Journal of Informetrics 16 (2022) 101295

Contents lists available at ScienceDirect

Journal of Informetrics

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

Further divided gender gaps in research productivity and collaboration during the COVID-19 pandemic: Evidence from coronavirus-related literature

Meijun Liu[a,b,c], Ning Zhang[d], Xiao Hu[e,k], Ajay Jaiswal[f], Jian Xu[g], Hong Chen[i], Ying Ding[f,j], Yi Bu[h,*]

[a] *Institute for Global Public Policy, Fudan University, Shanghai, China*
[b] *LSE-Fudan Research Centre for Global Public Policy, Fudan University, Shanghai, China*
[c] *Shanghai Center for Innovation and Governance, Fudan University, Shanghai, China*
[d] *School of Information Management, Nanjing University, Nanjing, Jiangsu, China*
[e] *Faculty of Education, University of Hong Kong, Hong Kong SAR, China*
[f] *School of Information, University of Texas at Austin, Austin, TX, United States of America*
[g] *School of Information Management, Sun Yat-sen University, Guangzhou, Guangdong, China*
[h] *Department of Information Management, Peking University, Beijing, China*
[i] *School of Information, University of Michigan, Ann Arbor, Michigan, State of, U.S.A*
[j] *Dell Medical School, University of Texas at Austin, Austin, TX, United States of America*
[k] *Musketeers Foundation Institute of Data Science, University of Hong Kong, Hong Kong SAR, China*

ARTICLE INFO

Keywords:
Gender inequality
Scientific collaboration
COVID-19
Scientific impacts
Mixed-gender collaboration

ABSTRACT

Based on publication data on coronavirus-related fields, this study applies a difference in differences approach to explore the evolution of gender inequalities before and during the COVID-19 pandemic by comparing the differences in the numbers and shares of authorships, leadership in publications, gender composition of collaboration, and scientific impacts. We find that, during the pandemic: (1) females' leadership in publications as the first author was negatively affected; (2) although both females and males published more papers relative to the pre-pandemic period, the gender gaps in the share of authorships have been strengthened due to the larger increase in males' authorships; (3) the share of publications by mixed-gender collaboration declined; (4) papers by teams in which females play a key role were less cited in the pre-pandemic period, and this citation disadvantage was exacerbated during the pandemic; and (5) gender inequalities regarding authorships and collaboration were enhanced in the initial stage of COVID-19, widened with the increasing severity of COVID-19, and returned to the pre-pandemic level in September 2020. This study shows that females' lower participation in teams as major contributors and less collaboration with their male colleagues also reflect their underrepresentation in science in the pandemic period. This investigation significantly deepens our understanding of how the pandemic influenced academia, based on which science policies and gender policy changes are proposed to mitigate the gender gaps.

* Corresponding author.
E-mail address: buyi@pku.edu.cn (Y. Bu).

https://doi.org/10.1016/j.joi.2022.101295
Received 8 July 2021; Received in revised form 12 April 2022; Accepted 3 May 2022
Available online 5 May 2022
1751-1577/

1. Introduction

The COVID-19 pandemic has disrupted our economy and society on a global scale. The pandemic-induced shock exerted impacts of different levels on various social groups and has hit the vulnerable groups the hardest, partly because of the social inequalities that already existed (Adams-Prassl, Boneva, Golin & Rauh, 2020; Brown & Ravallion, 2020; Johnson et al., 2020; Peters, 2021; Gomes & Avellaneda, 2021). Empirical evidence shows that the lockdown and other consequences caused by COVID-19 have exacerbated existing gender inequalities in various aspects, such as employment (Alstadsæter et al., 2020; Hossain, 2021), work conditions (Brubaker, 2020; Minello, 2020), psychological aspects (Liu, Zangerle, Hu, Melchiorre & Schedl, 2020) and health (Pieh, Budimir & Probst, 2020).

Changes originated by the COVID-19 pandemic forced scientists to work from home and amended their routines, especially those with young children, so that working and parenting became extremely challenging. The COVID-19 pandemic has amplified the unequal influence and further divided the gap between males and females (Chauhan, 2020b). Females have almost always assumed more responsibilities for parenting, and this was especially the case during this pandemic (Brubaker, 2020; Liang, Huang & Yang, 2020). City lockdowns and the closure of schools and/or daycare facilities have resulted in an enormous burden for female researchers, and their research productivity has plummeted (Minello, 2020).

Despite the prolific research about gender inequalities in terms of academic productivity, limited studies have provided a full picture of changes in scientific gender gaps during the COVID-19 pandemic, especially for gender inequalities concerning collaboration and scientific impacts. This study complements the extant studies on gender inequalities mainly from the perspective of productivity and collaboration including authorships, leadership in publications, gender composition of collaboration, and scientific impacts. Collaboration is one of the major influential factors of research productivity (Muriithi, Horner & Pemberton, 2013), helps address females' lack of social capital, and brings their better integration into the academic community (Abramo, D'Angelo & Murgia, 2013). However, it is unclear whether females' leadership and participation in teams changed during COVID-19. Hence, this study addresses this gap and sheds light on important mechanisms underlying the enlarged gender imbalance in science during COVID-19 by utilizing a difference in differences approach, and attempts to capture causality between the occurrence of COVID-19 and scientific gender gaps.

The objective of this study is to address currently underexplored research questions of whether, to what extent, and how gender inequalities increased with regard to the aggregate differences in the number and shares of males' and females' authorships, their leadership in publications, gender composition of collaboration, and scientific impacts during the COVID-19 pandemic. Some studies have claimed a narrowing trend of gender imbalance in publishing rates (Graham & Shaffer, 2004; Hart, Frangou & Perlis, 2019), while whether this trend has been disrupted during COVID-19 remains unclear. Most of the prior studies focused on gender gaps in authorship in COVID-19 period, and whether and how gender inequalities concerning collaboration and scientific impacts were enhanced remains to be further explored. To achieve the research objective, this study proposes three research questions:

Before and during COVID-19:
RQ1: How have females' and males' authorships, especially when they are major contributors in the papers, evolved?
RQ2: How has teams' gender composition developed?
RQ3: What is the association between teams' gender composition and scientific impacts of papers by teams?

2. Related work

2.1. Gender inequality in science

Gender gaps in the scientific community have been both theoretically and empirically investigated from various perspectives. Females' representation in academic careers has witnessed exponential growth since the 1970s (Cui, Ding & Zhu, 2021), while their progress has been much less dramatic in most mathematically intensive areas, such as physics, engineering, and mathematics. Since the large gender disparities in academic productivity were documented by Cole & Zuckerman (1984), researchers have been investigating the gender imbalance in science and the reasons for it, spanning a variety of forms including productivity, promotion, STEM education, and collaboration (Beaudry & Larivière, 2016; Derrick et al., 2019; Larivière, Ni, Gingras, Cronin & Sugimoto, 2013).

The gender disparities are pronounced in research productivity and scientific impacts. Cole & Zuckerman (1984) estimated that females published 57% as many papers as males and observed a large gender gap in research productivity, a phenomenon called the "productivity puzzle". Females' lower productivity relative to males is well documented in some early work (Fox & Faver, 1985; Long, 1992; Xie & Shauman, 1998), and has persisted in recent years (Luoto, 2020; Rachid et al., 2021). Prominent disciplinary differences were found regarding gender gaps of scientific productivity, with the largest gender gaps in engineering and technology, and natural sciences, and the closest gender gaps in medicine and agricultural sciences (Chan & Torgler, 2020; Ruggieri, Pecoraro & Luzi, 2021). Prior literature found that females are underrepresented when publishing in prestigious journals, and are less likely to be leading authors on a paper (Conley & Stadmark, 2012; West et al., 2013). A few studies have suggested the citation disadvantage of females (Beaudry & Larivière, 2016; Thelwall, 2018). For example, papers with females in leading author positions receive fewer citations than those with males in the same positions (Larivière et al., 2013).

In addition to underrepresentation in research performance, females are faced with barriers concerning collaboration. Evidence shows that females are more likely to publish single-authored papers, relative to males (Ferber & Teiman, 1980; McDowell & Smith, 1992; West et al., 2013). Studying a sample of 178 researchers with doctoral degrees from the top U.S. universities, McDowell

& Smith (1992) found that females were more likely than their male peers to co-author with another female. Females' propensity to collaborate with their peers of the same gender might lead to their lack of social capital because of the historic overrepresentation of men in science (Abramo et al., 2013). Prior literature has discussed the reasons for females' propensity for single authorship, spanning from gender sorting in academia and gender disparities in the propensity for collaboration, to gender segregation in the choice of research sub-fields (Boschini & Sjögren, 2007; McDowell, Singell Jr & Ziliak, 2001; Thelwall, Bailey, Tobin & Bradshaw, 2019). One experiment indicated that a team's productivity is affected by its members' genders as gender composition could influence team members' willingness to contribute (Ivanova-Stenzel & Kübler, 2011). That study further found that females underperform in gender-mixed teams and perform best in purely female teams .

The mechanisms that drive the gender gap in academia have been investigated primarily from the individual- and institutional-level perspectives. Numerous researchers have explored the reasons for the underrepresentation of females in science, especially in STEM, while the conclusions are still controversial. Hypotheses vary from biological reasons (Eals & Silverman, 1994) to sociocultural factors (Ceci, Williams & Barnett, 2009; Halpern et al., 2007). Discrimination against females in hiring, journal reviewing, and funding allocation is considered a powerful reason that accounts for their underrepresentation in science (Ceci & Williams, 2010; Hill, Corbett & Rose, 2010). One experiment showed that discrimination is important in explaining females' underrepresentation in science by asking 127 faculty members to evaluate applications for the position of lab manager; it was observed that identical applications were more highly evaluated if the applicant had a male name (Moss-Racusin, Dovidio, Brescoll, Graham & Handelsman, 2012). Yet, only trivial or nonexistent biological gender differences are found in inherent aptitude and abilities for math and science (Hyde & Linn, 2006; Spelke, 2005).

Prior literature documented that scientific gender disparities might be attributed to females' interest in careers involving social relations, their decision to raise children, taking care of elderly parents, or looking for a job within a limited geographical area for family reasons (Bell, 2010; Ceci & Williams, 2010; Su, Rounds & Armstrong, 2009). Females are responsible for the major part of child- and family-care (Beddoes & Pawley, 2014), which limits their engagement in academic activities and thus dampens their productivity and impacts. This fact is considered the primary reason for gender differences in science (Ceci & Williams, 2011; Moss-Racusin et al., 2012). Based on a survey of 17,519 first and last authors publishing between 2007 and 2017, Derrick et al. (2019) suggested a connection between parenting arrangements and differing research productivity for males and females, and demonstrated that the amount of parental responsibility is a key factor that explains gender differences in academic productivity. Because of parental responsibilities, females spend less time working at their jobs than both their male colleagues and female peers with no children, which might reduce their productivity and lower their likelihood of earning tenure positions (Ceci & Williams, 2010).

2.2. Gender gaps during the COVID-19 period

Gender gaps have already been extensively discussed in the COVID-19 pandemic period. Pinho-Gomes et al. (2020) noted that females' lower representation as authors of COVID-19-related research papers published in biomedical science journals persisted in the early stages of the outbreak. In STEM and medical fields, Krukowski, Jagsi and Cardel (2021) observed that the significant gender imbalance in academic productivity was amplified during COVID-19, with no change in men's productivity and a significant decrease in females' self-reported article submissions. Fewer females submitted manuscripts to medical-related professional journals during the pandemic (Kibbe, 2020), and their research productivity has been influenced more than that of their male peers (Andersen, Nielsen, Simone, Lewiss & Jagsi, 2020). The widening gender gaps in research productivity or journal submissions were also found in many other bio-medical research fields (Ribarovska, Hutchinson, Pittman, Pariante & Spencer, 2021; Wehner, Li & Nead, 2020), such as ophthalmology (Nguyen, Trinh, Kurian & Wu, 2021) and public health (Bell & Fong, 2021). The above studies showed a more severe impact of COVID-19 on female academics who are engaged in collaborative research, although the falling curves are not the same due to methodological differences.

The major causes of scientific gender inequalities during COVID-19 include city lockdown, social-distancing policies, and gendered roles and social norms in domestic and caring work. These factors disproportionately increased the time females spent on housework and childcare, and reduced the time they spent on research (Adams-Prassl et al., 2020; Minello, 2020). Collins, Landivar, Ruppanner & Scarborough (2020) analyzed the U.S. Current Population Survey panel data spanning February through April 2020, and found that the gender difference in working hours increased by 20–50% during the COVID-19 pandemic. A survey of U.S. and European female scientists by Myers et al. (2020) indicated that female scientists with young children spent significantly less time conducting research during the pandemic than they did in the pre-COVID-19 period. In addition to childcare, housework has been found to be a factor increasing gender inequality during the COVID-19 pandemic (Farré, Fawaz, González & Graves, 2020). Amano-Patiño, Faraglia, Giannitsarou & Hasna (2020) speculated that, during the COVID-19 pandemic, female researchers in their early and mid-careers may have been less willing to take risks, and were therefore less likely to start new projects during that time, which may have prevented them from investing in new and high-cost research. A more recent survey showed that females experienced a substantial decline in starting new projects during COVID-19 (Gao, Yin, Myers, Lakhani & Wang, 2021).

The pandemic-related impacts are highly unequal and disproportionately detrimental to females, and the effects at this time might damage their careers going forward. Vincent-Lamarre, Sugimoto & Larivière (2020) explicitly pointed out that females are less likely to be leading authors, which may negatively affect them in their early stages. The decline in the number of projects registered by female scholars compared to previous years, according to some registry report databases, could have a long-term adverse effect on their scientific research (Viglione, 2020). Given the ongoing global pandemic and the continued lockdown measures, it is undoubtedly necessary and urgent to explore policies to promote gender equality and to avoid the worsening of such gender imbalances. For example, Witteman, Haverfield & Tannenbaum (2021) studied grant application policies and found that factoring gender into COVID-19 grant requirements would increase the number of applications from female scientists.

Existing literature provides evidence to show the differential impact of COVID-19 on females and males in academia, especially for productivity. Gender imbalance was generally found to be strengthened, but whether the inequality differed at different times of the outbreak has not been well explored. On the other hand, most of the studies are descriptive analyses that look at changes in numbers and proportions, and lack inferential investigation with rigorous modeling. In addition, other dimensions, such as the impact on gender-specific scientific impacts and collaboration, remain missing. To address these research gaps, this study applies a causal inference approach, i.e., a difference in differences strategy, to investigate the changes in gender inequalities in the coronavirus-related domain before and during COVID-19, regarding the differences in the numbers and shares of females' and males' authorships, leadership in publications, gender composition of collaboration and scientific impacts.

3. Data and methodology

3.1. Data source

3.1.1. Publication data

Records on coronavirus-related papers were downloaded from the COVID-19 Open Research Dataset (hereafter CORD).[1] This dataset contains 332,458 research articles about COVID-19 and related historical coronaviruses, such as SARS and MERS, that were published between 1951 and October 2020. This dataset includes title, abstract, author name, PubMed ID, publication date and so forth. Papers in this dataset are sourced from PubMed Central, bioRxiv, and medRxiv, with titles, abstracts, or full text including the following keywords:

"COVID-19″ OR "Coronavirus" OR "Corona virus" OR "2019-nCoV" OR "SARS CoV" OR "MERS-CoV" OR "Severe Acute Respiratory Syndrome" OR "Middle East Respiratory Syndrome". CORD includes 122,310 papers between January 2018 and September 2020. Based on pmid/s2_id recorded in CORD, we obtain citations of these papers from Semantic Scholar (Ammar et al., 2018).

3.1.2. Patient data

We use the patient data on COVID-19 derived from the website of Our World in Data that covers 211 countries from December 2019 to September 2020,[2] to capture the timing when the first COVID-19 case was reported, and the daily number of new COVID-19 cases in each country during the study period.

3.2. Predicting authors' gender and identifying country information in CORD

Advanced inferential analytics methods have been recently developed to predict genders (Das & Paik, 2021; Fabris, Purpura, Silvello & Susto, 2020; Fosch-Villaronga, Poulsen, Søraa & Custers, 2021). In this study, Gender Guesser, a Python Package, is used to predict gender information of authors based on their first names. In scientific research, inferring gender by names with quantitative methods is one of the mainstream solutions. In this study, we use Gender Guesser, which is considered reliable and has been extensively used in related work (Feramisco et al., 2009; Squazzoni et al., 2021), to predict authors' gender. The gender prediction of this tool is presumed to be of high quality because it involves manual checks by native speakers of various countries (Santamaría & Mihaljević, 2018). Gender Guesser maps first names to the genders those names are associated with by using a database including 45,000 names and classifying each name as "male", "female", "mostly male", "mostly female", "androgynous", and "unknown". We only keep authors' names that are categorized as "female" or "male" for higher accuracy. Papers with all authors with genders not identified are removed, with 98,742 papers kept. Figure A1 in Supplementary Information (SI) presents the results of gender identification based on the coronavirus-related literature using Gender Guesser.

We aim to conduct country-based regression analyses. To this end, CORD papers are linked to their PubMed versions to obtain the data on authors' address information, by which their country information is captured. To standardize country names, we manually merge variations of country names, including ISO two- or three-letter country codes, alternative country names, and country names with typos, into the same country. Finally, standard country names corresponding to authors' affiliations in 36,447 CORD papers are found. We keep papers published between January 2018 and September 2020, and published by authors from the 50 most productive countries ranked by the number of coronavirus-related papers in the CORD dataset,[3] with 33,104 CORD papers and 16,429 unique authors remaining.

[1] Accessible at https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge. This dataset was downloaded on October 8, 2020. Publications in October 2020 are removed due to the incomplete coverage of this month.

[2] Accessible at https://ourworldindata.org/covid-cases. This dataset was downloaded on October 9, 2020.

[3] The dependent variables for countries that did not publish or published very few coronavirus papers would be missing. The observations that have a missing value for any one of the variables used in the regression model would be dropped by Stata. This is the reason why we limit our analyses to the most prolific countries in the CORD dataset. To measure a country's productivity, we use a full counting method based on the authors' address information. For example, for a paper authored by two scientists with Chinese affiliations, one scientist with a U.S. affiliation and three scientists with U.K. affiliations, China, the U.S., and the U.K. get two, one, and three papers, respectively. For scientists with several affiliations belonging to different countries, we consider the country information of the affiliation listed as the first one.

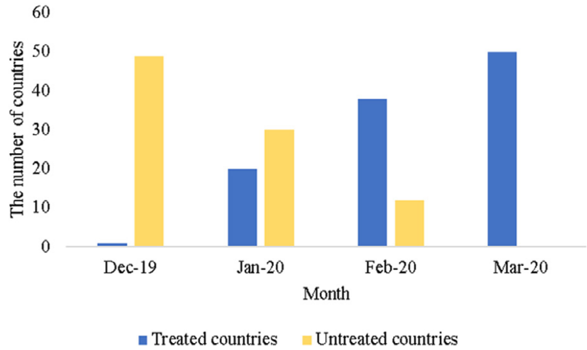


Fig. 1. The distribution of treated countries and untreated countries by month. In and after March 2020, all sampled countries were exposed to COVID-19 and thus were treated countries. The Y-axis indicates the number of countries. The X-axis means the month in which the first COVID-19 case has been confirmed in a country.

3.3. Variables

3.3.1. Independent variables

The binary variable, whether the first COVID-19 case (*COVID19*) has been confirmed in the country by the month, is the major explanatory variable in this study. The month when the first case of COVID-19 in the 50 sampled countries occurred is identified based on the patient data (Table A1 in SI). Once the first case of COVID-19 has been officially reported, the country gets treated in the month and the succeeding months. The treatment status of each country varies with time. From January 2018 to November 2019, since there was no COVID-19 case on a global scale, all the sampled countries were untreated. In January 2020, 19 countries became treated countries as the first COVID-19 cases were confirmed in those countries in that month (see Table A1 in SI). The distribution of treated countries (i.e., the countries where the first COVID-19 case has been confirmed) and untreated countries (i.e., the countries where the first COVID-19 case has not been confirmed) by the month is indicated in Fig. 1. Note that all sampled countries have been exposed to COVID-19 by March 2020.

3.3.2. Dependent variables

From four dimensions, i.e., the numbers and shares of authorships, leadership in publications, gender composition of collaboration, and scientific impacts, we explore the gender inequalities before and during the COVID-19 pandemic.

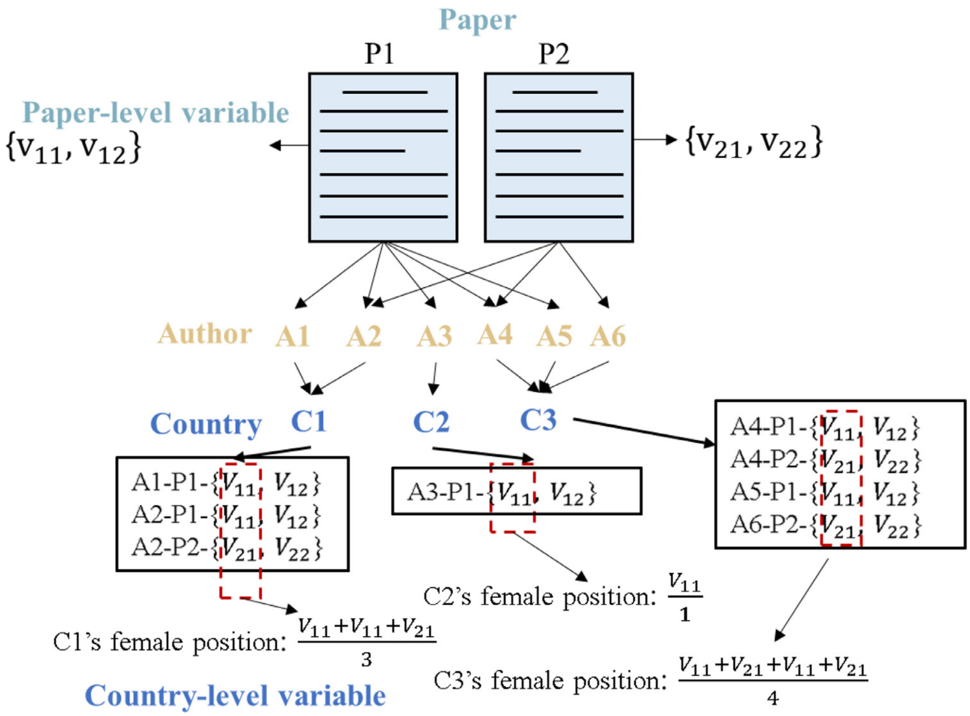
The numbers and shares of authorships: To measure females' and males' monthly authorships in each country, we use a full counting method (Waltman, 2016) based on authors' address information. For example, for a paper authored by one male scientist in a Chinese affiliation and two female scientists in a U.S. affiliation, China and the U.S. get zero and two papers regarding females' authorships, respectively. Based on the papers' publication dates, we generate the number of authorships by females (*female all*) and by males (*male all*) in each country in each month. The proportion of females' authorships to a country's total authorships, *female all(share)*, in each month, and that of males' authorships, *male all(share)*, to a country's total authorship count in each month could thus be obtained. It is noted that the sum of a country's *female all(share)* and its *male all(share)* in a month equals one.

Leadership: To investigate the changes in females' and males' leadership in coronavirus-related papers before and during the pandemic, we differentiate females'/males' publications in which they are leading authors (defined as the first and/or the last author) from the pool of their publications. We quantify the fraction of publications with a female as the first author, *female first(share)*, and that with a female as the last author, *female last(share)*, to the total publications of a country in each month. In addition to the leading authors, females' importance as non-leading authors is considered, which is defined in Eq. 1:

$$Female\ position = \frac{\sum S_j}{(N-2)\times F} \tag{1}$$

where N indicates the number of authors in paper i; F shows the number of female authors in addition to the first and the last authors in paper i; j indicates a female author in non-leading author position in paper i; S refers to the author sequence of a female in a paper. $Female\ position$ operationalizes the average importance of females in non-leading author positions in a paper. For example, for a paper with seven authors, two females and five males, one female is the third author and the other is the fourth author. $Female\ position$ for this paper equals $(3/5 + 4/5)/2 = 0.7$. The smaller the value of this indicator, the more important females are as non-leading authors in a paper. If there is no female as non-leading authors in a paper, the female position of this paper is set as "undefined" since the denominator of Eq. (1) cannot have the value of zero. Male position is calculated similarly.

Gender composition of collaboration: Scientific teams in coronavirus-related fields are categorized into three groups based on gender composition, i.e., all-female, all-male, and mixed-gender teams. The number of publications for these three groups by a country in a month could be calculated separately, as well as the fraction of publications by these three groups to the total publications (*all-female*, *all-male*, and *mixed-gender*) by the country in a month.



	Female position	Team size
C1	(V11+V11+V21)/3	(V12+V12+V22)/3
C2	V11	V12
C3	(V11+V21+V11+V21)/4	(V12+V22+V12+V22)/4

Fig. 2. The way to aggregate paper-level variables to country-level variables.

Scientific impact: We explore the association between teams' gender composition and papers' citations before and during the pandemic. Papers' citations are used to approximate their scientific impacts (Liu & Hu, 2021; Liu, Shi & Li, 2017; Min, Bu, Wu, Ding & Zhang, 2021). We use a natural log of citation count of each paper, ln (citation_count+1).

There are two paper-level variables that need to be aggregated to the country level for further country-level analyses, i.e., *female position* and *team size* (i.e., the control variable included in the analysis). The way to aggregate the paper-level variable to the country level is annotated in Fig. 2. *Female position* and *team size* for paper *i* are denoted by v_{i1} and v_{i2}, respectively. There are two papers, P1 by five authors from three countries (i.e., C1, C2, and C3), and P2 by three authors from two countries (C1 and C3), respectively. A1, who is an author of P1, and A2, who is in the author lists for P1 and P2, both belong to country C1. v_{11} and v_{21} indicate female position for P1 and P2, respectively. C1's average female position is the sum of v_{11}, v_{11}, and v_{21} weighted by the unique number of author-paper pairs (i.e., A1-P1, A2-P1, and A2-P2), i.e., three. Similarly, C3's female position is equal to the sum of v_{11}, v_{21}, v_{11}, and v_{21} weighted by four author-paper pairs (i.e., A4-P1, A4-P2, A5-P1, and A6-P2). In a similar way, we transform paper-level team size to the country level.

3.4. Difference in differences approaches

The unexpected outbreak of the COVID-19 pandemic is used as a quasi-experiment to investigate gender inequalities. We apply a difference in differences (DID) approach based on the data on 50 sampled countries over 33 months from 2018 January to 2020 September. DID is applied to mimic a natural experimental research design and infer potential causality based on observational study data (Angrist & Pischke, 2008; Liu et al., 2021).

The major goal of this study is to estimate the relationship between countries' monthly changes in gender inequalities in the four aspects, and the occurrence of the first COVID-19 case in the country. To address RQ1 and RQ2, we regress the major dependent variables, as aforementioned, on whether or not the first COVID-19 case in the country has been officially confirmed by the month, as shown in Eq. (2). To control the possible impact of the changes in team size on the dependent variables, team size is added as a control. An Ordinary Least Square linear model is applied to estimate Eq.(2). The fixed effects for country, θ_i, and those for month, δ_t, are incorporated to control the time- and country-invariant factors. The coefficient on COVID19 is a before-after estimate of the

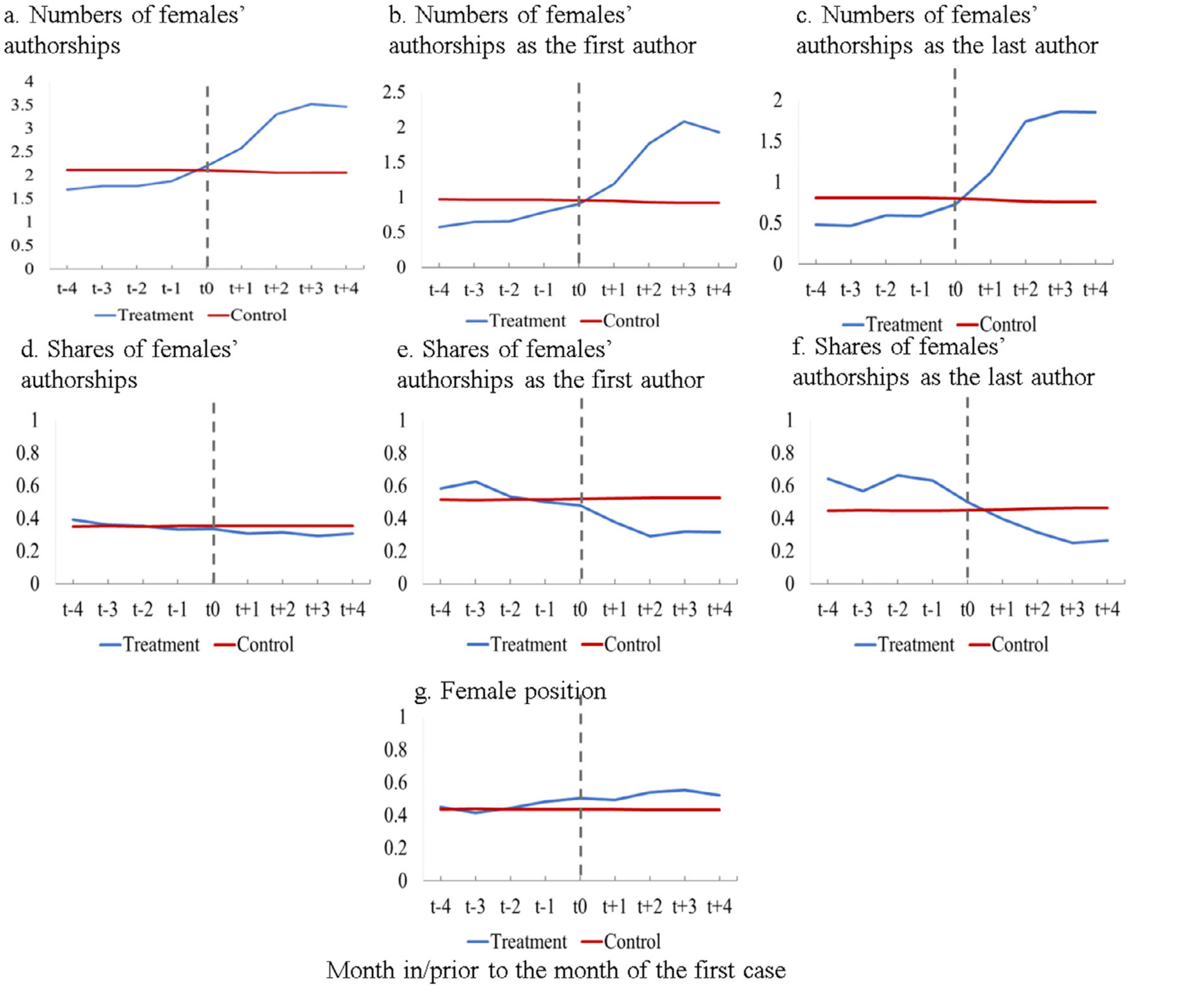


Fig. 3. The evolution of the dependent variables with regard to females' authorships of the treatment groups and control groups. The Y-axis indicates the mean of the dependent variables; t-n/t+n indicate n months prior to/after the month of the first COVID-19 case, and t0 refers to the month of the first COVID-19 case; the dashed gray lines indicate t0.

effect of COVID-19 on gender inequalities in the four dimensions.

$$Dependent\ variables_{i,t} = \alpha + \beta COVID19_{i,t} + \gamma Control_{i,t} + \delta_t + \theta_i + \varepsilon \quad (2)$$

To examine the association between the severity of COVID-19 in the country and gender inequalities in the four aspects, we regress the dependent variables on the monthly number of new COVID-19 cases. Fixed effects of country and month are added as well.

To address RQ3, we perform a paper-level analysis to explore gender inequalities in scientific impacts in the pre-COVID-19 period and during the pandemic. The independent variables are whether a female is the first author in a paper, whether a female is the last author in a paper, and female position. The dependent variable is the natural log-transformation of a paper's citations.

3.5. Tests of the parallel trend assumption

DID has an assumption that changes or growth rates in dependent variables of treatment groups and those of control groups are similar, which is also known as the parallel trend assumption. In this study, this assumption is satisfied if and only if the dependent variables of the two groups have parallel trends before the occurrence of the first COVID-19 case in a country. In other words, the evolution of the difference in the dependent variables between the two groups should be "flat" over time before the intervention. The most ideal situation is that the differences in dependent variables between the two groups are close to zero and insignificant before the treatment (Roth, 2019)

First, following previous literature (Li, Ding, & Yang, 2020; Stagni, Fosfuri, & Santaló, 2021), we plot the mean of the dependent variables of treatment groups and control groups over periods (Figs. 3-5). In general, before t0, the month of the first COVID-19 case, the evolution of the dependent variables shows a generally parallel trend for treatment groups and control groups, and the

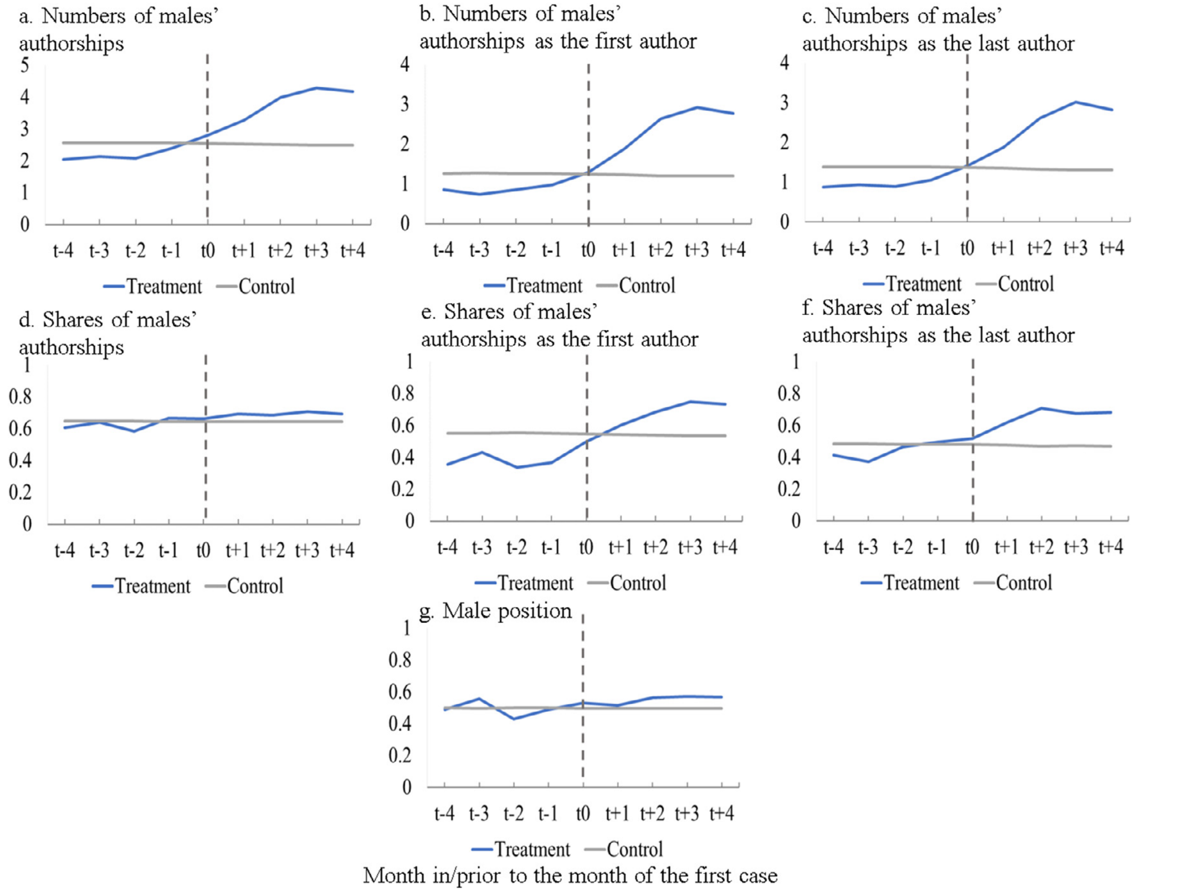


Fig. 4. The evolution of the dependent variables with regard to males' authorships of the treatment groups and control groups. The Y-axis indicates the mean of the dependent variables; t-n/t+n indicate n months prior to/after the month of the first COVID-19 case, and t0 refers to the month of the first COVID-19 case; the dashed gray lines indicate t0.

differences in the dependent variables between the two groups are small and constant before t0. However, after t0, the differences in the dependent variables between the two groups turn prominent and become increasingly larger, relative to the pre-treatment periods.

Then, we apply statistical tests to perform the pretreatment tests. Following prior literature (Agrawal, 2013; Wang, Yin & Yu, 2021), we replace the variable COVID-19 in Eq. (2) with nine dummy variables that relate the dependent variables to the occurrence of the first COVID-19 case in the country in, prior to and after the month when the first COVID-19 case was confirmed (see Eq. (3)). The coefficients of the dummy variables we introduce, i.e., t-n and t0, indicate the difference in the dependent variables between treated and untreated countries before and in the month of the first case. If the coefficients are small and even close to zero, and insignificant, the differences in the dependent variables between the two groups are slight and statistically insignificant, which indicates a decent satisfaction of the parallel trend assumption.

$$\begin{aligned} \textit{Dependent variables}_{i,t} &= \alpha + \beta_1(t-4)_{i,t} + \beta_2(t-3)_{i,t} + \beta_3(t-2)_{i,t} + \beta_4(t-1)_{i,t} + \beta_5(t0)_{i,t} \\ &+\beta_6(t+1)_{i,t} + \beta_7(t+2)_{i,t} + \beta_8(t+3)_{i,t} + \beta_9(t+4)_{i,t} + \gamma \textit{Control}_{i,t} + \delta_t + \theta_i + \varepsilon \end{aligned} \tag{3}$$

Where t0 is a binary variable that relates to the month when the first COVID-19 case was confirmed in the country; t-n and t+n indicate whether the observation occurs n month(s) before or after the month of the first case, respectively. The coefficients of t0 and t-n indicate the difference in dependent variables between the two groups in and before the occurrence of the first COVID-19 case.

As shown in Figs. 6-8, the difference between the two groups is insignificant for most of the time in and prior to the month of the first COVID-19 case, which is evidenced by the insignificant coefficients of t0 and t-n. Consequently, the parallel trend assumption is generally satisfied.

3.6. The structural changes of authors in CORD papers

The outbreak of the pandemic led to a surge in coronavirus-related literature and attracted authors from diverse backgrounds who had not investigated coronavirus prior to the pandemic to publish papers in the field. If those new entrants exhibit structural

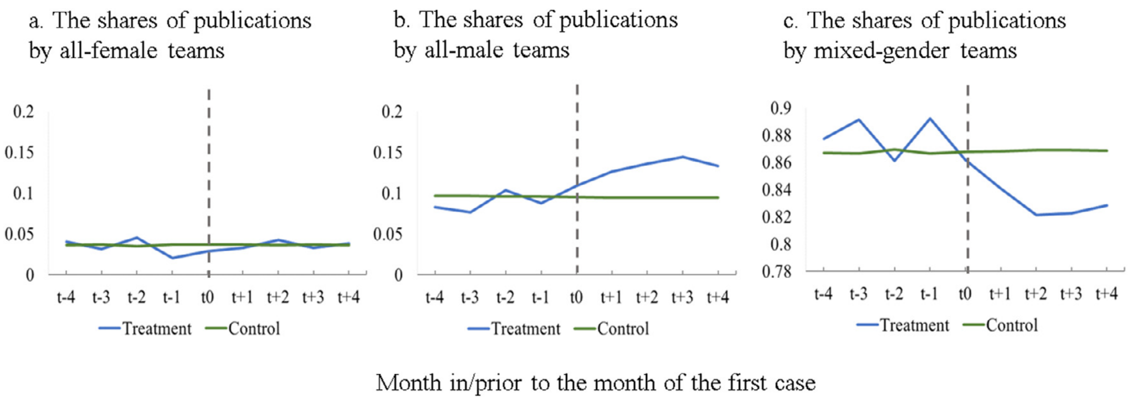


Fig. 5. The evolution of the dependent variables with regard to gender composition of collaboration of the treatment groups and control groups. The Y-axis indicates the mean of the dependent variables; t-n/t + n indicate n months prior to/after the month of the first COVID-19 case, and t0 refers to the month of the first COVID-19 case; the dashed gray lines indicate t0.

Table 1
The comparisons of the three groups of authors in CORD papers.

Group	# of obs.	% of female	% of authors publishing at least one PubMed article before December 2019
1. Authors (2018.01–2019.11)	7010	35.33%	97.83%
2. Authors (2019.12–2020.09)	10,374	31.77%	91.99%
3. New Entrants	9419	31.92%	91.23%

differences relative to incumbents, the changes in the outcome variables concerning gender inequalities we observed might be caused by those structural changes, rather than the true impact of COVID-19.

To mitigate the above concerns, from three dimensions, i.e., gender, country, and whether or not the author published at least one PubMed article before December 2019, we compare the following groups of authors:

- Group 1: authors who published at least one CORD paper before the pandemic
- (January 2018 to November 2019);
- Group 2: authors who published at least one CORD paper during the pandemic
- (December 2019 to September 2020); and
- Group 3: new entrants, defined as authors who published their first CORD paper during the pandemic.

The comparisons across the three groups could illustrate whether or not structural changes of authors in CORD papers occurred. Whether or not the author published at least one PubMed article (i.e., papers indexed in PubMed but which are not CORD papers) implies their research subjects. If an author published articles indexed in PubMed before the outbreak of COVID-19, it is plausible that they worked in biomedical or life sciences domains before the pandemic.

The results of the comparisons are shown in Table 1 and Fig. 9. Among 16,429 authors in CORD papers, 57.33% of authors are new entrants who had never published in the coronavirus-related domains before the outbreak of the pandemic and entered these fields during the pandemic. Generally, despite some slight differences, we do not observe substantial structural changes in gender compositions, research subjects, or authors' country across the three groups. The above evidence suggests that the divided gender gaps we show in Section 4 cannot be simply attributed to the structural changes of researchers in CORD papers.

4. Results

Since the first global COVID-19 case, the number of publications in coronavirus-related topics by females and that by males have both considerably increased. Fig. 10(a) and (b) indicate a dramatic leap in females' and males' authorships since January 2020. The growth of authorships by males is sharper than that by females. The fraction of coronavirus-related publications by teams with a female first author sharply decreased after the occurrence of the first global COVID-19 case, as shown in Fig. 10(c). It is the same for the fraction of publications by teams with a female as the last author. In contrast, the proportion of publications by teams with a male as the first or the last author suddenly increased after December 2019 (Fig.10(d)).

The results of the regression analyses consistently show that females and males both witnessed a sharp growth in publication counts during COVID-19, with a 24.5% increase for females (Fig. 11(a)) and a 53.7% one for males (Fig. 11(b)). However, the proportion of females with regard to the total authorships decreased by 5% (Fig. 11(a)). Those results suggest that the relative authorships of females have declined though females' absolute authorships have grown. Publications by teams with a female as the first author decreased by 21% during COVID-19 (Fig. 11(a)). There is no significant change in the publications by teams with a female as the last author, or in the position of females in teams (coefficient: 0.05,p-value<0.1) (see Fig. 11(a)). The regression results of the relationship

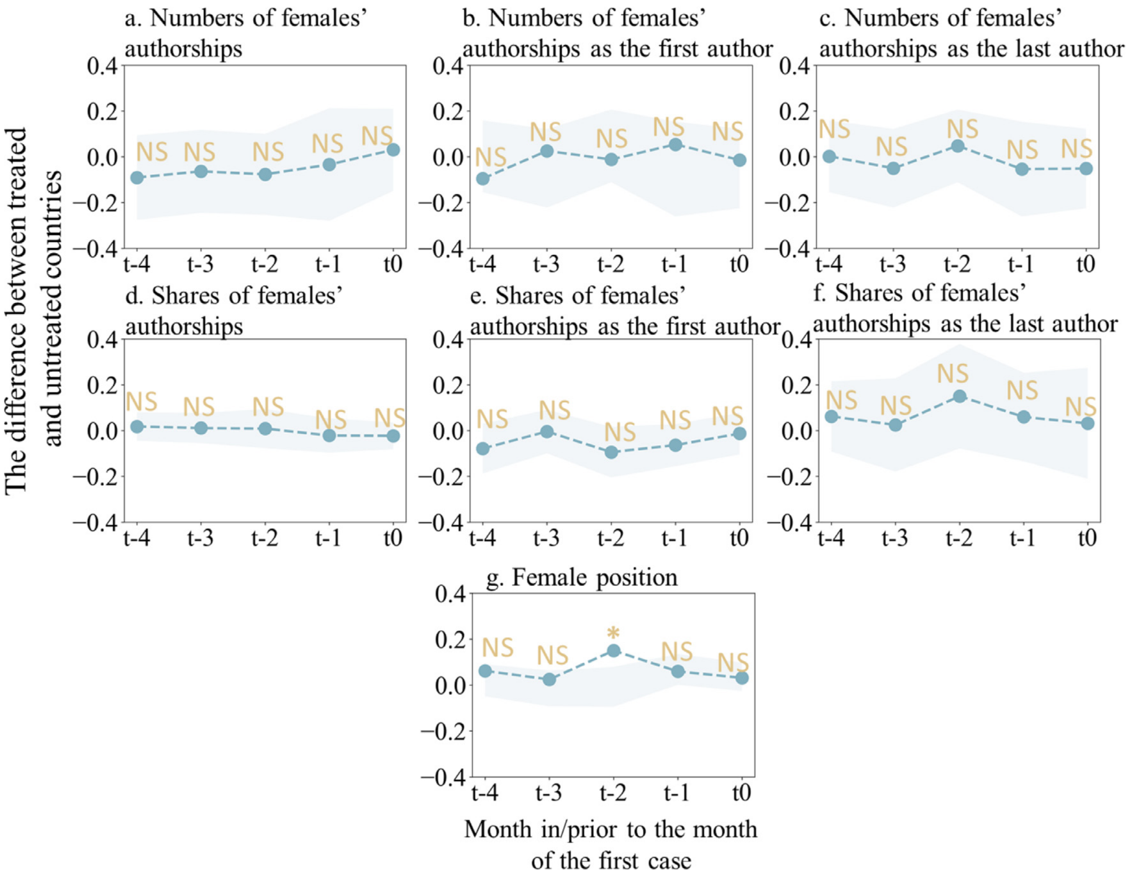


Fig. 6. The difference between treatment groups and control groups in dependent variables regarding females' authorships in and before the month of the first COVID-19 case. The shaded area represents ± 1.96 * std. error of each point estimate; t-n indicate n months prior to the month of the first COVID-19 case, and t0 refers to the month of the first COVID-19 case; the coefficient means the difference between treated countries and untreated countries in dependent variables; ***, **, and * represent significance at the 1, 5, and 10% levels and NS stands for not significant.

between the occurrence of the first COVID-19 case in the country and females'/males' monthly authorships and importance in teams are shown in Tables A2 and A3 in SI.

We further find that the share of publications by mixed-gender teams was hampered during COVID-19, while the share of publications produced by all-male teams went up. From December 2019, the fraction of papers authored by mixed-gender teams suddenly decreased at the initial stage of the pandemic, then gradually went up, and remained at a level similar to where it was before COVID-19 (Fig. 10(f)). In contrast, the proportion of papers by all-male teams increased, and reached a peak in March 2020 (Fig. 10(f)). The regression analyses provide consistent results. The share of publications by all-female and all-male teams increased by 1.8% and 5% (see Fig. 11(c)), respectively. However, countries' share of publications by mixed-gender teams have declined by 6.8% since the first COVID-19 case was officially confirmed. The regression results of the relationship between the occurrence of the first case of COVID-19 in the country and the share of publications by the three types of teams are presented in Table A4 in SI.

Gender inequalities regarding the differences in the share of authorships and collaboration widened immediately after the first global COVID-19 case was confirmed, while this imbalance was gradually mitigated and evolved to the level of the pre-pandemic period with the development of COVID-19. Fig. 10(c) illustrates a sudden decrease in the share of authorships of females as the first and as the last author after December 2019, and reached the bottom in March 2020. After this month, the share of authorships of females as the leading authors gradually grew and was nearly restored to the pre-pandemic level by the end of the study period. In contrast, the share of authorships of males as the leading authors increased sharply at the initial stage of the pandemic, and reached the peak in March 2020, and finally returned to the pre-pandemic state (Fig. 10(d)). We also observe a similar trend for the share of publications by single-gender and mixed-gender collaborations (Fig. 10(f)). All the evidence shows that gender imbalance in terms of the differences in the shares of authorships and collaboration were suddenly strengthened after the COVID-19 pandemic happened, and then gradually went back to the pre-pandemic level.

Gender inequalities seem to be strengthened with the increasing severity of COVID-19 in a country. The authorships of both females and males increased with the growth of new COVID-19 cases in a country, while the magnitude of the increase for females is smaller than that for males. A 1% increase in new cases is related to a 0.036% increase (see Column 1 of Table A5 in SI) in females'

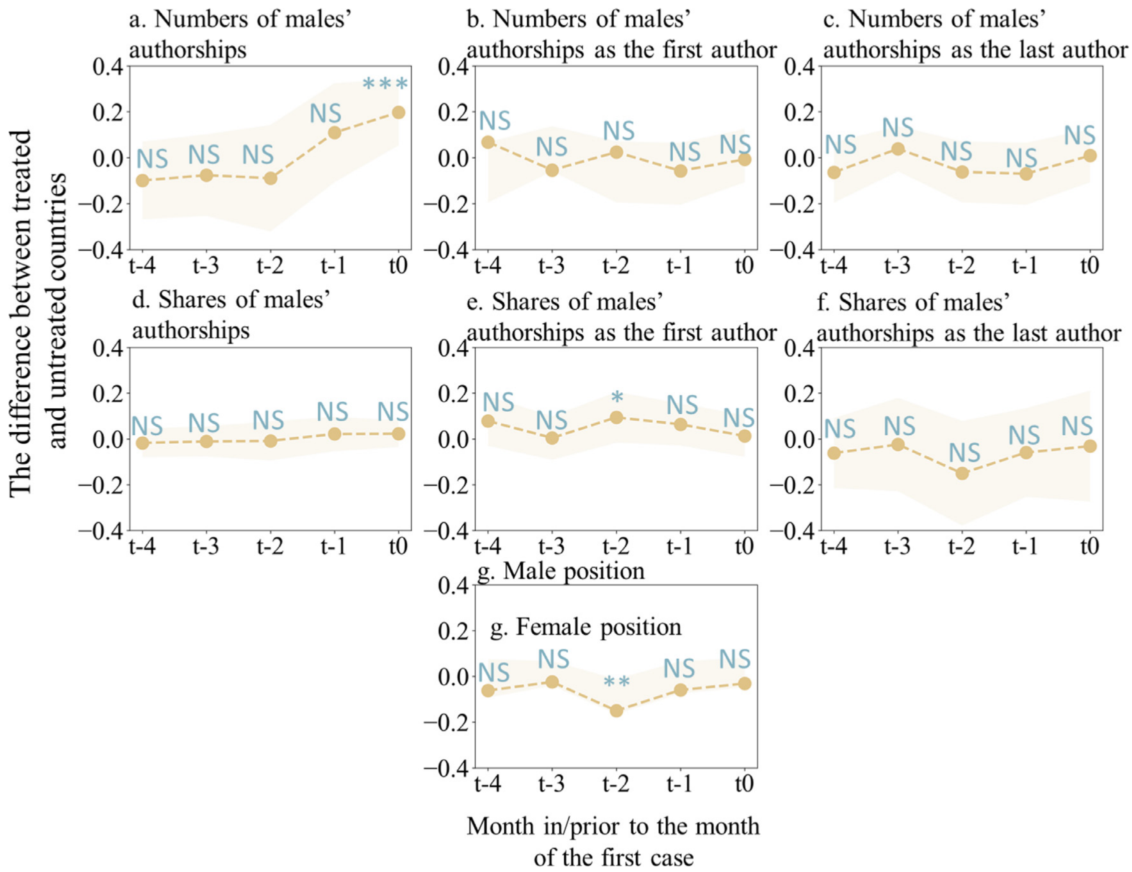


Fig. 7. The difference between treatment groups and control groups in dependent variables regarding males' authorships in and before the month of the first COVID-19 case. The shaded area represents ± 1.96 * std. error of each point estimate; t-n indicate n months prior to the month of the first COVID-19 case, and t0 refers to the month of the first COVID-19 case; the coefficient means the difference between treated countries and untreated countries in dependent variables; ***, **, and * represent significance at the 1, 5, and 10% levels and NS stands for not significant.

authorships and a 0.065% growth (Column 1 of Table A6 in SI) of males' authorships. With new cases going up, the fraction of papers in which females are listed as the first author decreased (Column 3 of Table A5 in SI). Additionally, the proportion of papers by single-gender teams grew (Columns 1 and 2 of Table A7), while that for mixed-gender papers shrank with the growth of new cases (Column 3 of Table A7).

During the COVID-19 pandemic, as compared to publications by teams with males as the first author, publications in which the first author is female received 8% lower citations (see Fig. 12(b) and Column 5 of Table A8). However, papers' citations and whether or not the first author is female were not statistically correlated before the COVID-19 pandemic (Fig. 12(a)). Regardless of before or during the COVID-19 pandemic, papers by teams with a female as the last author are cited less than those by teams with a male as the last author (Fig. 12(b) and columns 3 and 6 of Table A8). There is no significant relationship between the proportion of females in a team and citations papers received before and during the COVID-19 pandemic (Fig. 12 and columns 1 and 4 of Table A8). Those results generally suggest that papers in which females play a key role are less cited, and this form of inequality was strengthened during COVID-19.

To ensure the robustness of the findings, we perform robustness checks using various strategies. (1) We perform author-level analyses (Tables A9 and A10 in SI); (2) explore whether the findings hold for non-coronavirus-related domains by investigating the evolution of gender inequalities during COVID-19 in the field of cancer (Tables A11 to A13 in SI); (3) use Genderize.io, a widely used gender prediction tool, to identify authors' gender information (Tables A14 to A16); and (4) apply an inverse-probability-weighting estimator (Table A17 in SI). In general, the major findings hold, while we find that gender inequalities concerning the various aspects aforementioned did not change during COVID-19 in the field of cancer.

5. Discussion and conclusion

The outbreak of the COVID-19 pandemic has spurred studies of its impact on gender inequalities. However, how gender inequalities in science differ before and during a pandemic remains unclear. In coronavirus-related domains, we find that gender gaps widened

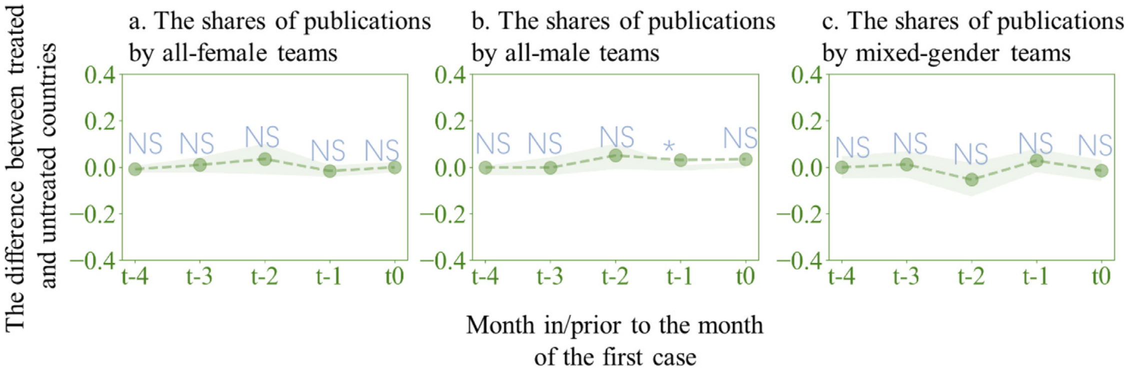


Fig. 8. The difference between treatment groups and control groups in dependent variables regarding the share of publications by the three types of teams in and before the month of the first COVID-19 case. The shaded area represents ± 1.96 * std. error of each point estimate; t-n indicate n months prior to the month of the first COVID-19 case, and t0 refers to the month of the first COVID-19 case; the coefficient means the difference between treated countries and untreated countries in dependent variables; ***, **, and * represent significance at the 1, 5, and 10% levels and NS stands for not significant.

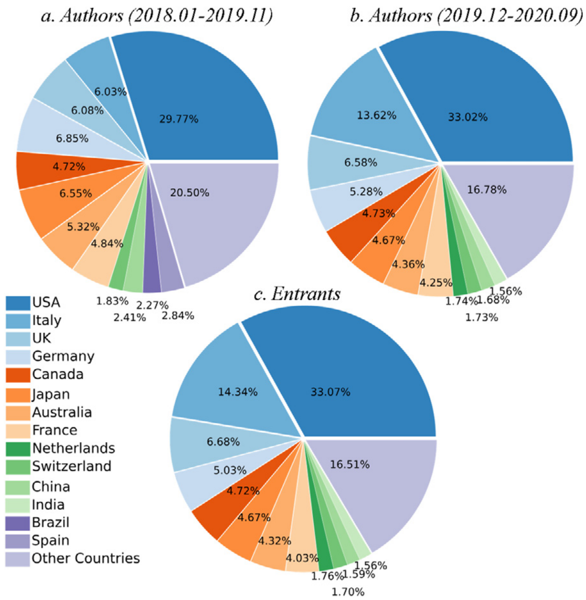


Fig. 9. The top 12 countries the three groups of CORD authors belong to.

because of the occurrence of the pandemic by comparing the performance of female and male scientists along a variety of dimensions, such as authorships, leadership in publications, gender composition of collaboration, and citations before and during COVID-19.

In this study, we find that the number of authorships by both male and female scientists increased dramatically during COVID-19, with a sharper magnitude in the growth of males' authorships. However, the number of authorships by teams with a female as the first author decreased by 21%, indicating that females' leadership in research teams as the major contributor has been negatively affected. In contrast, the number of authorships by teams with a male first author went up by 7.4% after the first global COVID-19 case. The proportion of females' authorships declined by 5%. These results suggest that, although both females and males published more papers during the pandemic, relative to the pre-pandemic period, the gender gaps in the shares of authorships have been strengthened due to the larger increase in males' authorships. These results provide evidence that gender inequalities have been amplified during COVID-19 from diverse perspectives. The closure of public places and the introduction of social-distancing measures have forced scientists to work from home. Factors including gender roles and the social division of labor determine that females have to spend more time taking care of their families than males (Chauhan, 2020a), which might dampen their participation in science, especially if they play a key role in scientific teams.

We further observe that the share of publications by single-gender collaboration grew during COVID-19, while that by mixed-gender collaboration declined, which suggests a negative effect of the pandemic on collaboration between females and males. The empirical evidence shows that the fraction of publications by all-female teams and that by all-male teams increased by 1.8% and 5%,

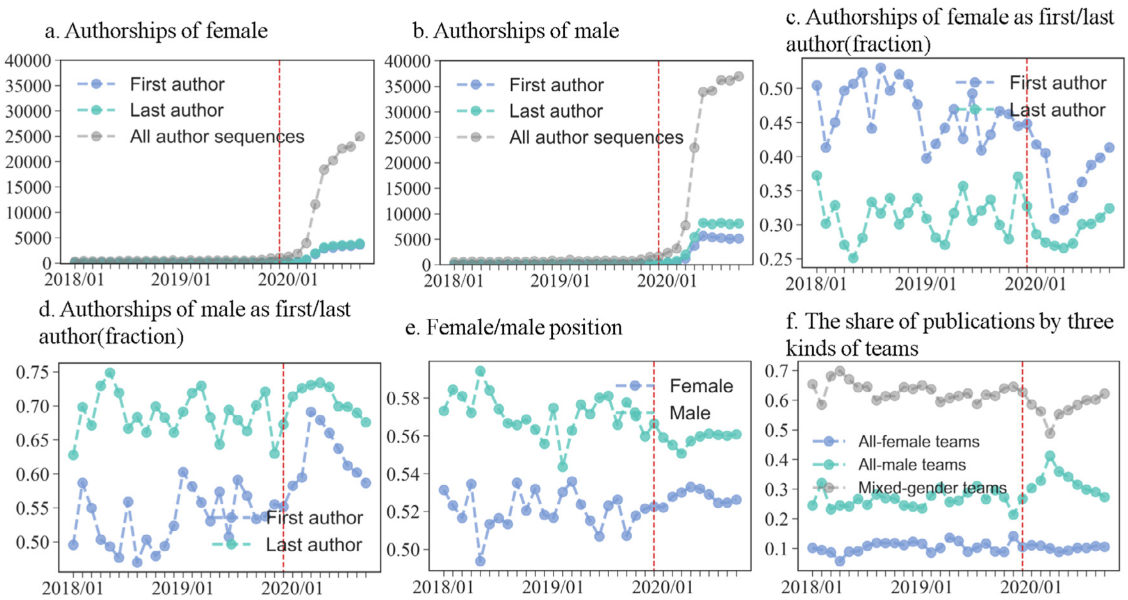


Fig. 10. Evolution of dependent variables for the 50 sampled countries before and during COVID-19. The dashed lines indicate the month of the global first COVID-19 case, i.e., December 2019.

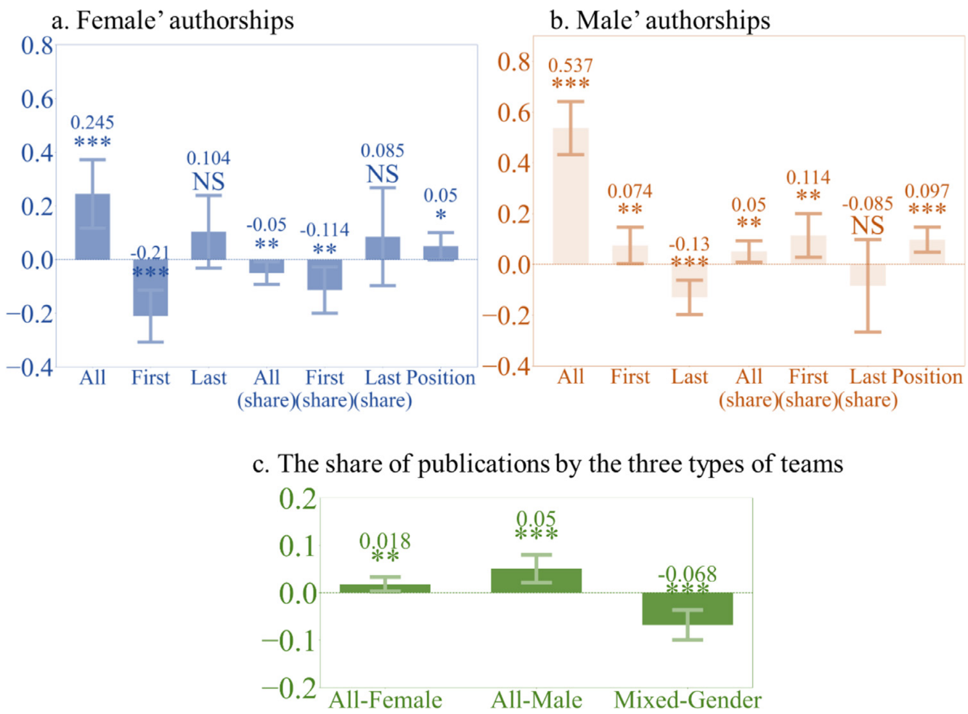


Fig. 11. The relationship between COVID-19, and females' authorships and males' authorships in teams, and the shares of publications by three types of teams. ***, **, and * represent significance at the 1, 5, and 10% levels and NS stands for not significant.

respectively, whereas the proportion of publications by mixed-gender teams declined by 6.8% during the pandemic. Previous studies suggest that people of both genders are inclined to collaborate with peers of the same gender (Gallivan & Ahuja, 2015; Holman & Morandin, 2019). The sudden outbreak of COVID-19 prompted an increase in same-gender collaboration, especially among males, and led to an initial decline of mixed-gender collaboration. As the pandemic progressed, collaboration between females and males gradually returned to the pre-pandemic level. "Homophily" or "similarity" determines the construction and maintenance of scientific collaboration because people with similar backgrounds tend to connect with each other (Freeman & Huang, 2015; Zhang, Bu, Ding

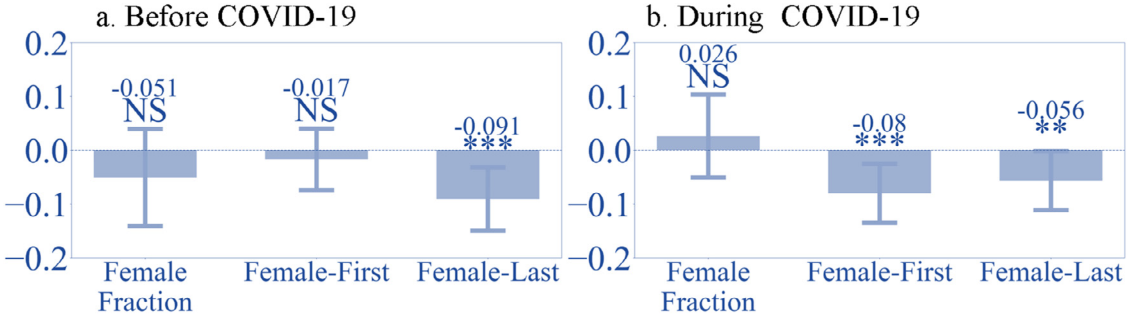


Fig. 12. The relationship between gender composition of teams and citations papers received. ***, **, and * represent significance at the 1, 5, and 10% levels and NS stands for not significant.

& Xu, 2018). Gender-based homophily in science has been recently highlighted by a number of authors (Jadidi, Karimi, Lietz & Wagner, 2018; Kwiek & Roszka, 2021a, 2021b; Wagner, 2018). A recent study reveals a pattern that scientists publish predominantly with those of the same gender, which is known as the gender homophily principle and was found to apply to male scientists (Kwiek & Roszka, 2021a). Collaboration between females and males might entail more communication costs and barriers, and thus might be more vulnerable due to the negative consequences of the COVID-19 pandemic. With fewer communication and coordination costs, collaboration between scientists of the same gender might be favored during the pandemic. Yet, more detailed, causal-level conclusions cannot be drawn based on the current findings.

The existing gender inequalities in citations existed prior to COVID-19. Specifically, we observe that papers with a female as the last author received 9.1% fewer citations than those with a male as the last author before COVID-19. During COVID-19, papers with a female as the main contributor (either as the first author or the last author) were less cited than their counterparts. Those results suggest that the existing gender gaps in citations have been amplified during COVID-19. Previous literature suggests that there is no statistically significant gender disparity regarding creative potential and ability (Baer & Kaufman, 2008; Runco, Cramond & Pagnani, 2010). However, papers with a female as the last author were less cited than those with a male as the last author, regardless of whether it was before or during the pandemic period.

This study further shows that gender gaps concerning various aspects had recovered to the pre-pandemic level by the end of the study period. This might indicate that females' relative authorships and collaboration with their male peers were immediately affected by COVID-19, but this negative effect was not long-lasting and was alleviated with the progression of COVID19.

The longstanding gender inequalities in science have been detrimental to female scientists (Huang, Gates, Sinatra & Barabási, 2020). This study shows that COVID-19 has undoubtedly exacerbated this situation and undermined female scientists' career development. Our results indicate that universities and funding agencies may support female scientists to be resilient, productive, and collaborative with potential policy changes that improve gender equity since female scientists were considerably vulnerable during the pandemic in aspects of authorship, leadership, collaboration, and scientific impacts.

Prior literature claimed a narrowing or stable trend of gender inequalities in some disciplines or some aspects of science (Gruber, Mendle, Lindquist, Schmader & Williams, 2020; Hart et al., 2019), while this study suggests that the trend has been reversed during COVID-19. Our results showed that COVID-19 has hampered females' leadership as major contributors in teams, and their collaboration with males. Males often take leading positions in science and might largely shape coronavirus-related research (Pinho-Gomes et al., 2020). Considering the fact that females are still the minority in many disciplines, the negative impact of COVID-19 on their participation as team leaders and their collaboration with males might bring larger estrangement from males, and make females more isolated from the core of the scientific community. Science policies should be implemented to encourage females' engagement in research projects as leaders, and facilitate their collaboration with males. The policy supports and funding supports should be provided to female scientists, especially the field that was most affected by and related to the pandemic, to help overcome the challenges female scientists faced and mitigate the enlarged gender gaps. Additionally, mixed-gender collaboration in the most affected field should be encouraged as empirical evidence shows that the research quality measured by journal prestige level of mixed-sex publications, is higher than that of same-sex publications for both male and female scientists (Kwiek & Roszka, 2021a).

The limitation of this study is that our analyses do not cover papers in other domains, especially those in humanities and social sciences. The findings are only applicable to the coronavirus-related domain, the field that was most related to the pandemic, and cannot fully reflect the influence of such an event on different genders throughout the whole scientific community. There might be a publication delay in that the publication time of journal papers in the CORD dataset should be later than the time when papers are produced. However, all the sampled papers in this study are coronavirus-related papers that might have received fast-track peer review in journals before being published due to the severity of COVID-19. Another limitation is that only publications with both country and gender information of authors are included in our analyses, accounting for only a third of the overall coronavirus-related literature. Additionally, there might be differences in publication intensities, author hierarchies, and citation practices across fields in the category of biomedicine, which are not considered. The citation data in this study is up to January 2021. The use of citation information of COVID-related papers in a short time window is limited. It is necessary to explore the impact of COVID-19 on gender inequalities concerning scientific impacts in a longer time window in future studies.

CRediT authorship contribution statement

Meijun Liu: Conceptualization, Writing – original draft, Data curation, Investigation. **Ning Zhang:** Data curation, Investigation. **Xiao Hu:** Data curation, Writing – original draft. **Ajay Jaiswal:** Data curation, Writing – original draft. **Jian Xu:** Data curation, Writing – original draft. **Hong Chen:** Data curation, Writing – original draft. **Ying Ding:** Conceptualization, Supervision, Writing – review & editing. **Yi Bu:** Conceptualization, Supervision, Investigation, Writing – review & editing.

Acknowledgments

Ying Ding would like to acknowledge the funding support from NSF RAPID 2028717 for this research. This work is supported by the Youth Program of National Natural Science Foundation in China (No: 72104054 and 72104007), Youth Project of Humanities and Social Sciences of the Ministry of Education of China (No.: 21YJC870001), Shanghai Pujiang Program (21PJC026) and "ISTIC-Taylor & Francis Joint Lab for Academic Frontier Observation Open Fund". The authors would like to thank the constructive comments from the anonymous reviewers.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.joi.2022.101295.

References

Abramo, G., D'Angelo, C. A., & Murgia, G. (2013). Gender differences in research collaboration. *Journal of Informetrics, 7*(4), 811–822.

Adams-Prassl, A., Boneva, T., Golin, M., & Rauh, C. (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics, 189*(13183). 10.1016/J.JPUBECO.2020.104245.

Agrawal, A. K. (2013). The impact of investor protection law on corporate policy and performance: Evidence from the blue sky laws. *Journal of Financial Economics, 107*(2), 417–435.

Alstadsæter, A., Bratsberg, B., Eielsen, G., Kopczuk, W., Markussen, S., Raaum, O., et al. (2020). The first weeks of the coronavirus crisis: Who got hit, when and why? Evidence from norway. *NBER working papers*. National Bureau of Economic Research Retrieved from https://www.nber.org/system/files/working_papers/w27131/w27131.pdf .

Amano-Patiño, N., Faraglia, E., Giannitsarou, C., & Hasna, Z. (2020). Who is doing new research in the time of COVID-19? Not the female economists. *Publishing and Measuring Success in Economics*, 13.

Ammar, W., Groeneveld, D., Bhagavatula, C., Beltagy, I., Crawford, M., Downey, D. et al. (2018). *Construction of the Literature Graph in Semantic Scholar.*Paper presented at the Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 3 (Industry Papers).

Andersen, J.P., .Nielsen, M.W., .Simone, N.L., .Lewiss, R.E., .& Jagsi, R. (2020). COVID-19 medical papers have fewer women first authors than expected. *eLife*, 9, e58807. doi:10.7554/eLife.58807

Angrist, J. D., & Pischke, J.-. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Baer, J., & Kaufman, J. C. (2008). Gender differences in creativity. *The Journal of Creative Behavior, 42*(2), 75–105.

Beaudry, C., & Larivière, V. (2016). Which gender gap? Factors affecting researchers' scientific impact in science and medicine. *Research policy,* 45(9), 1790–1817. doi:HTtps://doi.org/10.1016/j.respol.2016.05.009

Beddoes, K., & Pawley, A. L. (2014). Different people have different priorities': Work–family balance, gender, and the discourse of choice. *Studies in Higher Education, 39*(9), 1573–1585.

Bell, & Fong, K. C. (2021). Gender differences in first and corresponding authorship in public health research submissions during the COVID-19 pandemic. *American Journal of Public Health, 111*(1), 159–163.

Bell, N.E. (.2010). *Graduate enrollment and degrees: 1999 to 2009*: Council of Graduate Schools, Graduate Record Examinations Board.

Boschini, A., & Sjögren, A. (2007). Is team formation gender neutral? Evidence from coauthorship patterns. *Journal of Labor Economics, 25*(2), 325–365.

Brown, C.S., .& Ravallion, M. (2020). Inequality and the Coronavirus: Socioeconomic Covariates of Behavioral Responses and Viral Outcomes Across US Counties. *Nber Working Papers*.

Brubaker, L. (2020). Women Physicians and the COVID-19 Pandemic. *JAMA, 324*(9), 835–836. 10.1001/jama.2020.14797.

Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current directions in psychological science, 19*(5), 275–279.

Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences, 108*(8), 3157–3162. 10.1073/pnas.1014871108.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218–261.

Chan, H. F., & Torgler, B. (2020). Gender differences in performance of top cited scientists by field and country. *Scientometrics, 125*(3), 2421–2447.

Chauhan, P. (2020a). Gendering COVID-19: Impact of the pandemic on women's burden of unpaid work in India. *Gender issues*, 1–25. 10.1007/s12147-020-09269-w.

Chauhan, P. (2020b). Gendering COVID-19: Impact of the pandemic on women's burden of unpaid work in India. *Gender issues*, 1–25.

Cole, J. R., & Zuckerman, H. (1984). The productivity puzzle. *Advances in motivation and achievement. women in science*. Greenwich, CT: JAI Press.

Collins, C., Landivar, L.C., .Ruppanner, L., & Scarborough, W.J. (.2020). COVID-19 and the gender gap in work hours. gender, Work & organization. *Forthcoming*.

Conley, D., & Stadmark, J. (2012). Gender matters: A call to commission more women writers. *Nature, 488*(7413), 590.

Cui, R., Ding, H., & Zhu, F. (2021). Gender inequality in research productivity during the COVID-19 pandemic. *Manufacturing & Service Operations Management*.

Das, S., & Paik, J. H. (2021). Context-sensitive gender inference of named entities in text. *Information Processing & Management, 58*(1), Article 102423.

Derrick, G.E., .Jaeger, A., Chen, P.-.Y., Sugimoto, C.R., .Van Leeuwen, T., & Larivière, V. (2019). Models of parenting and its effect on academic productivity: Preliminary results from an international survey.

Eals, M., & Silverman, I. (1994). The Hunter-Gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. *Ethology and Sociobiology, 15*(2), 95–105. 10.1016/0162-3095(94)90020-5.

Fabris, A., Purpura, A., Silvello, G., & Susto, G. A. (2020). Gender stereotype reinforcement: Measuring the gender bias conveyed by ranking algorithms. *Information Processing & Management, 57*(6), Article 102377.

Farré, L., Fawaz, Y., González, L., & Graves, J. (2020). *How the COVID-19 lockdown affected gender inequality in paid and unpaid work in Spain*. IZA Discussion Paper No. 13434. Retrieved from SSRN: Https://ssrn.com/abstract=3643198

Feramisco, J. D., Leitenberger, J. J., Redfern, S. I., Bian, A., Xie, X.-. J., & Resneck Jr, J. S. (2009). A gender gap in the dermatology literature? Cross-sectional analysis of manuscript authorship trends in dermatology journals during 3 decades. *Journal of the American Academy of Dermatology, 60*(1), 63–69.

Ferber, M. A., & Teiman, M. (1980). Are women economists at a disadvantage in publishing journal articles? *Eastern Economic Journal, 6*(3/4), 189–193.

Fosch-Villaronga, E., Poulsen, A., Søraa, R. A., & Custers, B. (2021). A little bird told me your gender: Gender inferences in social media. *Information Processing & Management, 58*(3), Article 102541.
Fox, M. F., & Faver, C. A. (1985). Men, women, and publication productivity: Patterns among social work academics. *The Sociological Quarterly, 26*(4), 537–549.
Freeman, R. B., & Huang, W. (2015). Collaborating with people like me: Ethnic coauthorship within the United States. *Journal of Labor Economics, 33*(S1), S289–S318.
Gallivan, M., & Ahuja, M. (2015). Co-authorship, homophily, and scholarly influence in information systems research. *Journal of the Association for Information Systems, 16*(12), 2.
Gao, J., Yin, Y., Myers, K. R., Lakhani, K. R., & Wang, D. (2021). Potentially long-lasting effects of the pandemic on scientists. *Nature communications, 12*(1), 6188. 10.1038/s41467-021-26428-z.
Gomes, A. O., & Avellaneda, C. N. (2021). The role of pro-women institutions in addressing violencereports against women. *Global Public Policy and Governance, 1*(1), 39–60. 10.1007/s43508-021-00003-0.
Graham, J., & Shaffer, D. (2004). Occupational segregation, gender gaps and skill gaps. *Equal Opportunities Commission*, 155–168.
Gruber, J., Mendle, J., Lindquist, K. A., Schmader, T., & Williams, L. A. (2020). The future of women in psychological science. *Perspectives on Psychological Science* 1745691620952789.
Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological science in the public interest, 8*(1), 1–51.
Hart, K. L., Frangou, S., & Perlis, R. H. (2019). Gender trends in authorship in psychiatry journals from 2008 to 2018. *Biological psychiatry, 86*(8), 639–646.
Hill, C., Corbett, C., & Rose, A. S. (2010). Why so few? *Women in science, technology, engineering, and mathematics* 1111 Sixteenth Street NW, Washington, DC: American Association of University Women 20036. Tel: 800-326-2289; Tel: 202-728-7602; Fax: 202-463-7169; e-mail: Foundation@aauw.org; Web site: Http://www.aauw.org.
Holman, L., & Morandin, C. (2019). Researchers collaborate with same-gendered colleagues more often than expected across the life sciences. *PloS one, 14*(4), Article e0216128.
Hossain, M. (2021). Gender differences in experiencing coronavirus-triggered economic hardship: Evidence from four developing countries. *Research in Social Stratification and Mobility, 71*, Article 100555.
Huang, J., Gates, A. J., Sinatra, R., & Barabási, A.-. L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences, 117*(9), 4609–4616.
Hyde, J. S., & Linn, M. C. (2006). Gender similarities in mathematics and science. *Science (New York, N.Y.), 314*(5799), 599–600.
Ivanova-Stenzel, R., & Kübler, D. (2011). Gender differences in team work and team competition. *Journal of Economic Psychology, 32*(5), 797–808. 10.1016/j.joep.2011.05.0117.
Jadidi, M., Karimi, F., Lietz, H., & Wagner, C. (2018). Gender disparities in science? Dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems, 21*(03n04), Article 1750011.
Johnson, H., Irish, W., Strassle, P., Mahoney, S., Schroen, A., Josef, A., et al. (2020). Associations between career satisfaction, personal life factors, and work-life integration practices among US surgeons by gender. *JAMA Surgery*, 155. 10.1001/jamasurg.2020.1332.
Kibbe, M. R. (2020). Surgery and COVID-19. *JAMA, 324*(12), 1151–1152.
Krukowski, R. A., Jagsi, R., & Cardel, M. I. (2021). Academic productivity differences by gender and child age in science, technology, engineering, mathematics, and medicine faculty during the COVID-19 pandemic. *Journal of Women's Health, 30*(3), 341–347.
Kwiek, M., & Roszka, W. (2021a). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics, 15*(3), Article 101171.
Kwiek, M., & Roszka, W. (2021b). Why collaboration with men is dominating in science? Gender homophily among 25,000 academic scientists. *Nauka* 39-78-39-78.
Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature, 504*(7479), 211–213.
Li, Y., Ding, L., & Yang, Y. (2020). Can the Introduction of an Environmental Target Assessment Policy Improve the TFP of Textile Enterprises? A Quasi-Natural Experiment Based on the Huai River Basin in China. *Sustainability, 12*(4) 1696–1696. 10.3390/su12041696.
Liang, L., Huang, D.-. S., & Yang, T. (2020). Associations of various subspecialties in surgery with career satisfaction and personal life among surgeons by gender. *JAMA Surgery*. 10.1001/jamasurg.2020.5637.
Liu, M., Bu, Y., Chen, C., Xu, J., Li, D., Leng, Y., et al. (2021). Pandemics are catalysts of scientific novelty: Evidence from COVID-19. *Journal of the Association for Information Science and Technology*.
Liu, M., & Hu, X. (2021). Will collaborators make scientists move? A Generalized Propensity Score analysis. *Journal of Informetrics, 15*(1), Article 101113. 10.1016/j.joi.2020.101113.
Liu, M., Shi, D., & Li, J. (2017). Double-edged sword of interdisciplinary knowledge flow from hard sciences to humanities and social sciences: Evidence from China. *PloS one, 12*(9), Article e0184977.
Liu, M., Zangerle, E., Hu, X., Melchiorre, A., & Schedl, M. (2020). Pandemics, music, and collective sentiment: Evidence from the outbreak of COVID-19. *Paper presented at the 21th international society for music information retrieval conference*.
Long, J. S. (1992). Measures of sex differences in scientific productivity. *Social Forces, 71*(1), 159–178.
Luoto, S. (2020). Sex differences in people and things orientation are reflected in sex differences in academic publishing. *Journal of Informetrics, 14*(2), Article 101021.
McDowell, J. M., Singell Jr, L. D., & Ziliak, J. P. (2001). Gender and promotion in the economics profession. *ILR Review, 54*(2), 224–244.
McDowell, J. M., & Smith, J. K. (1992). The effect of gender-sorting on propensity to coauthor: Implications for academic promotion. *Economic Inquiry, 30*(1), 68–82.
Min, C., Bu, Y., Wu, D., Ding, Y., & Zhang, Y. (2021). Identifying citation patterns of scientific breakthroughs: A perspective of dynamic citation process. *Information Processing & Management, 58*(1), Article 102428.
Minello, A. (2020). The pandemic and the female academic. *Nature*. 10.1038/d41586-020-01135-9.
Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. In *Proceedings of the National Academy of Sciences of the United States of America: 109* (pp. 16474–16479).
Muriithi, P., Horner, D., & Pemberton, L. (2013). Understanding factors influencing the effect of scientific collaboration on productivity in a developing country: Kenya. *Proceedings of the American Society for Information Science and Technology, 50*(1), 1–10.
Myers, K. R., Tham, W. Y., Yin, Y., Cohodes, N., Thursby, J. G., Thursby, M. C., et al. (2020). Unequal effects of the COVID-19 pandemic on scientists. *Nature Human Behaviour, 4*(9), 880–883. 10.1038/s41562-020-0921-y.
Nguyen, A. X., Trinh, X.-. V., Kurian, J., & Wu, A. Y. (2021). Impact of COVID-19 on longitudinal ophthalmology authorship gender trends. *Graefe's Archive for Clinical and Experimental Ophthalmology, 259*(3), 733–744.
Peters, B. G. (2021). Governing in a time of global crises: the good, the bad, and the merely normal. *Global Public Policy and Governance, 1*(1), 4–19. 10.1007/s43508-021-00006-x.
Pieh, C., Budimir, S., & Probst, T. (2020). The effect of age, gender, income, work, and physical activity on mental health during coronavirus disease (COVID-19) lockdown in Austria. *Journal of psychosomatic research, 136*, Article 110186.
Pinho-Gomes, A.-. C., Peters, S., Thompson, K., Hockham, C., Ripullone, K., Woodward, M., et al. (2020). Where are the women? Gender inequalities in COVID-19 research authorship. *BMJ Global Health, 5*(7), Article e002922.
Rachid, E., Noureddine, T., Tamim, H., Makki, M., Naalbandian, S., & Al-Haddad, C. (2021). Gender disparity in research productivity across departments in the faculty of medicine: A bibliometric analysis. *Scientometrics*, 1–17.
Ribarovska, A. K., Hutchinson, M. R., Pittman, Q. J., Pariante, C., & Spencer, S. J. (2021). Gender inequality in publishing during the COVID-19 pandemic. *Brain, behavior, and immunity, 91*, 1.
Roth, J. (2019). *Pre-test with caution: Event-study estimates after testing for parallel trends*. Department of Economics, Harvard University *Unpublished manuscript*.
Ruggieri, R., Pecoraro, F., & Luzi, D. (2021). An intersectional approach to analyse gender productivity and open access: A bibliometric analysis of the Italian National Research Council. *Scientometrics, 126*(2), 1647–1673.
Runco, M. A., Cramond, B., & Pagnani, A. R. (2010). Gender and creativity. In *Handbook of gender research in psychology* (pp. 343–357). Springer.

Santamaría, L., & Mihaljević, H. (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science, 4*, e156.
Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *American psychologist, 60*(9), 950.
Squazzoni, F., Bravo, G., Farjam, M., Marusic, A., Mehmani, B., Willis, M., et al. (2021). Peer review and gender bias: A study on 145 scholarly journals. *Science advances, 7*(2), eabd0299.
Stagni, R. M., Fosfuri, A., & Santaló, J. (2021). A bird in the hand is worth two in the bush: Technology search strategies and competition due to import penetration. *Strategic Management Journal*, (42), 1516–1544. 10.1002/smj.3277.
Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological bulletin, 135*(6), 859–884. 10.1037/a0017364.
Thelwall, M. (2018). Do females create higher impact research? Scopus citations and Mendeley readers for articles from five countries. *Journal of Informetrics, 12*(4), 1031–1041.
Thelwall, M., Bailey, C., Tobin, C., & Bradshaw, N.-. A. (2019). Gender differences in research areas, methods and topics: Can people and thing orientations explain the results? *Journal of Informetrics, 13*(1), 149–169.
Viglione, G. (2020). Are women publishing less during the pandemic? Here's what the data say. *Nature, 581*(7809), 365–366.
Vincent-Lamarre, P., Sugimoto, C.R., .& Larivière, V. (2020). *Monitoring Women's Scholarly Production during the COVID-19 Pandemic*. Retrieved from http://scholar.googleusercontent.com/scholar?q=cache:FOHSrUJvXJEJ:SCholar.google.com/+Vincent-Lamarre,+P.,+Sugimoto,+C.+R.,+%26+Larivi%C3%A8re,+V.+(2020).+Monitoring+Women%27s+Scholarly+Production+during+the+COVID%E2%80%9019+Pandemic.+In.&hl=en&as_sdt=0,5
Wagner, C. S. (2018). *The collaborative era in science: Governing the network*. Springer.
Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365–391.
Wang, Z., Yin, Q. E., & Yu, L. (2021). Real effects of share repurchases legalization on corporate behaviors. *Journal of Financial Economics, 140*(1), 197–219.
Wehner, M. R., Li, Y., & Nead, K. T. (2020). Comparison of the proportions of female and male corresponding authors in preprint research repositories before and during the COVID-19 pandemic. *JAMA network open, 3*(9), Article e2020335.
West, J. D., Jennifer, J., King, M. M., Correll, S. J., Bergstrom, C. T., & Lilach, H. (2013). The role of gender in scholarly authorship. *PloS one, 8*(7), e66212.
Witteman, H. O., Haverfield, J., & Tannenbaum, C. (2021). COVID-19 gender policy changes support female scientists and improve research quality. *Proceedings of the National Academy of Sciences*, (6), 118.
Xie, Y., & Shauman, K. A. (1998). Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review*, 847–870.
Zhang, C., Bu, Y., Ding, Y., & Xu, J. (2018). Understanding scientific collaboration: Homophily, transitivity, and preferential attachment. *Journal of the Association for Information Science and Technology, 69*(1), 72–86.