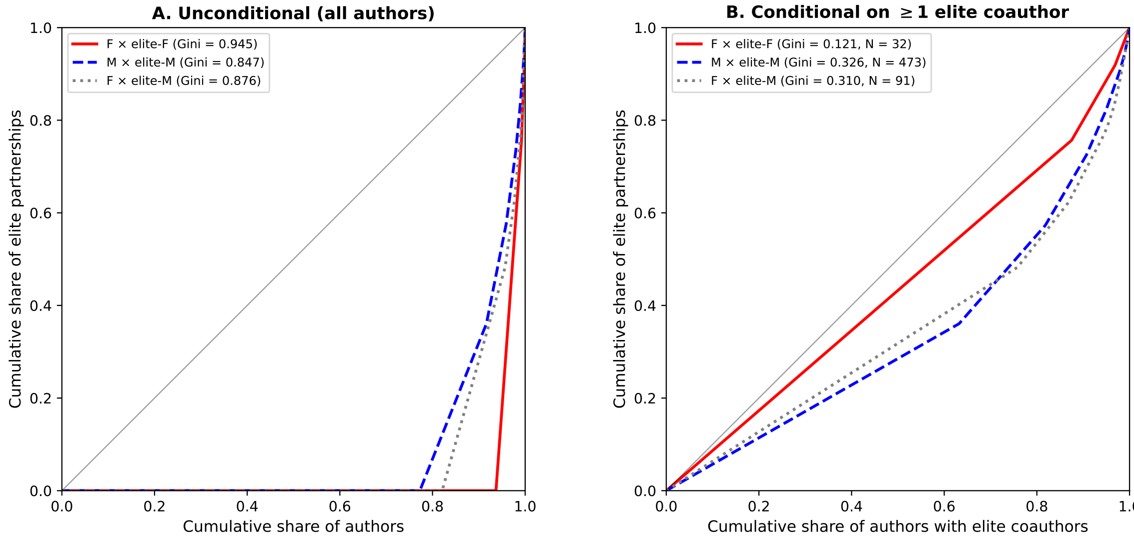

Figure 3: Lorenz Curves of Elite Same-Gender Coauthorship

Notes: Panel A plots Lorenz curves for all authors (including those with zero elite same-gender coauthors). Panel B restricts to authors with ≥ 1 elite coauthor in the relevant group. Elite defined as faculty in top-5 departments. F × elite-F = female authors' partnerships with elite female coauthors; M × elite-M = male authors with elite male coauthors; F × elite-M = female authors with elite male coauthors. Gini coefficients reported in legends.

Table 1: Descriptive Statistics, by Gender

	Female (1)	Male (2)	Diff (1)−(2)
Education			
PhD year	2005.7	1999.3	6.4***
Current Status			
Department rank 1–20	0.220	0.297	-0.077***
Department rank 21–45	0.271	0.268	0.003
Department rank 46–80	0.273	0.233	0.039*
Department rank 81+	0.236	0.201	0.035*
Assistant Professor	0.348	0.221	0.127***
Associate Professor	0.264	0.203	0.061***
Full Professor	0.386	0.573	-0.187***
Field			
Macro/International/Finance	0.253	0.313	-0.060***
Public/Health/Education	0.316	0.225	0.091***
Environmental/Urban	0.145	0.108	0.038**
Development/Political/History	0.217	0.166	0.051***
Industrial Organization	0.082	0.097	-0.015
Labor	0.257	0.165	0.092***
Theory	0.161	0.250	-0.089***
Econometrics	0.105	0.127	-0.022
Productivity (2010–2023)			
Total number of papers	4.78	6.03	-1.26***
Number of top-5 papers	0.69	1.11	-0.42***
Number of unique coauthors	8.71	11.07	-2.35***
Share male coauthorship pairs	0.70	0.81	-0.11***
N	572	2240	2812

Authors can be classified in multiple fields.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (two-sided t-tests).

Table 2: Share of Male Coauthors

	(1)	(2)	(3)
Female	-0.109***	-0.090***	-0.088***
	(0.012)	(0.012)	(0.012)
N	2812	2812	2812
Dep. var. mean	0.784	0.784	0.784
PhD year FE	No	Yes	Yes
PhD institution rank FE	No	Yes	Yes
Field FE	No	Yes	Yes
Current department rank FE	No	No	Yes
Academic rank FE	No	No	Yes

Robust standard errors in parentheses.

Outcome: share of coauthorship pairs who are male.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Share of Male Coauthors by Journal Type

	Cross-sectional		Author Fixed Effects	
	Top-5	Non-Top-5		
	(1)	(2)	(3)	(4)
Female	-0.155***	-0.101***		
	(0.026)	(0.018)		
Top-5 paper			0.026**	0.034***
			(0.010)	(0.010)
Female × Top-5				-0.047
				(0.032)
N	1217	1217	2434	2434
Dep. var. mean	0.808	0.782	0.795	0.795
PhD year FE	Yes	Yes	–	–
PhD institution rank FE	Yes	Yes	–	–
Field FE	Yes	Yes	–	–
Author FE	No	No	Yes	Yes

Robust standard errors in parentheses (clustered by author in cols 3-4).

Outcome: share of coauthorship pairs who are male.

Sample: authors with both Top-5 and Non-Top-5 publications (N = 1217).

Cols 1-2: Difference in Female coeff = -0.054 (p = 0.084).

Col 4: Top-5 effect for women = -0.012 (p = 0.684).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Access to Elite Coauthors

Panel A: All Elite Coauthors

	Any Top5	Any Top10	Any TopPr	N Top5	N Top10	N TopPr
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.045***	-0.043**	-0.062***	-0.119***	-0.178***	-0.227***
	(0.017)	(0.021)	(0.021)	(0.046)	(0.063)	(0.064)
N	2812	2812	2812	2812	2812	2812
Dep. var. mean	0.221	0.336	0.354	0.406	0.734	0.830

Panel B: Male Elite Coauthors

	Any Top5	Any Top10	Any TopPr	N Top5	N Top10	N TopPr
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.057***	-0.060***	-0.070***	-0.127***	-0.203***	-0.254***
	(0.017)	(0.020)	(0.020)	(0.040)	(0.053)	(0.056)
N	2812	2812	2812	2812	2812	2812
Dep. var. mean	0.206	0.310	0.331	0.350	0.618	0.716

Panel C: Female Elite Coauthors

	Any Top5	Any Top10	Any TopPr	N Top5	N Top10	N TopPr
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.007	0.018	0.007	0.007	0.024	0.027
	(0.011)	(0.015)	(0.014)	(0.013)	(0.022)	(0.021)
N	2812	2812	2812	2812	2812	2812
Dep. var. mean	0.050	0.094	0.093	0.056	0.116	0.115

Robust standard errors in parentheses.

All columns control for PhD year FE, PhD institution rank FE, and field FE.

Cols 1–3: dummy for any elite coauthor (full sample).

Cols 4–6: count of unique elite coauthors (unconditional, full sample).

Elite = coauthor at top-5/top-10 current institution or top productivity decile.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Concentration of Elite Same-Gender Collaborations

Panel A: Female Authors × Elite Female Coauthors

	N Elite Coauthors	Unique Coauthorship Pairs	HHI	Equiv. N (1/HHI)	Top 10 Share (%)
Top-5 department	20	37	0.0811	12.3	73.0
Top-10 department	40	78	0.0398	25.1	55.1
Top productivity decile	37	75	0.0432	23.1	58.7

Panel B: Male Authors × Elite Male Coauthors

	N Elite Coauthors	Unique Coauthorship Pairs	HHI	Equiv. N (1/HHI)	Top 10 Share (%)
Top-5 department	162	829	0.0097	103.4	18.6
Top-10 department	332	1,461	0.0047	212.5	10.5
Top productivity decile	310	1,720	0.0044	224.8	9.7

Panel C: Female Authors × Elite Male Coauthors

	N Elite Coauthors	Unique Coauthorship Pairs	HHI	Equiv. N (1/HHI)	Top 10 Share (%)
Top-5 department	80	143	0.0170	58.9	27.3
Top-10 department	150	247	0.0091	109.9	17.4
Top productivity decile	158	262	0.0087	114.4	17.2

Unit of analysis: unique author–coauthor pairs.

HHI = $\sum_j s_j^2$ where s_j is elite coauthor j's share of all unique pairs.

Equiv. N = 1/HHI, the number of equally-sized elite coauthors that would produce the same concentration.

Top 10 Share = fraction of unique pairs accounted for by the 10 most-connected elite coauthors.

Table 6: Fragility of Elite Coauthor Access: Hub Removal Counterfactual

	−1	−3	−5	−10
Panel A: Remove top elite-F coauthors				
P(elite-F | female)	−16%	−41%	−53%	−69%
P(any elite | female)	−4%	−5%	−6%	−9%
P(any elite | male)	0%	0%	0%	−1%
Panel B: Remove top elite-M coauthors				
P(elite-M | male)	−1%	−4%	−6%	−11%
P(elite-M | female)	0%	−1%	−1%	−9%
P(any elite | female)	0%	−1%	−1%	−8%
P(any elite | male)	−1%	−3%	−5%	−9%

Elite defined as coauthors at top-5 departments.
Columns show % change relative to baseline after removing top-*k* elite coauthors ranked by number of unique partnerships.
Baseline: P(elite-F|F) = 6.3%, P(elite-M|M) = 22.7%, P(elite-M|F) = 17.9%, P(any elite|F) = 20.7%, P(any elite|M) = 24.0%.
Results for top-10 dept and top productivity in Appendix Table A.10.

Appendix

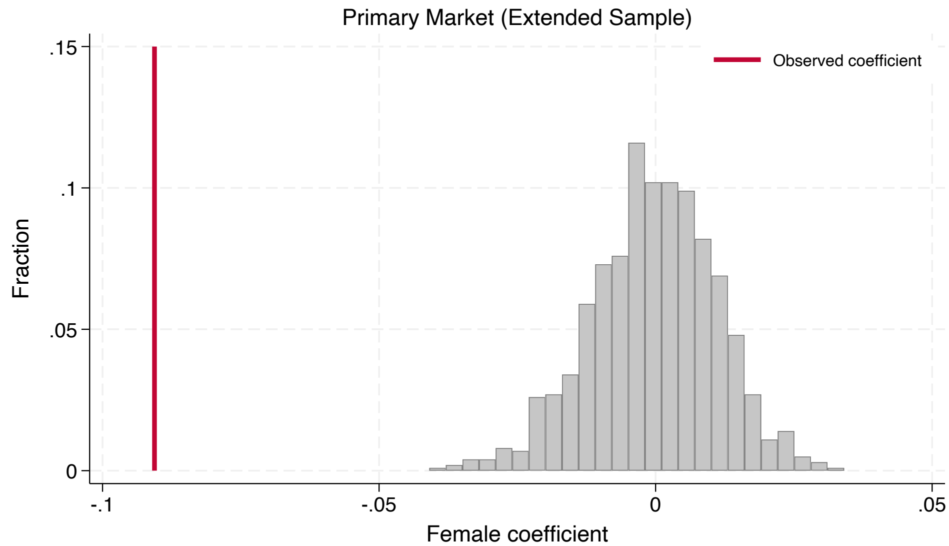

Figure A.1: Permutation Test: Within-Market Female Coefficient

Notes: Histogram shows the distribution of the Female coefficient from 1,000 within-market permutations. In each permutation, the Female indicator is reshuffled within market cells (field × PhD institution rank (top 10 vs. rest) × cohort (pre- vs. post-2003)) and a regression of Share Male on Female with market fixed effects, PhD year-cohort, and PhD rank controls is estimated (Table A.4 Column 2). The outcome (Share Male) remains fixed; only the Female indicator is permuted. The vertical line marks the observed coefficient. Under the null hypothesis of no within-market gender sorting, the distribution is centered at zero.

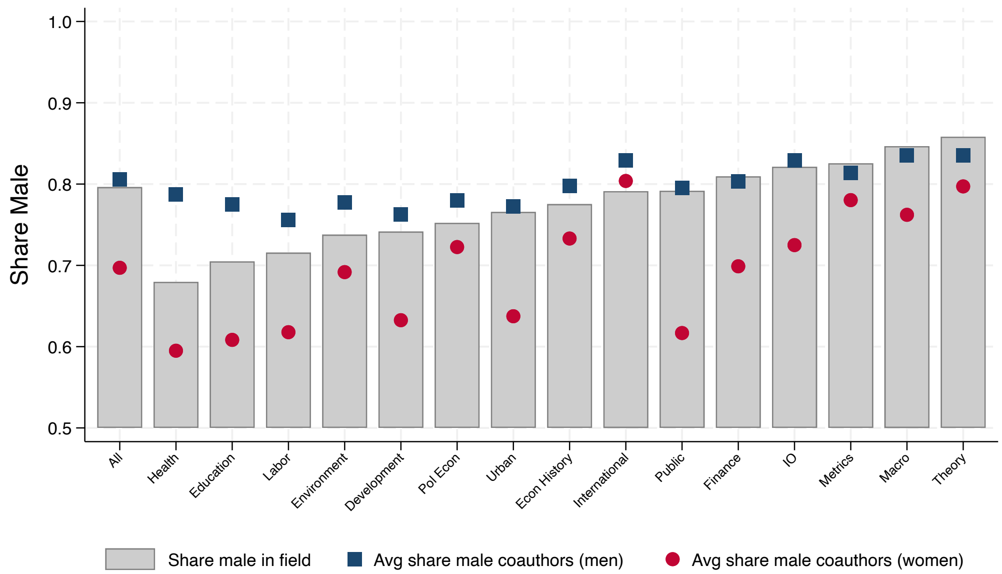

Figure A.2: Male Share of Scholars and Coauthors by Field

Notes: Bars show the male share of scholars in each field (the random-matching benchmark). Squares show the average share of male coauthors for male economists; circles show the average share of male coauthors for female economists. Fields are ordered by female share of scholars (highest to lowest). Under random matching, both markers would equal the bar height.

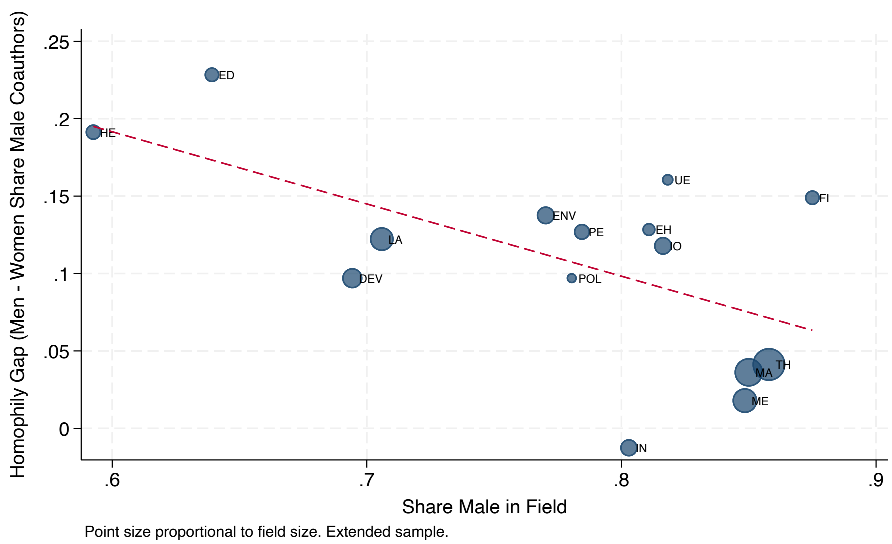

Figure A.3: Field-Level Homophily Gap and Share Male

Notes: Each point represents a field. The y-axis shows the homophily gap (men's average share male coauthors minus women's average share male coauthors). The x-axis shows the share of male scholars in the field. Point size is proportional to field size. The dashed line shows the linear fit.

Table A.1: Productivity by Gender

	N Papers	N Top 5	Share Top 5
	(1)	(2)	(3)
	(1)	(2)	(3)
Female	-0.472**	-0.307***	-0.046***
	(0.214)	(0.075)	(0.012)
N	2812	2812	2812
Dep. var. mean	5.776	1.025	0.163

Robust standard errors in parentheses.
All columns control for PhD year FE, PhD institution rank FE, and field FE.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Share of Male Coauthors by Tie Type

	Unique coauthors	All ties	New ties	Repeat ties
	(1)	(2)	(3)	(4)
Female	-0.083***	-0.090***	-0.083***	-0.111***
	(0.012)	(0.012)	(0.012)	(0.021)
N	2589	2812	2589	1852
Dep. var. mean	0.771	0.784	0.771	0.798

Robust standard errors in parentheses.
All columns control for PhD year FE, PhD institution rank FE, and field FE.
Col 1: each coauthor counted once. Cols 3–4: new = first paper with coauthor; repeat = subsequent papers.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Oster (2019) Bounds for Selection on Unobservables

	(1) Share Male	(2) Elite Access	(3) Top-5 Pubs
Panel A: Regression Results			
Outcome	Share male	Elite access	N top-5 pubs
Treatment	Female	Female	Elite access
β (short, no controls)	-0.109	-0.038	2.386
R^2 (short)	0.036	0.001	0.219
β (controlled)	-0.090	-0.045	1.929
SE	(0.012)	(0.017)	(0.127)
R^2 (controlled)	0.093	0.205	0.316
N	2,812	2,812	2,812
Panel B: Oster Bounds			
R^2_{max} (= 1.3 $\times R^2$)	0.121	0.267	0.411
δ	2.33	-1.92	4.11
β^* (assuming δ = 1)	-0.081	-0.047	1.483

Notes: Column (1): outcome is share of male coauthors, treatment is female. Column (2): outcome is elite access (any coauthor at top-5 dept), treatment is female. Column (3): outcome is number of top-5 publications, treatment is elite access (any coauthor at top-5 department), controls include female. All controlled regressions include PhD year FE, PhD institution rank FE, and field FE. δ measures how strong selection on unobservables would need to be (relative to observables) to fully explain the effect. $|\delta| > 1$ suggests robustness to omitted variable bias. A negative δ (Column 2) indicates that adding controls moves the coefficient *away* from zero, so unobservables would need to operate in the opposite direction of observables to explain the effect (Oster, 2019).

Table A.4: Market-Adjusted Share of Male Coauthors

	(1)	(2)	(3)	(4)
Female	-0.078***	-0.081***	-0.079***	-0.081***
	(0.012)	(0.012)	(0.012)	(0.014)
N	2812	2812	2812	2781
Dep. var. mean	-0.013	-0.013	-0.013	-0.014
Market def.	Field	Field×Rank(2)×Cohort(2)	Field×Rank(2)×Cohort(2)	Field×Cohort(4)×Rank(5)
PhD year FE	Yes	Yes	Yes	Yes
PhD inst. rank FE	Yes	Yes	Yes	Absorbed
Dept. rank & acad. rank	No	No	Yes	No

Robust standard errors in parentheses.

Outcome is Excess Share Male = observed − LOO expected share within market.

LOO = leave-one-out expected share male within market cells.

"Absorbed" indicates the variable is mechanically controlled by the market definition.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Share of Male Coauthors by Cohort and Department Rank

Panel A: By PhD Cohort				
	1955-1989	1990-2002	2003-2012	2013-2022
	(1)	(2)	(3)	(4)
Female	-0.020	-0.100***	-0.096***	-0.101***
	(0.035)	(0.023)	(0.020)	(0.025)
N	681	695	749	687
Dep. var. mean	0.802	0.766	0.785	0.782

Panel B: By Current Department Rank				
	Rank 1-20	Rank 21-45	Rank 46-80	Rank 81-126
	(1)	(2)	(3)	(4)
Female	-0.095***	-0.125***	-0.076***	-0.056*
	(0.022)	(0.025)	(0.024)	(0.029)
N	792	756	679	585
Dep. var. mean	0.806	0.785	0.774	0.762

Robust standard errors in parentheses.
All columns control for PhD year FE, PhD institution rank FE, and field FE.
Department rank based on RePEc ranking of current institution.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Share of Male Coauthors by Field

	Macro/ Int/Fin (1)	Health/ Ed/Pub (2)	Env/ Urban (3)	Dev/Pol/ Hist (4)	IO (5)	Labor (6)	Theory (7)	Metrics (8)
Female	-0.062***	-0.163***	-0.082**	-0.097***	-0.102**	-0.143***	-0.035	-0.038
	(0.023)	(0.025)	(0.038)	(0.027)	(0.046)	(0.028)	(0.025)	(0.037)
N	847	686	324	496	264	517	651	344
Dep. var. mean	0.814	0.744	0.753	0.748	0.810	0.717	0.830	0.808

Robust standard errors in parentheses.

All columns control for PhD year FE and PhD institution rank FE.

Authors can be classified in multiple fields.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Excess Elite Access: LOO Market Mean vs. Random-Matching Benchmark

Panel A: All Elite Coauthors

	Top 5		Top 10		Top Prod	
	LOO	Random	LOO	Random	LOO	Random
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.053***	-0.032*	-0.047**	-0.019	-0.072***	-0.033
	(0.018)	(0.017)	(0.021)	(0.020)	(0.021)	(0.020)
N	2812	2812	2812	2812	2812	2812
Dep. var. mean	-0.000	-0.050	0.000	-0.070	-0.000	-0.097

Panel B: Male Elite Coauthors

	Top 5		Top 10		Top Prod	
	LOO	Random	LOO	Random	LOO	Random
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.063***	-0.044***	-0.063***	-0.034*	-0.081***	-0.042**
	(0.017)	(0.016)	(0.020)	(0.019)	(0.020)	(0.020)
N	2812	2812	2812	2812	2812	2812
Dep. var. mean	0.000	-0.040	0.000	-0.057	-0.000	-0.082

Panel C: Female Elite Coauthors

	Top 5		Top 10		Top Prod	
	LOO	Random	LOO	Random	LOO	Random
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.002	0.007	0.014	0.023*	0.006	0.015
	(0.011)	(0.011)	(0.015)	(0.014)	(0.014)	(0.014)
N	2812	2812	2812	2812	2812	2812
Dep. var. mean	0.000	-0.002	0.000	-0.007	0.000	-0.010

Robust standard errors in parentheses. All columns control for PhD year FE and PhD institution rank FE. Market: field × PhD rank (top 10 vs rest) × cohort (pre/post 2003). Field FE absorbed by market. LOO: Excess = AnyElite$_i$ − leave-one-out market mean of AnyElite. Random: Excess = AnyElite$_i$ − $[1 - (1 - q_m)^{k_i}]$, where q_m = LOO fraction of unique coauthors who are elite (computed from pair-level data), k_i = number of unique coauthors. Because k_i may itself be affected by homophily, this benchmark likely overcontrols. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Comparative Statics: Elite Composition and Access to Elite Coauthors

	Top-5 Dept	Top-10 Dept	Top Prod. Decile
	(1)	(2)	(3)
Female	0.049	0.093	0.128**
	(0.065)	(0.067)	(0.051)
Female × Elite Male Share	-0.126*	-0.173**	-0.244***
	(0.067)	(0.080)	(0.064)
N	2231	2395	2385
Mean dep. var.	0.218	0.336	0.356

All columns: Market FE, PhD year FE, PhD rank FE.

Market = field × PhD rank (top 10 vs. rest) × cohort (pre- vs. post-2003).

Elite Male Share = among elite authors in the market, fraction who are male.

Elite definition matches the column. Market-level main effect absorbed by FE.

SEs clustered at market level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: Access to Elite Coauthors – Robustness

Panel A: Any Coauthor in Top-5 Department

	Baseline	+ Controls	Assistants	Non-elite
	(1)	(2)	(3)	(4)
Female	-0.045***	-0.036**	-0.068**	-0.035**
	(0.017)	(0.017)	(0.029)	(0.016)
N	2812	2812	693	2365
Dep. var. mean	0.221	0.221	0.196	0.146
Sample	Full	Full	Assistants	Non-elite

Panel B: Any Coauthor in Top Productivity Decile

	Baseline	+ Controls	Assistants	Non-elite
	(1)	(2)	(3)	(4)
Female	-0.062***	-0.052**	-0.049	-0.052**
	(0.021)	(0.021)	(0.038)	(0.022)
N	2812	2812	693	2365
Dep. var. mean	0.354	0.354	0.329	0.276
Sample	Full	Full	Assistants	Non-elite

Robust standard errors in parentheses.

All columns control for PhD year FE, PhD institution rank FE, and field FE.

Col 2: adds controls for number of unique coauthors and number of papers.

Col 3: sample restricted to assistant professors.

Col 4: sample excludes authors in top-10 departments.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Fragility of Elite Coauthor Access: Hub Removal Counterfactual (All Definitions)

	Top-5 Dept				Top-10 Dept				Top Prod. Decile			
	−1	−3	−5	−10	−1	−3	−5	−10	−1	−3	−5	−10
Panel A: Remove top elite-F coauthors												
P(elite-F | female)	−16%	−41%	−53%	−69%	−8%	−24%	−34%	−47%	−5%	−22%	−27%	−47%
P(any elite | female)	−4%	−5%	−6%	−9%	−1%	−5%	−5%	−6%	−1%	−4%	−4%	−6%
P(any elite | male)	0%	0%	0%	−1%	0%	0%	0%	−1%	0%	0%	−1%	−1%
Panel B: Remove top elite-M coauthors												
P(elite-M | male)	−1%	−4%	−6%	−11%	0%	−2%	−3%	−5%	0%	−1%	−2%	−5%
P(elite-M | female)	0%	−1%	−1%	−9%	0%	−1%	−1%	−2%	0%	−1%	−2%	−3%
P(any elite | female)	0%	−1%	−1%	−8%	0%	−1%	−1%	−2%	0%	−1%	−1%	−2%
P(any elite | male)	−1%	−3%	−5%	−9%	0%	−2%	−2%	−4%	0%	0%	−2%	−4%

Columns show % change relative to baseline after removing top-k elite coauthors ranked by number of unique partnerships.
Panel A removes elite-F coauthors; Panel B removes elite-M coauthors. Elite definition matches the column group.
Baselines — Top-5: P(elite-F|F) = 6.3%, P(elite-M|M) = 22.7%, P(elite-M|F) = 17.9%, P(any|F) = 20.7%, P(any|M) = 24.0%.
Top-10: P(elite-F|F) = 11.6%, P(elite-M|M) = 33.3%, P(elite-M|F) = 28.0%, P(any|F) = 31.9%, P(any|M) = 35.7%.
Top Prod.: P(elite-F|F) = 10.8%, P(elite-M|M) = 36.8%, P(elite-M|F) = 30.3%, P(any|F) = 33.7%, P(any|M) = 39.1%.

A Appendix: Data Construction

A.1 Sample of Departments and Faculty

Our sample is based on the top 100 U.S. economics departments as ranked by RePEc (Research Papers in Economics) in May 2023. We excluded business schools when a university had both an economics department and a business school, retaining only faculty listed in economics departments.

We collected names of all tenured and tenure-track faculty by scraping department websites in 2023, supplemented by the 2020–2021 Economics Faculty Directory (Hasselback). We excluded lecturers, visitors, emeritus faculty, and courtesy appointments. For each faculty member, we recorded name, affiliation, academic rank, PhD institution, PhD year, and a photograph.

A.2 Publication Data

Faculty publications were extracted from the Web of Science (WOS) in December 2023–January 2024. We searched for publications in two sets of journals: (i) the top-5 economics journals (AER, Econometrica, JPE, QJE, REStud) and (ii) the top-125 economics journals as classified by WOS. Searches were limited to journal articles from 2010–2023; we excluded earlier years because WOS coverage of economics journals was expanding through the 2000s, making pre-2010 data less comparable.

Names were searched in batches by university to facilitate disambiguation of common names. We transformed the data from one record per publication to one record per author-paper pair. For authors with multiple affiliations, we retained the highest-ranked university. We excluded papers with more than 10 coauthors, which are typically review articles or large data collection efforts.

A.3 Name Cleaning

Substantial effort was devoted to correcting name variations due to spelling errors, nicknames, middle name variations, and hyphenation differences. Over 200 hours were spent cleaning author names from top-5 journals. As discussed in the main text, significant effort was made to try to identify and rectify name changes which were detected primarily by examining faculty with zero reported publications, looking at names in lists of PhD recipients at

AER/JEL and via AI queries.

A.4 Gender Assignment

For faculty in our core sample, gender was assigned using profile photographs. For coauthors outside the core sample, we used the NAMSOR algorithm, which assigns gender probabilities based on a database of over one billion names worldwide. Names were classified as female when the predicted probability exceeded 0.5, and analogously for male.

A.5 Field Assignment

Research fields were assigned using a hierarchical approach combining multiple sources. Our primary classification uses NBER working paper program codes and publication in field-specific journals. NBER programs were aggregated into broader categories (e.g., Labor Studies → Labor; Corporate Finance and Asset Pricing → Finance; Economic Fluctuations and Monetary Economics → Macro). Field journals include outlets such as the *Journal of Labor Economics*, *Journal of Health Economics*, and *Journal of Monetary Economics*. An author is assigned to a field if at least 35% of their papers fall in that field or if they have more than 9 papers in that field. Authors may be assigned to multiple fields.

For authors without NBER or field-journal information, we used ChatGPT to assign JEL codes to paper titles and abstracts, mapping JEL codes to fields (e.g., J → Labor; I → Health/Education; E → Macro). For authors with no other field information, we used dissertation field when available.

B Additional Results and Proofs

B.1 Mapping the Model to the Extensive Margin

In the model, each economist of gender g draws k collaborators independently, and each collaborator is elite with probability $p_g \equiv \Pr(E \mid g)$. As a result, the number of elite collaborators $\#E$ follows a binomial distribution with parameters (k, p_g).

Lemma 1. *In the model, the probability that an economist of gender g has at least one elite collaborator is*

$$\Pr(\#E \geq 1 \mid g) = 1 - (1 - p_g)^k,$$

which is strictly increasing in p_g *and in the expected number of elite collaborators* $\mathbb{E}[\#E \mid g] = kp_g$. *Consequently, any comparative static that increases* $\Pr(E \mid g)$ *or* $\mathbb{E}[\#E \mid g]$ *also increases the probability of observing at least one elite collaborator.*

Proof. Independence of collaborator draws implies $\Pr(\#E = 0 \mid g) = (1 - p_g)^k$. The result follows immediately. Differentiating yields

$$\frac{\partial}{\partial p_g}\left[1 - (1 - p_g)^k\right] = k(1 - p_g)^{k-1} > 0$$

for $p_g \in [0, 1)$ and $k \geq 1$. □