Journal of Economic Behavior and Organization 221 (2024) 477–492

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Research Paper

Gender homophily, collaboration, and output ☆

Lorenzo Ductor [a,*], Anja Prummer [b]

[a] University of Granada, Spain
[b] Johannes Kepler University Linz, Austria

ARTICLE INFO

JEL classification:
D85
J16
O30

Keywords:
Homophily
Collaboration
Diversity
Research quality

ABSTRACT

We consider the implications of gender homophily in Economics, which has persisted despite the significant increase in women in the field. With women still underrepresented, gender homophily could hinder collaboration. Additionally, it may result in less gender-diverse co-author teams, potentially diminishing research quality. Our findings reveal that gender homophily does not limit collaboration or reduce output quality. Merely increasing the number of women in Economics may not suffice to address entrenched gender disparities.

1. Introduction

Many workplaces strive to attract women. More female colleagues increase the share of women in teams. The influx of women either leads to greater gender diversity or more collaboration between women. The level of gender diversity observed may depend on which teams produce the best results. If collaborators are chosen freely, then preferences or differential interaction rates between men and women could also affect the gender composition of a team, potentially inducing too little diversity.

In this paper, we describe how female and male economists collaborate across gender from 1970 to 2017 and connect gender diversity to research output.

Authors may display gender homophily, meaning that they collaborate relatively more with authors of the same gender. In contrast, if authors work disproportionately with Economists of the opposite gender, they exhibit heterophily. As in Currarini et al. (2009), we interpret homophily as an equilibrium phenomenon that does not only depend on preferences but also interaction rates between men and women.[1] The simplest measure of homophily, *Relative Homophily*, compares the share of same-gender co-authors with the share of same-gender authors among all economists. Both men and women display relative homophily. However, economics has become significantly more gender-balanced. At the beginning of our sample period, in 1970, only 5% of authors were female.

☆ We thank our editor, Friederike Mengel, as well as two anonymous referees for excellent comments. We are indebted to Sanjeev Goyal for helpful comments and his contribution to the homophily section. We are grateful to René Böheim, Matthias Fahn, Erin Hengel, Nagore Iriberri, Michael Rose and Rudolf Winter-Ebmer for their valuable comments. Thi Phuong Anh Doan provided excellent research assistance. We thank various audiences for comments and suggestions. Lorenzo Ductor acknowledges the financial support of the *Fundación Ramón Areces*, project XVIII National Competition for the Allocation of Research Grants in Social Sciences, and the Ministry of Science and Innovation of Spain, Grants PID2022-142943NB-I00 and PID2019-111708GA-100, funded by MICIU/AEI/ 10.13039/501100011033 and, ERDF/EU. This research was funded in whole or in part by the Austrian Science Fund (FWF) V 991-G. Funding for open access charge: Universidad de Granada / CBUA.

* Corresponding author.
E-mail addresses: lductor@ugr.es (L. Ductor), anja.prummer@jku.at (A. Prummer).

[1] For a detailed discussion of homophily, see McPherson et al. (2001).

https://doi.org/10.1016/j.jebo.2024.03.027
Received 8 March 2023; Received in revised form 22 March 2024; Accepted 25 March 2024

Available online 10 April 2024

This share has increased to almost 30% in 2017 (Ductor et al. (2021)). Therefore, both female and male authors have now more opportunities to co-author with women. This increase in the share of female authors requires us to use a measure of homophily that takes into account the varying gender composition in our population, *Inbreeding Homophily*. This measure normalizes the same gender co-authors by their share in the population. Both male and female economists display inbreeding homophily, collaborating disproportionately with same-gender co-authors. Men display a higher level of homophily compared to women. Remarkably, the level of homophily has remained stable throughout our observation period. The influx of women is neither correlated with a relative increase in female or male co-authors. Rather, the proportions remained unchanged for both men and women in economics.

The stability and persistence of homophily leads us to investigate whether the predominantly same-gender collaborations are correlated with research output. We consider two channels, namely (i) the smaller number of same type co-authors for women and (ii) the association between team diversity and research quality.

In the presence of homophily, women, as the underrepresented group in Economics, face a smaller pool of potential collaborators. The number of collaborators is a key factor in predicting research output: a higher number of co-authors is tied to more high-quality publications in the future (Ductor et al. (2021)).[2] If homophily was indeed a factor limiting the number of collaborators, then an increase in the share of women would ameliorate the shortage of co-authors, an idea formalized by Currarini et al. (2009). We test this prediction using the variation in the share of women over time as well as across fields. Different fields in Economics attract distinct shares of women. Perhaps surprisingly, women's average number of collaborators does not correlate with the share of female authors present. This indicates that women's number of collaborators has not increased over time relative to men's and the gender gap in number of collaborators is unaffected by research field. Even though economists display homophily, it does not impact the gap in the number of collaborators. This implies that the mere increase in the share of women is not correlated with women acquiring more co-authors relative to men. Our findings indicate that a higher number of women in Economics may not per se decrease the gender gap in number of collaborators, which in turn is an important contributor to the continued gender gap in research.[3]

Homophily has implications for the gender diversity of the research team. While homophily and team diversity are related, they differ in how they are constructed. Homophily is an *author-specific* measure, while team diversity is assessed for each research team that produces an article, making it *article-specific*. Male teams write 54% of all articles, while all-female teams only make up 2%. Mixed teams account for 17% of publications.[4] Gender diverse teams make up for a lower share of publications relative to same-gender teams, consistent with homophily.

Research teams exclusively comprised of men publish articles in higher-impact journals compared to mixed teams, which perform as well as all-female teams for our sample of 100 journals in Economics.[5] An all-female team is correlated with articles published in journals with an 8% lower impact factor compared to an all-male team.

In contrast, female teams accrue the highest citations compared to mixed and male teams, controlling for journal impact factor and other observables, in line with Card et al. (2020). All female teams are connected to 26% higher citations compared to male teams, while mixed teams have 5% higher citations.

Consequently, there is no clear benefit to collaborate with an author of a different gender. For male authors working with a woman, the drop in the journal quality is at least associated with a slightly higher citation count, conditional on observables. However, women do not attain any benefit from collaborating with men. If a female author opts to work with a male collaborator, then the publication quality does not improve relative to a paper written with a female author, and such cross-gender collaboration tends to acquire lower citations.

While both men and women display gender homophily, homophily does not seem to be limiting them in finding collaborators and publishing as gender diversity does not improve research quality. While an increase in the share of women in Economics is an important goal in and of itself, it does not seem sufficient to erase gender disparities.

Finally, our results raise questions about the correct metric for evaluating research contributions given the notable gender differences between measures. The article influence score, a sophisticated measure of journal quality, seems in practice particularly relevant as publications play a crucial role in advancing academic careers. The journal an article is published in matters for recruitment, tenure, promotions and recognition, especially at early stages when citations have not accumulated yet (Laband (2013)). Given that male teams are more likely to publish in higher-ranked journals, simply counting these publications means a gender bias in evaluating candidates. However, such a bias would persist if raw citations were added to the evaluation criteria–as male teams also outperform female and mixed teams on this metric. A more gender-neutral approach would require to account for citations conditional on the publication quality.

Related literature We contribute to the literature on homophily in social networks. Homophily has been extensively studied in sociology (for a summary see McPherson et al. (2001)). In Economics, gender homophily has been shown to arise in job search networks (Torres and Huffman (2002), Zhu (2018)), referral networks (Beaman et al. (2013), Zeltzer (2020)) as well as in the lab (Mengel (2020)). Jackson et al. (2023) document gender homophily among college students, distinguishing between friendships

[2] Note that the number of co-authors is based on previous and current collaboration, while research output is assessed in future periods to rule out simultaneity.

[3] Ductor et al. (2021) show that the gender differences in past number of collaborators explain approximately 10% of the variation in research output differences across gender, over and above past performance. Doan et al. (2023) shows that the relationship between number of collaborators and research output is causal. They found that loosing a co-author leads to a decline in long-run output of 11%.

[4] For the remainder, the gender composition cannot be identified. If we cannot identify an author's gender, then we can also not assess the gender composition of the research teams.

[5] We use the Top 100 journals and follow in our classification of journal quality Ductor et al. (2020).

and study groups. We complement their findings by documenting gender homophily in team formation. Concurrently to this paper, Davies (2022) shows gender homophily in NBER working papers, while we cover all journals in Economics over a 47 year period.

More broadly, we add to the literature documenting the constraints women in economics face, see e.g., Ginther and Kahn (2004), Sarsons et al. (2021), Wu (2018), Hengel (2016), Chari and Goldsmith-Pinkham (2017), Boring (2017), Mengel et al. (2019), Card et al. (2020), Paredes et al. (2020), Dupas et al. (2021), Alexander et al. (2021). Ductor et al. (2018) shows that differing network structures between men and women can account for a substantial part of the gender gap in research output. In particular, a higher number of co-authors is a key predictor for future research output. We document gender homophily and investigate whether it serves as a constraint on finding collaborators for women in Economics.

Currarini et al. (2009), Bramoullé et al. (2012) model matching processes that tie homophily to the number of connections, formalizing the idea that homophily can serve as a constraint for underrepresented groups. Our paper highlights that the influx of women in Economics has not reduced the gender differences in the number of co-authors. This contrasts with Currarini et al. (2009)' finding on interracial friendships, emphasizing the importance of distinguishing between (i) friendship ties and work collaborations and (ii) race and gender.

Our finding on gender team diversity and research output confirms Card et al. (2020), who show that publications of women attract a higher number of citations in four journals, when compared to publications by men with similar observable characteristics. This citations gap is robust in that it persists for the Top 100 journals in Economics. The gap also emerges if we exclusively consider publications in Top 5 journals as in Hengel and Moon (2020).[6]

In light of Koffi (2021), our result can be interpreted as a lower bound on the citations gap. She documents that papers with a more female co-author team are cited less relative to the citations they should be receiving, according to the quality of the article (measured using the degree of novelty of the article). To assess the quality and novelty of an article, Koffi (2021) relies on textual analysis techniques and machine learning tools. While we find that women are cited more, the citation bias documented by Koffi (2021) could indicate that women should be cited even more.[7]

The gender composition of teams has been analyzed in the context of boardrooms, in committees and in experiments, see the surveys by Azmat and Petrongolo (2014), Azmat and Boring (2020). Results as to whether gender diversity is beneficial for better outcomes are mixed, in line with the theory. Theoretical work has explored how diversity in groups worsens communication across members (Lazear 1999, Hong and Page 2001, Prat 2002, Kets and Sandroni 2021, Dong and Mayskaya 2023, Dong et al. 2023, Hughes et al. 2023), but how important this effect depends on the specific environment.[8] This warrants an investigation into economics, especially as Sarsons et al. (2021) have shown that women do not receive credit for papers co-authored with men. Their finding provides one channel of why gender diversity may not be beneficial. Complementing this finding, we show that gender diversity in teams does not seem to be rewarded in terms of generating better research quality.

The rest of the paper proceeds as follows. We introduce a theoretical framework in Section 2. Section 3 describes our data. In Section 4 we document homophily. Section 5 considers the correlation between homophily and the number of co-authors. In Section 6 we turn to team diversity and research output. Alternative specifications and robustness checks are presented in Section 7. Section 8 concludes.

2. Models of homophily and diversity

To discipline our analysis, we introduce the model of homophily developed by Currarini et al. (2009). Their model lends itself to our environment as agents can be one of two types (gender in our setting). Moreover, it incorporates both preferences and differential meeting rates to explain homophily and generates sharp, testable predictions. Subsequently, we introduce Prat (2002). The model presented there connects team diversity and output.

We consider here two distinct models as to the best of our knowledge there does not yet exist a model of team formation that considers several attributes of workers that also anticipates how that team will perform.[9]

2.1. Homophily measures

Denote the fraction of male authors in the population as w_m and the share of women by $w_f = 1 - w_m$. Let H_m denote the average share of male co-authors among men, while H_f represents women's average share of female collaborators. Women exhibit *relative homophily* if their share of female collaborators is larger than the share of women among authors. Formally, $H_f > w_f$. Similarly, men exhibit relative homophily if $H_m > w_m$, that is if men's share of male co-authors is greater than the share of men in our population. *Inbreeding homophily*, a measure first introduced by Coleman (1958), explicitly takes into account the gender shares among the population of economists. It is defined as

$$h_s = \frac{H_s - w_s}{1 - w_s} \tag{1}$$

[6] In contrast to both Hengel and Moon (2020) and us, Maddi and Gingras (2021) find that gender-diverse teams are associated with higher citations in economics, not controlling for journal quality of the publication.

[7] We take into account a measure of article quality but do not assess novelty.

[8] Hughes et al. (2023) experimentally show that diversity does not significantly reduce communication.

[9] This would be a multidimensional matching problem between workers. While there is work on multidimensional matching of firms and workers, it is unclear how this extends to matching within a given group.

We shall say that there is inbreeding homophily if the index is positive, inbreeding heterophily if it is negative.[10] Intuitively, the index compares the proportion of collaborations with the same gender with the fraction of this gender in the relevant population, weighted by the maximal gender bias authors could display.[11]

2.2. Homophily and collaboration

Each author can be male or female, which we refer to as their type $s \in \{f, m\}$. An author of type s forms e_s equal-type collaborations and o_s other-type collaborations. They derive utility $U(e_s, o_s)$ from their co-authors. Currarini et al. (2009) impose the usual technical assumptions on the utility function: it increases in the number of co-authors, and is continuous and concave.[12] Additionally, the utility function displays overall diminishing returns to collaboration, $U(ae, ao) < aU(e, o)$ for any constant $a > 1$.

Authors then acquire collaborators through a matching process. In our setting, the matching process could be interpreted as going to a conference. The matching process is a continuous-time process. At each unit of time there is an influx of new authors, while other economists leave. Matching is costly, capturing that going to conferences does not come for free– there are both opportunity costs as well as travel and conference costs. These costs are denoted by $c > 0$ and are incurred for each time unit spent in the matching process, independently of the outcome. Therefore, the cost of an economist, who spends time t_s looking for collaborators is given by ct_s.

As finding co-authors is costly and every economist benefits from having collaborators every economist accepts every collaborator they encounter. The question is therefore, how much time an economist spends on meeting other economists. Denote the meeting rate for same-gender economists in the matching pool by q_i. Then, each economist maximizes their utility

$$\max_{t_s} \quad U(q_s t_s, (1 - q_s)t_s) - ct_s \tag{2}$$

The assumptions on the utility function, in particular, the diminishing returns to collaboration, together with positive matching costs ensure that an economist spends a finite amount of time matching.

The utility function depends on two variables that capture the two features that determine homophily. First, a potential preference for same versus other types is captured by distinguishing between $q_s t_s$ and $(1 - q_s)t_s$. Second, there can be biases in meeting rates. If the matching process is unbiased then the probability of meeting someone of a given gender is the same for both men and women, formally $q_s = 1 - q_{-s}$. If the matching process is biased in favor of the same type, then men are meeting male authors at a faster rate than female economists. Similarly, women meet other women more quickly.

If the utility is type-neutral, every economist ends up with the same number of collaborators (from Proposition 1 of Currarini et al. (2009)), independently of the matching process they face. This is summarized in Observation 1.[13]

Observation 1. *If female and male economist display differences in the number of collaborators, economists must have a preference for same-gender collaborators.*

If an economist does not care about the gender of their collaborator, then all of them spend the same time matching and acquire the same number of collaborators.

This is no longer the case, if an author has a preference for same gender collaborators. Every author, both female and male, is more likely to meet a male collaborator given that men are in the majority. If authors display a same-type bias, men derive a higher utility from meeting men. Therefore, they spend more time matching and the matching pool skews even more towards men. Ultimately, men acquire a higher number of collaborators.

Observation 2. *Assume same-gender bias in collaboration preferences.*
(1) If there are more male than female economists, men acquire more collaborators.
(2) As the share of women increases, the collaboration gap between female and male economists decreases. As gender parity is reached, then female and male economists have the same number of co-authors.

A key feature of the model is the gradual adjustment in the gender gap in co-authors as the share of men and women becomes more balanced. Therefore, if the preference for working with same gender collaborators is crucial, an influx of women should per se diminish the collaboration gap. This collaboration gap accounts for a substantial share in the research output (Ductor et al. (2021)) and so closing it would improve the career outcomes of women substantially. The model therefore provides us with a clear prediction about the implications of a preference for same gender co-authors.[14]

[10] Note that the index lies between $-\frac{w_s}{1-w_s}$ and 1.

[11] An alternative measure has been suggested by Zeltzer (2020). His measure is similar to the relative homophily we consider but primarily pertains to directed networks (we consider undirected networks). His index is defined as the difference in the share of male collaborators of men versus women. Therefore, his index contains the same information as relative homophily.

[12] For further details, see p. 1013f.

[13] These observations are merely non-technical summaries of results stated in Currarini et al. (2009).

[14] Observation 2 builds on Proposition 4 and Proposition 9 from Currarini et al. (2009).

With an unbiased matching process, if one gender displays gender homophily, the other must display gender heterophily. If men have more male connections, then women also have more male connections. An immediate implication of this result is that for both men and women to display homophily, we require a biased matching process.[15] Therefore, for homophily to emerge both same-type preferences as well as biased meeting rates are needed. We summarize this insight in the following observation.

Observation 3. *For both men and women to display homophily, gender-based preferences as well as biased meeting rates are needed.*

Finally, we connect the extent of homophily and group size.[16] Currarini et al. (2009) show that the larger group displays a greater level of inbreeding homophily than the smaller group.[17]

Observation 4. *Assume same-type preferences and a biased matching process. Then, men, as the larger group, display greater inbreeding homophily compared to women.*

Currarini et al. (2009) find empirical support for their theory analyzing high-school friendships according to race. In contrast we focus on work collaborations across gender.

2.3. Team diversity and output

If economists display gender homophily, then collaboration teams may exhibit little diversity. We therefore investigate whether the gender composition of the research team impacts output building on Prat (2002). He analyzes the role of team diversity more generally and finds that it depends on the production environment.

In this model, a state of the world, x, is realized. In the context of academic collaboration, a state can be interpreted as current trends in research, a new method and whether it is appropriate for the research question at hand, or new findings that impact the project at hand. The state is unknown and every author i has a common prior belief about it.

There are n authors on the team and the payoff depends on the action of each agent, denoted by a_i, as well as the state of the world. Each agent chooses an information structure, which induces a partition on the set of states. A collaborator then chooses an action based on the information he has. However, the information structure comes at a cost. Overall, the team want to maximize the payoff minus the cost of all information structures across agents.

Team diversity or homogeneity is captured by the information structures. If all team members have the same information structure, then we refer to a team as homogeneous, while different information structures imply diversity. In the context of men and women, this could mean varied perspectives or differential access to information.

If men and women have different information structures, then we should observe different outcomes for male, female and mixed teams. If they don't then all three types of teams should perform the same. This motivates our next observation.

Observation 5. *If men and women have the same information structures, then female, male and mixed teams perform the same.*

If men and women have different information structures, then we require an analysis of whether the same or different information structures are beneficial. This depends on the features of the production function. In particular, it matters whether the production function is super-or sub-modular in collaborator's actions. Sub-modular functions imply that same actions are better than distinct actions. This means that one person choosing a higher action, while the other person chooses a lower action is worse than both choosing an intermediate action. The chosen action depends on the individual information structure. Therefore, different structures induce varying actions. If the production function is sub-modular, then the payoff with varying information structures and actions is worse than the payoff if everyone takes the same action. Ultimately, for a sub-modular production function, homogeneity is best.

Observation 6. *If the production function is sub-modular, then a homogeneous team performs best. If the production function is super-modular, then a heterogeneous team is best.*

To illustrate how Prat (2002) could be interpreted in the context of collaborations, we present an example with two co-authors. One co-author focuses on the theory section of the paper, the other on the empirical part and the quality of each section produced determines how well the paper can be published. Denote by $a_1 \in [0,1]$ the quality of the theory part, while $a_2 \in [0,1]$ measures the quality of the empirical part. The state of the world $x \in [0,1]$ could for instance indicate the taste among editors for a paper that combines theory and empirical work. Then, consider the following team production function:

$$x \min\{a_1, a_2\} - k(a_1 + a_2), \tag{3}$$

[15] An example of a biased process is $q_m^\beta + q_f^\beta = 1$ with $\beta > 1$.

[16] We base our next observation on Proposition 4 from Currarini et al. (2009).

[17] Their results are derived using unbiased matching, in which case the smallest group displays heterophily. This result can be reversed by allowing for a biased matching process as discussed on p. 1025.

where k is some strictly positive constant. It is easy to verify that this is a sub-modular production function for any x.[18] Therefore, a team with the same information structure performs best. Here, the qualities of the different sections are complements. If collaborator 1 learns that x is high, while co-author 2 faces greater uncertainty, resulting in lower quality, then the publication outcome is the same as if both collaborators have co-author 2's information. However, in the second case production costs are lower, therefore making the same information structure beneficial. The contrary would hold if the quality of the different sections were substitutes. In this case, we would face a super-modular function and different information structures were beneficial.

The model introduced by Prat (2002) allows us to identify whether homogeneous or diverse teams are best. However, it does not provide a reason for why an all female may perform better than an all male team or vice versa. Standard explanations for disparities in performance across gender can be parsed into demand and supply side factors.

On the demand side, women may be held to higher standards as most recently documented in Hengel (2016), Alexander et al. (2021).[19] Such an explanation could be incorporated in our model by condition the state on gender, x_s. Then, x_m could be higher than x_f and women's output is lower for the same information structure.

Alternatively, on the supply side, women could face a higher cost of generating the same information structure than men, leading to lower actions. This would be reflected in a gender specific cost multiplier $k_f > k_m$. These cost disparities may reflect that men have larger professional networks as documented in Lindenlaub and Prummer (2021). Additionally, women may face higher costs due to family constraints or additional requirements on the job (Morgan et al. (2021), Antecol et al. (2018), Babcock et al. (2017)). This motivates our next observation.

Observation 7. *If same gender teams perform differently, then the production function differs in the state of the world, x_s and/or costs, k_s.*

Finally, we connect the performance of the three different teams, female, male, diverse back to the structure of the production function, state of the world and costs. If both same gender teams outperform mixed teams or mixed teams produce better than both same gender teams, then the shape of the production function is crucial in determining output. If male teams perform better than mixed teams which in turn outperform female teams (or the reverse holds true), then supply and demand factors seem more important than the structure of the production function. After all, in this case, the discrepancy between same gender teams is greater than the difference between any of them and the diverse team.

Notably, Prat (2002) does not allow for communication between agents. If one agent receives information, he cannot simply share it. This assumption has been recently micro-founded by theoretical work that demonstrates how diversity can lead to reduced communication (Kets and Sandroni (2021), Dong and Mayskaya (2023), Dong et al. (2023), Hughes et al. (2023)), indicating that this restrictive assumption does not invalidate the model.[20]

3. Data

We use two data sets for our analysis, EconLit and the Web of Science, which we describe in turn.

The EconLit database is a bibliography of journals in economics compiled by the editors of the *Journal of Economic Literature*. The database provides information on 921,976 articles published between 1970 and 2017, in 1990 journals. We do not cover working papers and work published in books. We focus on research papers with at most 3 co-authors as EconLit does not report the names of all the authors for articles published by more than three authors before 1999; therefore, we exclude these articles from the analysis for the period 1970-1999. Articles published by four or more authors represent 1.6% of all the articles published between 1970-1999.[21]

Each article registered in EconLit contains information on the journal (including name of the journal, volume, issue, first and last page), title, the last and first name of each author, affiliations of each author, JEL codes, keywords and the abstract.[22] Authors are identified by their first and last name, as in Goyal et al. (2006), leaving us with 470,309 authors. Using information about all the articles published by an author in our sample period, 1970-2017, we construct a panel that starts for each individual with their first publication and extends to the last observed publication of the author (or to 2017). This allows us to describe an author's collaborators.

For our measures of research quality, we supplement the EconLit data with citations and references from the Web of Science (hereafter, WoS) (Clarivate Analytics, 2018). For this latter exercise, we focus on the 100 most established journals in economics according to IDEAS/RePEc, see also Ductor et al. (2020).[23] The citation and reference data set includes information on 275,670 articles and the number of citations they received yearly until 2017.

We identify the gender of an author using their first names and *gender-api.com*, a source that provides first names and the estimated gender for 201 countries. We identify an author's gender if the author's first name is associated with a single estimated gender in the

18 See Eq (3) in Prat (2002) as well as the Appendix A.1. on p. 1203.

19 This is consistent with women experiencing discrimination.

20 However, Hughes et al. (2023) show in the lab that information sharing is not significantly reduced, even though this would increase payoffs. Therefore, further work is needed to understand whether information sharing is indeed impacted by diversity. See also the overview provided in Page (2008).

21 Goyal et al. (2006) show that the co-authorship network statistics are unaffected when articles with four or more authors are included.

22 Affiliations are only available for articles published after 1989. Abstracts are only available for articles published after 1999.

23 More precisely, we take the top 100 journals from the Simple Rank list over all years.

201 countries, at least 95% of the time. Using this standard approach allows us to identify the gender of 78% of the authors (367,441 out of 470,309 authors).[24]

4. Homophily

We begin by documenting gender homophily among economists.

4.1. Estimation

Considering that the share of female authors in economics has increased substantially, we measure the share of each gender across a 5-year window and index this measure with t. The share of each gender at time t is the number of authors of gender s, present in year $t-4$ to year t. We select a 5-year window in line with Ductor et al. (2018), to capture the available collaborators at the beginning of a project. We focus on a five-year time frame as it is well known that there are long lags in publication. Therefore, available co-authors of a paper do not coincide with the publication date.[25] We amend our previous measures of homophily by incorporating a time index. The proportion of collaborators of gender s at time t is denoted by H_{st}. Similarly, w_{st} measures the proportion of authors at time t and gender s. Similarly, the inbreeding homophily h_{st} at time t for gender s is denoted by

$$h_{st} = \frac{H_{st} - w_{st}}{1 - w_{st}}. \tag{4}$$

The measure h_{st} provides the average inbreeding homophily for men and women at time t. We additionally measure inbreeding homophily at the author level: we replace the overall share of same-gender collaborations with H_{it}, the number of same-gender collaborators author i has and obtain h_{it}.[26]

Focusing on a given author allows us to condition on their observable characteristics. We estimate the following regression using Pooled OLS,

$$h_{it} = \alpha + x_{it}\beta + \epsilon_{it}, \tag{5}$$

separately for each gender. We control for a number of variables collected in x_{it}. Authors may display different collaboration patterns over the course of their career. We therefore account for experience through career time indicators, which are defined as the number of years since the first publication by the author.[27] Collaboration could also vary across fields of research. We therefore control for it. Following Fafchamps et al. (2010), we categorize 19 different fields using the first digit of JEL codes and include a measure of the proportion of publications in each JEL code. These codes capture the fields of specialization of the author. We also include year fixed effects to account for time trends as collaborations have changed over time. Further, we control for past research output, which is the accumulated output from the first publication until $t-1$, evaluated by their article influence score. This takes into account that some collaborators may be more appealing due to past output, which could also affect their homophily. Formally, past research output is defined as

$$Q_{it} = \sum_{p=1}^{P} \frac{AIS_p}{\#\text{authors}_p}, \tag{6}$$

where P are all publications of author i up to time $t-1$ and AIS_p denotes the article influence score. We weight every article by the number of authors.

Overall, the empirical model described by (5) allows us to generate a conditional homophily measure.

4.2. Findings

Men display relative homophily if their average share of male co-authors is higher than the fraction of male authors in the population. We compute in Table 1 the percentage of links within gender and find that on average 80.5% of men's collaborations are with other men: this is higher than the fraction of men in the population, 72%. Women also exhibit relative homophily as their collaboration with other women, 34.1% exceeds the fraction of women in the population, 28%. Therefore, both men and women display relative homophily.[28]

While relative homophily does not take into account the varying shares of men and women in the profession, inbreeding homophily incorporates this change. Fig. 1a documents that both men and women exhibit inbreeding homophily, collaborating more with authors of the same gender. This pattern is more pronounced for men whose inbreeding homophily is higher compared to the

[24] Given that our data spans the universe of authors publishing in Economics over a 47 year horizon, it is prohibitively costly to conduct a google search to identify the gender of the remaining 100 000 authors.

[25] Alternative time windows lead to qualitatively identical results.

[26] We omit here an explicit marker of gender as we keep track of it through the individual author.

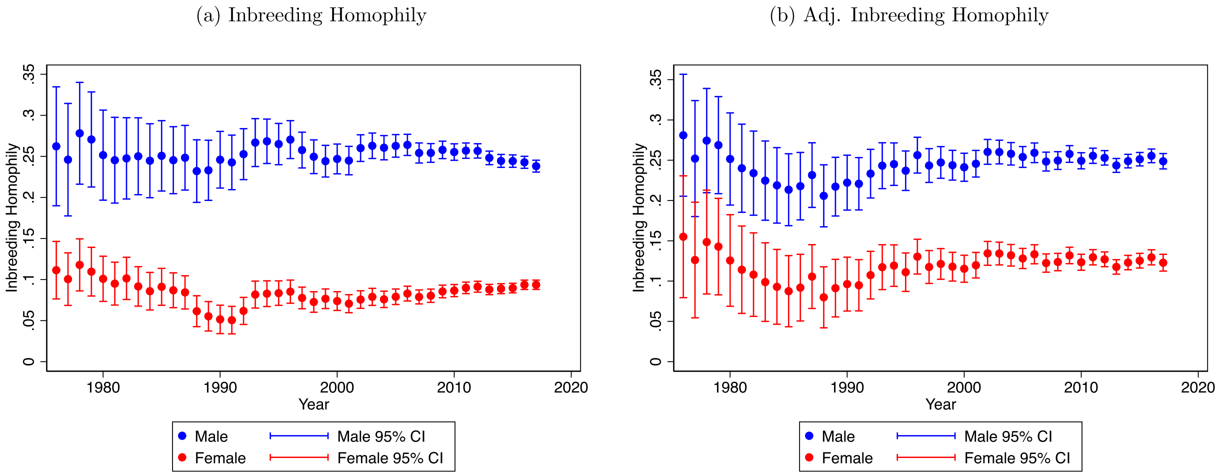
[27] While the Ph.D. graduation date is arguably a better proxy for experience, since the timing of the first publication may differ across gender, we refrain from doing so as gathering this information for over 367,000 authors is prohibitively costly.

[28] Figure A.1. in the Supplementary Appendix highlights that relative homophily exists for the entire sample period.

Table 1
Summary statistics.

	Men	Women	Men-Women
Population Share	0.72	0.28	0.44***
Men's Collaborators (share)	0.805	0.195	0.610***
Women's Collaborators (share)	0.659	0.341	0.318***
Inbreeding Homophily	0.30	0.08	0.22***
# Co-authors	2.31	2.13	0.18***
Co-authorship	0.66	0.72	-0.06***
Experience	10.37	7.08	3.29***
Past output	1.13	0.68	0.45***

Note: The sample includes all articles published in EconLit from 1970 to 2017, where the gender of at least one author is identified. The unit of analysis is at the author level. The # coauthors is the number of different coauthors accumulated from $t-4$ to t. Coauthorship is the ratio between the number of coauthored articles and the total number of articles published from $t-4$ to t. Experience refers to time since first publication, past output as defined in expression (6). Column 3 shows the differences in means between men and women. ***p < 0.01 implies that the difference between men and women is significant at the 1% level.



Note: Adjusted inbreeding homophily is the conditional mean of inbreeding homophily, calculated from model (5). Controls include career time, share of JEL codes, past research output, and year.

Fig. 1. Inbreeding homophily.

one for women, as predicted by Currarini et al. (2009). Remarkably, the level of inbreeding homophily is stable over time, both for men and women. Considering the whole sample period, 1970-2017, the inbreeding homophily of men and women are equal to 0.30 and 0.08 (see Table 1), respectively. Inbreeding homophily remains stable even if we control for experience, field of research, past research output and year fixed effects, see Fig. 1b, which is based on empirical model (5).

Our result highlights that despite the increase in the share of women in Economics, authors' share of male and female collaborators has remained unchanged. Interestingly, the increase in the share of women has not led to a higher share of female co-authors, both for men and women.

4.3. Discussion

Unfortunately, our data does not permit us to distinguish between the potential drivers of homophily, such as differential meeting rates within groups, preferences or a mix of the two. This is a common issue in the literature on the evaluation of homophily, see for instance Graham (2016). Recently, Jackson et al. (2023) have tackled these questions, focusing on malleable (for instance hours playing video games) and permanent characteristics (such as gender). They find that both matter. Therefore, we follow the interpretation that homophily is an equilibrium phenomenon, shaped by both preferences and differential meeting rates. This interpretation is also in line with the theory developed by Currarini et al. (2009), where a biased meeting process as well as a same-gender preference are required to generate homophily for both men and women.

5. Homophily and collaboration

Having established gender homophily, we analyze if it limits collaboration opportunities. Collaborations are increasingly important in Economics. In particular, a higher number of co-authors is associated with a higher quality research output (Ductor et al. (2018), for a theoretical foundation, see Lindenlaub and Prummer (2021)).

5.1. Estimation

We begin by investigating whether there is a positive correlation between the relative size of a group in the population and the average number of co-authors. We first exploit the variation in the share of women over time and second, the variation in women across fields. We study if the gender gap in the number of coauthors from $t-4$ to t diminishes over time, as the share of women in economics increases. The number of collaborators is denoted by d_{it}, which abbreviates the degree of the author. We focus on cohorts as co-authorship may also be driven by similarity in career time. Specifically, we estimate the following model:

$$d_{it} = \alpha F_i + \sum_{c=1975}^{2017} \gamma_c C_{itc} + \sum_{c=1975}^{2017} \delta_c F_i \times C_{itc} + x_{it}\beta + \epsilon_{it}. \tag{7}$$

The variable F is an indicator variable equal to one if the author is female. The cohort denoted by C is an indicator variable equal to one for the year of first publication. The additional controls are identical to those included in empirical model (5) and as such are denoted once again by x_{it}. In this specification, we define past output as the accumulated output from the first publication until $t-5$ to avoid simultaneity with the number of co-authors. We focus on the coefficients of the interaction terms between female and the cohort indicator variables. They capture if the gender gap in the number of co-authors is changing across cohorts allowing us to investigate whether an increase in the share of women can erase potential disparities in terms of the number of collaborators.

Second, we exploit variation in gender shares across fields. Here we use the first two digits of the JEL codes, to define 124 different fields and obtain the average number of co-authors per JEL code, l, and the share of women per JEL code, w_{fl}. We de-trend degree by regressing degree on year indicators, the residual from this regression is the de-trended degree.[29] Finally, we estimate the association between de-trended degree d_l^{det} and the share of women in a field using the linear model

$$d_l^{det} = \alpha + \beta w_{fl} + \epsilon_l. \tag{8}$$

5.2. Findings

On average, men have a significantly higher number of co-authors relative to women, see Table 1: the average number of co-authors for men amounts to 2.31 while women's average number across a five-year period is 2.13.

We investigate how the number of collaborators have responded to changes in the gender composition in the economics profession. We first connect the share of women in a given cohort to the change in the gender gap in the number of collaborators. Our findings are presented in Fig. 2. The right vertical axis refers to the share of female authors in a given cohort, that is the share of women who published for the first time in a given year. The share of women publishing has been steadily increasing over the years. While women account for less than 10% of authors in the 1975 cohort, they represent 35% of first publications at the end of our sample.[30]

We estimate the relationship between the number of co-authors and the share of women per cohort using the empirical model described in (7). To evaluate how the gender gap evolved throughout our period of observation, we provide the coefficients as well as the 95% confidence interval of the interaction terms between the cohort indicators and the female indicator on the left vertical axis of Fig. 2. All the estimates are relative to the base cohort, 1974. Contrary to Currarini et al. (2009)'s prediction, we find that the gender difference in the number of collaborators is even increasing for the most recent cohort of economists: women who published their first article in 1974 have 0.18 fewer co-authors than men, while women of the 2011 cohort have 0.87 fewer co-authors than men.[31] We find that the discrepancy in the number of collaborators for men and women was less pronounced in 1974 than in 2017. This documents that the gender collaboration gap has increased despite the rise in the share of women.[32]

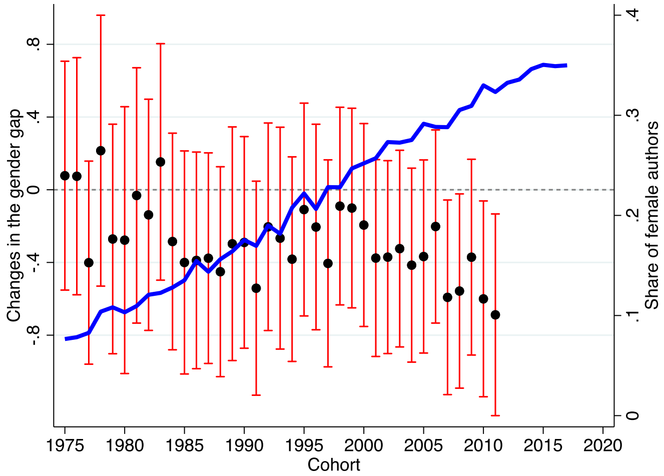
Second, we exploit variation in gender shares across fields and tie it to the number of collaborators using empirical model (8). In Fig. 3, we observe that the relationship between degree and relative group size is weak. Regressing the average degree per field on the relative group size per field we obtain a slope coefficient of 0.014, which is statistically insignificant (p-value = 0.18).

[29] The results are robust to other de-trending methods, available upon request.

[30] This is consistent with the higher share of women in Economics documented in Ductor et al. (2021).

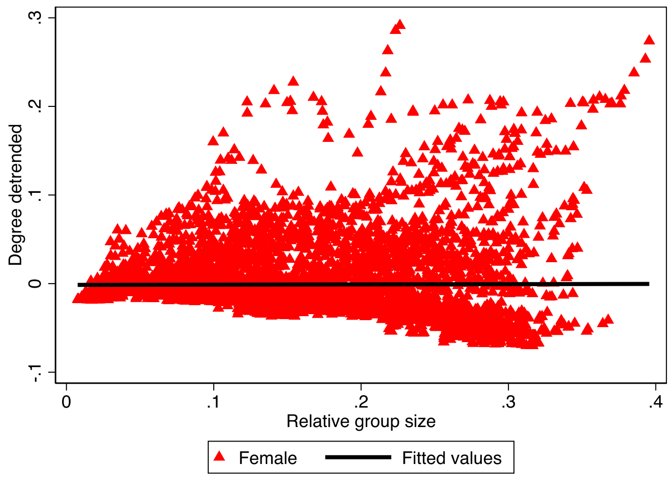
[31] The p-value of an F-test on the joint significance of the coefficients of the interaction terms of gender and time is 0.09 suggesting that the observed increase in degree over recent cohorts is jointly significant at the 9% level.

[32] Our results provide a lower bound on the gender gap in the number of collaborators: if we focus instead on the share of women at a given time and how it relates to the number of collaborators, we see an even larger increase in the gender gap.



Note: The blue line shows the share of female authors across different cohorts. The black dots and red lines are the coefficients and 95% confidence intervals of the interaction terms between cohort indicators and the female indicator added to the degree network model (7) estimated using POLS, the base cohort year is 1974.

Fig. 2. Gender gap in # of co-authors across cohort. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)



Note: Degree is detrended is the residual of a linear regression of degree on year indicators. Regressing the degree detrended on relative group size, we obtain: $\hat{d}_l^{det} = -.004 + 0.014 w_{fl}$, the p-value of the intercept and slope coefficients are 0.01 and 0.18, respectively.

Fig. 3. Degree and fraction of women, across fields.

An increase in the share of women in the Economics profession does not ameliorate the gender gap in collaboration, which is key to a reduction in the output gap.[33] Even though economists display gender homophily, the influx in the number of female economists has not led to a higher number of collaborators for women, and thus in turn, it has not helped close the gender output gap.

5.3. Discussion

Our results are at odds with the theoretical predictions of Currarini et al. (2009). For a given matching process and same-type preferences, an increase in the share of women should ameliorate the collaboration gap. Given that this closure has failed to emerge, it must hold that either the matching process has become more biased and/or same-type preferences are stronger.

Such a change is not only plausible in the context of the theory, but also supported by additional empirical evidence. For instance, Gertsberg (2022) has documented that the gender gap in co-authors has been exacerbated by the Me Too movement. Women collaborate less with their colleagues at their own universities and are not able to increase the number of co-authors from

[33] Doan et al. (2023) document a causal effect of number of collaborators on research output, e.g. loosing a coauthor leads to a long-run decline in research output of approximately 11%.

outside their institution.[34] We speculate that in light of the incidents of the Me Too movement a stronger meeting bias or preference for same-gender authors could have emerged.

6. Team diversity and research output

We examine gender diversity at the team level and its relationship to research quality. Gender homophily has implications for the gender composition of research teams. While homophily is measured at the author level, we also track gender diversity at the research team or article level.

6.1. Estimation

To assess gender diversity we create a categorical variable for which we classify teams as male if they exclusively consist of men, female if only women collaborate and mixed if there are female and male co-authors. If we can identify one author as male, another one as female, while the third remains unidentified, the team is categorized as mixed.[35] We further create a variable that measures the share of women on a research team as well as an explicit measure of gender diversity, the Blau index. A drawback of the Blau index is that it does not distinguish between all female and all male teams, which is problematic if male and female teams differ in terms of output. For further details on these alternative measures, see the Supplementary Appendix.

We assess the connection between gender team diversity and output quality. We focus on two measures for the quality of the research output. Our first measure is the article influence score, an impact factor of the journal an article is published in. We construct the article influence score, AIS_p, for each article p following Ductor et al. (2020). The advantage of the AIS is that it is time varying, it excludes self-citations, and it considers the influence of the citing journal (see Bergstrom et al. (2008) and Palacios-Huerta and Volij (2004) for further discussions on the virtues of AIS).[36]

Second, we consider the number of citations an article attracts. Citations capture whether a publication has an impact on future research. Higher citations may represent articles that introduced novel concepts, methods, or findings shaping the perspectives of researchers on a particular topic.

Research quality, q_{pt}, can either refer to the article influence score or the number of citations. We use Pooled OLS to estimate

$$log(q_{pt}) = \rho D_p + x'_{pt}\beta + \epsilon_{pt}, \tag{9}$$

where D_p denotes the team diversity. We include once again a number of controls, which contain a constant, and denote them by x'_{pt}. We control for the number of authors on the team as a larger team may allow for different skills, potentially affecting research quality. Additionally, we control for the specialization of the research team. This captures how many articles the collaborators have published previously, taking into account the research field. Having more experience of publishing in a given field may relate to quality. As in Ductor (2015), we count the number of papers the research team has published in a given field and calculate the Herfindahl index $\sum_{l=1}^{F}\left(\frac{n_{lt}}{n_t}\right)^2$. The variable n_{lt} denotes the number of publications in field l, while n_t captures the overall number of publications by the research team of article p up to t. Articles with multiple JEL codes are assigned proportionally to each field. If there is no previous article published by a given team, we normalize the Herfindahl index to zero.

We further control for the past output of the research team, defined in the same spirit as past output per author: we count the overall number of papers, weighted by journal quality, published by the team that produced article p. This variable captures the track record of the team. We also take into account the average past output of the co-authors. As above, we calculate the past output for each co-author and take its average. This variable control for the track record of the team's members.[37]

We include team experience as the number of years since the first publication of the research team, which relates to seniority. We take into account the length of the article p by counting the number of pages which has shown to be related to quality (Card and DellaVigna (2013)). In addition, we include year and field fixed effects, the latter measured by JEL codes. Standard errors are clustered at the team level as research quality is correlated over time.

We focus here on team characteristics given that this is the relevant unit of observation. We further provide results that take into account individual characteristics, for instance, each author's past output, and experience in the Supplementary Appendix.[38]

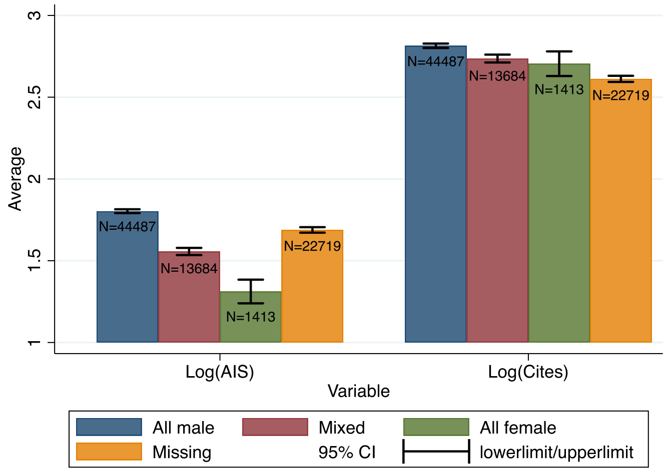
[34] Female academics across disciplines collaborate more locally in general, see Kwiek and Roszka (2021) and online interactions do not seem to have ameliorated this problem.

[35] Our classification of teams differs from Card et al. (2020) and Hengel and Moon (2020) as we focus explicitly on gender diversity.

[36] We consider the Article Influence Score as the main measure of journal quality because it is the only journal ranking method that satisfies the axiomatic properties of weak consistency, weak homogeneity, invariance to reference intensity and invariance to splitting of journals (Palacios-Huerta and Volij, 2004). Besides, the AIS is used in some universities – e.g., the Erasmus University of Rotterdam – to evaluate the research performance of their faculty. The National Agency for Quality Assessment and Accreditation of Spain (ANECA) also uses a version of the Article Influence Score (the Eigenfactor Score) to evaluate the research output of Spanish researchers.

[37] We focus here on averages. Alternatively, we could control for the highest output of a co-author or the highest seniority of the co-author team. Including such controls does not affect our results.

[38] Our sample consists of papers with two and three authors. To control for author-specific characteristics we are required to make an assumption about the missing third author's characteristics on a two-researcher team as e.g. his past output is not defined. We set missing characteristics of the third author to zero and include an indicator for a two person team.



Note: Sample includes articles published by two and three authors in the 100 leading journals from 1970 to 2017. *Men* corresponds to articles published only by male authors, *Mixed* is a research team composed of male and female authors, *Women* include articles published only by female authors, *Missing* is a research team with an unidentified gender . The left-bars present the average Log(AIS) and 95% confidence interval for each gender team composition, the right bars show the average Log(Cites+1) and 95% confidence interval for each gender team composition.

Fig. 4. Gender team composition: descriptive statistics.

6.2. Findings

We begin by providing descriptive statistics on the gender composition of research teams, before connecting it to the quality of research output. As citations are not available for all journals listed in EconLit, we restrict attention to articles published in the 100 most established journals in economics.[39]

Descriptive statistics for research teams, or equivalently, at the article level are provided in Fig. 4. We focus on research teams consisting of two and three authors.[40]

Exclusively male teams make up 54% of all collaborations. In comparison, only 2% of all research teams are entirely female. Mixed teams comprise 17%, and the remainder are teams with a missing gender composition.

We find unconditional differences in the quality of the research output according to gender team composition. The article influence score is significantly higher for male teams compared to mixed teams. In turn, gender-diverse teams possess a significantly higher article influence score compared to purely female teams. This pattern partially carries over to citations: male teams produce research articles that are more highly cited relative to those written by mixed or female teams. However, there is no distinction between the average number of citations that mixed vs female teams produce, see column 1 of Table 3.[41]

Moving beyond descriptives, we estimate model (9) using the article influence score as the dependent variable. Our results are presented in Table 2.

When controlling for observables, a mixed team publishes, on average, articles in journals with 6.1% lower article influence score than those published by male team, while female teams publish on journals with 8% lower article influence score than those of male teams. The gap between mixed and male teams is not significantly smaller than between female and male teams.

A higher number of authors on a team correlates with a lower article influence score.[42] Authors' average past output and team past output are positively associated with journal quality, which reflect a positive impact of past success on publications. Controlling for authors' past output and team past output, team experience is negatively associated with journal quality. Authors that collaborate with the same coauthors tend to publish in worse journals than authors who collaborate with a different set of coauthors. This seems to indicate that diversity matters– just not in terms of gender. Consistent with Card and DellaVigna (2013), the length of an article is positively associated with the article influence score.

[39] This reduces our sample from 367 k authors to 327 k authors. We opted to use all available information to measure homophily. Our homophily results remain unchanged if we perform our analysis for the authors that published at least once in one of the Top 100 journals.

[40] Articles published by more than three authors represent 4.3% of all the articles published between 1970 to 2017 in the 100 journals. We include single-authored articles as a robustness check when we consider the fraction of women as independent variable, see Tables B.1 and B.2 in the Supplementary Appendix. Our results are qualitatively unchanged, but do not capture the impact of team diversity on research output.

[41] Research teams with a missing gender composition have predominantly Asian last names. It seems like Asian research teams are cited less compared to other papers.

[42] This is in contrast to the results of Bramoullé and Ductor (2018) who find a positive relationship between the number of authors and journal impact. This difference is driven by the exclusion of single-authored papers.

Table 2
Team gender composition and article influence score.

Variables	Dependent Variable	
	Log(AIS)	
	(1)	(2)
Mixed	-0.246***	-0.061***
	(0.015)	(0.011)
Female	-0.491***	-0.080**
	(0.041)	(0.033)
Missing gender	-0.114***	-0.043***
	(0.012)	(0.009)
Past output team		0.429***
		(0.051)
Avg. past output authors		0.827***
		(0.008)
Team specialization		-0.004
		(0.021)
Team experience		-0.023***
		(0.002)
Pages		0.026***
		(0.000)
# authors		-0.033***
		(0.008)
Constant	1.803***	1.377***
	(0.007)	(0.074)
Year FE	–	✓
JEL codes FE	–	✓
Test Mixed = Female (p-value)	32.7(0.00)	0.31(0.58)
Observations	82,303	82,303
R-squared	0.007	0.409

Note: In columns 1 to 2, we estimate the relationship between Article Influence Score and gender diversity, the dependent variable is in log. Sample restricted to two- and three-authored papers. Pages is the number of pages of the article; Number of authors is the number of authors publishing the article; Team specialization is a Herfindahl index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; Past output team is the number of papers adjusted by quality that the research team has published together in the past; Avg. past output authors is the average number of papers adjusted by quality published by the authors of the research team in the past. All the regressions use clustered standard errors at the team level. ***p < 0.01, **p < 0.05, *p < 0.10.

We turn to citations and relate gender team composition to the number of citations the article receives. Our results are presented in Table 3. While male teams unconditionally generate higher citations, this pattern is reversed once we control for observables. A female team tends to receive higher citations followed by a gender diverse team, see column (2) of Table 3. On average, all-female teams publish articles that are 25% more cited than those of all-male teams. Mixed teams publish articles correlated with 4% higher citations compared to male collaborators. The gender gap in citations persists if we additionally control for the article influence score, that is the journal quality in which the paper was published. Then a hypothetical switch from a male to an otherwise observationally equivalent all-female team is associated to an increase in citations by 26%. Similarly, the citations gap between mixed and male teams increases to 5%.

Authors' past average output as well as the team's past output are positively associated with citations. In contrast, team experience is negatively associated with the number of citations, consistent with the findings presented in Table 2. Specialization is positively related to citations, in line with the existence of positive returns to specialization in the publication process. Papers written by more authors are also more cited.

In sum, we find that a woman on the team ties in with a lower article influence score. However, conditional on authors' and team characteristics, female teams generate a higher number of citations.

6.3. Discussion

Gender homophily has implications for the gender composition of research teams, which are often homogeneous. Gender homogeneous teams generally outperform gender-diverse teams: male teams outperform female and mixed teams in terms of journal

Table 3
Team gender composition and citations.

Variables	Dependent Variable		
	Log(Citations+1)		
	(1)	(2)	(3)
Mixed	-0.078***	0.044***	0.054***
	(0.016)	(0.013)	(0.013)
Female	-0.110***	0.246***	0.258***
	(0.041)	(0.031)	(0.031)
Missing gender	-0.202***	-0.081***	-0.075***
	(0.014)	(0.011)	(0.011)
Team specialization		0.049*	0.049**
		(0.025)	(0.025)
Past output team		0.647***	0.578***
		(0.064)	(0.065)
Avg. past output authors		0.313***	0.181***
		(0.010)	(0.010)
Team experience		-0.030***	-0.026***
		(0.002)	(0.002)
# authors		0.142***	0.147***
		(0.010)	(0.010)
Pages		0.031***	0.027***
		(0.000)	(0.000)
Log(AIS)			0.160***
			(0.004)
Constant	2.814***	1.671***	1.452***
	(0.008)	(0.076)	(0.075)
Year FE	–	✓	✓
JEL codes FE	–	✓	✓
Test Mixed = Female (p-value)	0.54(0.46)	40(0.00)	40(0.00)
Observations	82,303	82,303	82,303
R-squared	0.004	0.333	0.345

Notes: In columns (1) to (3), we estimate the relationship between cumulative citations from year of publication to 2017 and gender diversity, the dependent variable is in $\log(x+1)$. Sample restricted to two- and three-authored papers. Pages is the number of pages of the article; Number of authors is the number of authors publishing the article; Team specialization is a Herfindahl index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; Past output team is the number of papers adjusted by quality that the research team has published together in the past; Avg. past output authors is the average number of papers adjusted by quality published by the authors of the research team in the past. All the regressions use clustered standard errors at the team level. ***p < 0.01, **p < 0.05, *p < 0.10.

quality, while articles written by women attract the highest citations across the 100 leading journals in Economics, conditional on journal quality.[43] Notably, male teams perform unconditionally better across both research metrics.

In light of Prat (2002), we interpret our findings as indicating that men and women have different information structures. Only then, do female, diverse and male teams differ in terms of their research output, conditional on facing the same state and identical costs. Moreover, there seems to be an advantage to a homogeneous team: the gender diverse team never outperforms a homogeneous team. The research production function is therefore more likely to be sub-modular, meaning the contributions of co-authors are complements rather than substitutes. Given that men perform differently compared to women, it must hold that the production function for men and women differs in the state of the world and/or costs. The former captures demand side explanations, while the latter incorporates supply side factors.

A common explanation for gender disparities are differences in preferences. In our setting this could mean that women prefer to write papers that accrue higher citations, while men write papers that publish better. Note however that men's papers are cited more. Therefore, even if women prefer to maximize citations they would still want to publish in the highest ranked journals. Alternatively, differences in preferences could be captured through different states: if women are more interested in research that is currently less fashionable, then this may result in lower publication quality. However, their publications could gain more citations over time if interest in their topics grows.

[43] Our finding confirms Card et al. (2020), who document the same pattern for four journals.

7. Alternative measures, methods and samples

Given our data, we cannot establish a causal relationship between team composition and research quality. However, we document in this section that our correlations are robust to alternative measures of diversity and estimation methods. These results are provided in the Supplementary Appendix.

Alternative measures of gender diversity First, we show that our results are robust to a different measure of gender team diversity. We consider the fraction of women on a team as the independent variable of interest. Our results are presented in Tables B.1 and B.2. Once again, having a higher share of women is related to a lower article influence score, but a higher number of citations. Moving from a team without a woman to an all-female team is related to a decrease in the article influence score by 8.8%, while citations rise by 15.4%.

We then use the Blau diversity index instead of the indicator variable for gender composition. Note that the Blau diversity is difficult to interpret as all female and male teams are assigned the same number– despite the differences in research quality these teams are connected with. With this caveat, we find that gender team diversity is negatively associated with the article influence score. Moving from a gender-homogeneous team to a gender-balanced one is related to a 5% lower article influence score. We further find that the association between team diversity and citations is positive and statistically significant. Switching from zero diversity to gender balance is associated with a 6.7% increase in citations.

Alternative econometric models We also estimate the negative binomial for citations and other measures of gender team composition. Our results remain qualitatively unchanged, see Table D.8.

Alternative controls We further control for individual past output and individual experience instead of team characteristics. The results presented in Tables C.4 and C.5 are robust to these different control factors.

We then account for past collaboration between a subset of authors if the team consists of three co-authors. We calculated an alternative experience measure, namely the average experience of bilateral collaborations. This means we calculate the experience of authors 1 and 2, 1 and 3 and 2 and 3 and divide by three. In the same specification, we also provide an additional measure for specialization to capture past collaboration between some of the authors on a team. We calculate the Herfindahl index for each pair of authors. We weight each of the indices by the number of years the authors worked together relative to the total years of the three pairs of authors and add them. Our results remain qualitatively unchanged and are quantitatively very similar, see Tables C.6 and C.7.

Top 5 publications Finally, we restrict attention to Top 5 publications, which are summarized in the Supplementary Appendix, Section E. These findings are qualitatively in line with our findings for 100 leading journals in Economics, as well as with Hengel and Moon (2020).

We further analyze the implications of measurement error for our estimates in the Top 5 sample. We compare gender differences in citations using the gender database created by Hengel and Moon (2020) with our gender identification approach that relies on genderapi.com. Hengel and Moon (2020) accurately identify the gender composition of all research teams publishing in a Top 5 journal. The results show that measurement error in the gender team composition is attenuating the observed gender differences in the citation model (see Table E.12). Female teams tend to accumulate 37.4% more citations when we use the database created by Hengel and Moon (2020), while it is 28.8% when we rely on genderapi.com. This suggests that the positive gender gap in citations for women serves as a lower bound in our full sample of journals.

8. Conclusion

We document that gender homophily is present in Economics and has remained remarkable stable over a 47 year period. Gender homophily together with the under-representation of women in Economics could explain why women have fewer collaborators compared to men. If this is the case, then according to Currarini et al. (2009), an increase in the share of women will ameliorate the gender gap in collaborators. We find no evidence of this. Our result is consistent with stronger gender preferences or a more biased meeting rate among authors over time. Homophily induces gender homogeneous teams, and it turns out these teams are correlated with the highest research output. Therefore, a preference for same-gender authors does not harm productivity. However, male teams deliver higher-impact publications, while female teams produce articles that accrue the highest citations conditional on the publication quality and other observables. These gender discrepancies across different quality metrics raise questions about how to evaluate research appropriately. Moreover, our data does not allow to distinguish between potential drivers and mechanism of these gender disparities. We leave this for future research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jebo.2024.03.027.

References

Alexander, D., Gorelkina, O., Hengel, E., Tol, R., 2021. Gender and the time cost of peer review.
Antecol, H., Bedard, K., Stearns, J., 2018. Equal but inequitable: who benefits from gender-neutral tenure clock stopping policies? Am. Econ. Rev. 108 (9), 2420–2441.
Azmat, G., Boring, A., 2020. Gender diversity in firms. Oxf. Rev. Econ. Policy 36 (4), 760–782.
Azmat, G., Petrongolo, B., 2014. Gender and the labor market: what have we learned from field and lab experiments? Labour Econ. 30, 32–40.
Babcock, L., Recalde, M.P., Vesterlund, L., Weingart, L., 2017. Gender differences in accepting and receiving requests for tasks with low promotability. Am. Econ. Rev. 107 (3), 714–747.
Beaman, L., Keleher, N., Magruder, J., 2013. Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. Working Paper.
Bergstrom, C.T., West, J.D., Wiseman, M.A., 2008. The eigenfactor™ metrics. J. Neurosci. 28 (45), 11433–11434.
Boring, A., 2017. Gender biases in student evaluations of teaching. J. Public Econ. 145, 27–41.
Bramoullé, Y., Currarini, S., Jackson, M.O., Pin, P., Rogers, B.W., 2012. Homophily and long-run integration in social networks. J. Econ. Theory 147 (5), 1754–1786.
Bramoullé, Y., Ductor, L., 2018. Length. J. Econ. Behav. Organ. 150, 311–324.
Card, D., DellaVigna, S., 2013. Nine facts about top journals in economics. J. Econ. Lit. 51 (1), 144–161.
Card, D., DellaVigna, S., Funk, P., Iriberri, N., 2020. Are referees and editors in economics gender neutral? Q. J. Econ. 135 (1), 269–327.
Chari, A., Goldsmith-Pinkham, P., 2017. Gender representation in economics across topics and time: Evidence from the nber summer institute. Technical report. National Bureau of Economic Research.
Coleman, J., 1958. Relational analysis: the study of social organizations with survey methods. Human Organ. 17 (4), 28–36.
Currarini, S., Jackson, M.O., Pin, P., 2009. An economic model of friendship: homophily, minorities, and segregation. Econometrica 77 (4), 1003–1045.
Davies, B., 2022. Sex-based sorting among economists: Evidence from the nber.
Doan, T.P.A., Ductor, L., Rascon-Ramirez, E., 2023. The effect of networks on research output. Technical report, Mimeo.
Dong, M., Mayskaya, T., 2023. Does reducing communication barriers promote diversity? Available at SSRN 4164162.
Dong, M., Mayskaya, T., Smirnov, V., Taylor, O., Wait, A., 2023. Diversity in teams. Available at SSRN 4343715.
Ductor, L., 2015. Does co-authorship lead to higher academic productivity? Oxf. Bull. Econ. Stat. 77 (3), 385–407.
Ductor, L., Goyal, S., Prummer, A., 2018. Gender & collaboration. Technical report, Working Paper.
Ductor, L., Goyal, S., Prummer, A., 2021. Gender and collaboration. Rev. Econ. Stat., 1–40.
Ductor, L., Goyal, S., v. der Leij, M., Paez, G., 2020. On the influence of top journals. Unpublished.
Dupas, P., Modestino, A.S., Niederle, M., Wolfers, J., et al., 2021. Gender and the dynamics of economics seminars. Technical report. National Bureau of Economic Research.
Fafchamps, M., Goyal, S., van der Leij, M.J., 2010. Matching and network effects. J. Eur. Econ. Assoc. 8 (1), 203–231.
Gertsberg, M., 2022. The unintended consequences of # metoo-evidence from research collaborations. Available at SSRN 4105976.
Ginther, D.K., Kahn, S., 2004. Women in economics: moving up or falling off the academic career ladder? J. Econ. Perspect. 18 (3), 193–214.
Goyal, S., Van Der Leij, M.J., Moraga-González, J.L., 2006. Economics: an emerging small world. J. Polit. Econ. 114 (2), 403–412.
Graham, B.S., 2016. Homophily and transitivity in dynamic network formation. Technical report. National Bureau of Economic Research.
Hengel, E., 2016. Publishing while female. Technical report. University of Liverpool.
Hengel, E., Moon, E., 2020. Gender and equality at top economics journals. University of Liverpool. Paper 3111520. https://livrepository.liverpool.ac.uk/3111520/.
Hong, L., Page, S.E., 2001. Problem solving by heterogeneous agents. J. Econ. Theory 97 (1), 123–163.
Hughes, N., Khan, Z.U., Mengel, F., 2023. Diversity in committees. Available at SSRN 4495035.
Jackson, M.O., Nei, S.M., Snowberg, E., Yariv, L., 2023. The dynamics of networks and homophily. Technical report. National Bureau of Economic Research.
Kets, W., Sandroni, A., 2021. A theory of strategic uncertainty and cultural diversity. Rev. Econ. Stud. 88 (1), 287–333.
Koffi, M., 2021. Innovative ideas and gender inequality. Technical report, Working Paper Series.
Kwiek, M., Roszka, W., 2021. Gender disparities in international research collaboration: a study of 25,000 university professors. J. Econ. Surv. 35 (5), 1344–1380.
Laband, D.N., 2013. On the use and abuse of economics journal rankings. Econ. J. 123 (570), F223–F254.
Lazear, E.P., 1999. Culture and language. J. Polit. Econ. 107 (S6), S95–S126.
Lindenlaub, I., Prummer, A., 2021. Network structure and performance. Econ. J. 131 (2), 851–898.
Maddi, A., Gingras, Y., 2021. Gender diversity in research teams and citation impact in economics and management. J. Econ. Surv.
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annu. Rev. Sociol. 27 (1), 415–444.
Mengel, F., 2020. Gender differences in networking. Econ. J. 130 (630), 1842–1873.
Mengel, F., Sauermann, J., Zölitz, U., 2019. Gender bias in teaching evaluations. J. Eur. Econ. Assoc. 17 (2), 535–566.
Morgan, A.C., Way, S.F., Hoefer, M.J., Larremore, D.B., Galesic, M., Clauset, A., 2021. The unequal impact of parenthood in academia. Sci. Adv. 7 (9), eabd1996.
Page, S., 2008. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton University Press.
Palacios-Huerta, I., Volij, O., 2004. The measurement of intellectual influence. Econometrica 72 (3), 963–977.
Paredes, V.A., Paserman, M.D., Pino, F., 2020. Does economics make you sexist? Technical report. National Bureau of Economic Research.
Prat, A., 2002. Should a team be homogeneous? Eur. Econ. Rev. 46 (7), 1187–1207.
Sarsons, H., Gërxhani, K., Reuben, E., Schram, A., 2021. Gender differences in recognition for group work. J. Polit. Econ. 129 (1), 101–147.
Torres, L., Huffman, M.L., 2002. Social networks and job search outcomes among male and female professional, technical, and managerial workers. Sociol. Focus 35 (1), 25–42.
Wu, A.H., 2018. Gendered language on the economics job market rumors forum. In: AEA Papers and Proceedings, vol. 108, pp. 175–179.
Zeltzer, D., 2020. Gender homophily in referral networks: consequences for the medicare physician earnings gap. Am. Econ. J. Appl. Econ. 12 (2), 169–197.
Zhu, M., 2018. Job networks through college classmates: effects of referrals for men and women. Technical report, Working Paper.