The effect of the COVID-19 pandemic on the gender gap in research productivity within academia

Kiran Gok Lune Lee1,2, Adele Mennerat3, Alecia Carter4, Dieter Lukas5,§, Hannah Dugdale1,§, Antica Culina6,7,§,

1 Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands

2 Department of Animal & Plant Sciences, University of Sheffield, Sheffield, UK

3 Department of Biological Sciences, University of Bergen, Bergen, Norway

4 Department of Human Behavior, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

5 Rudjer Boskovic Institute, Zagreb, Croatia

6 Netherlands Institute of Ecology, NIOO-KNAW, Wageningen, The Netherlands

§Shared last authorship

# Abstract

**Background**:

The number of published articles is one of the most commonly-used metrics for evaluating academic success. It thus directly influences success in getting academic jobs, promotion, funding, and collaboration. However, this metric focuses on processes that are gender-biased, selecting against female academics. Novel social conditions during the COVID-19 pandemic may have exacerbated this gender bias if female academics performed extra work in caregiving, domestic, service and teaching roles at the expense of research productivity. We investigate the pandemic effect on the gender gap in research productivity through a systematic review and meta-analysis of published articles across scientific disciplines.

**Methods**:

Our systematic review identified 50 relevant studies with 115 effect sizes that measure the impact of the COVID-19 pandemic on gender-specific research productivity, measured mainly as the numbers of submitted/published articles (n=97), but including other self-reported measures too. We then conducted a meta-analysis to: 1) investigate the effect of the pandemic on the gender gap in research productivity within academia and 2) test hypotheses on how research field, breadth of gender gap before the pandemic, and authorship position influence this effect.

**Results**:

We find that, overall, the gender gap in research productivity within academia has increased during the COVID-19 pandemic compared to before, especially in social sciences and medicine, fields that were previously nearest to being gender equal. We did not detect an influence of authorship position on the effect.

**Conclusions**:

We detected that gender biases favouring the productivity and perceived impact of men in academia overall strengthened during the pandemic. We encourage academic and funding institutions to consider a range of different metrics to evaluate academic impact, and to further acknowledge and accommodate individual circumstances.

Subjects: Academia, Gender Bias, Research Productivity

Key words: COVID-19, Pandemic, Academia, Gender bias, Women researchers, Research productivity

# Introduction

### Background

Many traditionally used metrics for evaluating academic merit have clear biases against individuals historically underrepresented in academia (Davies {\i{}et al.}, 2021; Shandera {\i{}et al.}, 2021). Such metrics include the number (and perceived prestige) of manuscripts or publications (Larivière *et al.*, 2013; West *et al.*, 2013; Huang *et al.*, 2020), or the amount of funding acquisition (Shen, 2013; Valentova *et al.*, 2017; James, Chisnall and Plank, 2019; Safdar *et al.*, 2021). Those who publish less, or acquire less funds are commonly perceived as having lower research productivity. This can negatively impact other metrics for evaluating merit and career progression, “snowballing” the strength of selection against certain individuals in academia (Bol, Vaan and Rijt, 2018). Academic research comes with many traditional restrictions and constraints that bias access to merit and career progression towards those in the right socio-economic circumstances (Morgan *et al.*, 2022). Many of these circumstances intersect with gender, leading to a large body of research showing disadvantages can compound for women compared to men when success is measured using traditional metrics, despite no actual differences in contribution and impact of research (Armstrong and Jovanovic, 2015; Zeng *et al.*, 2016; Langin, 2019; Huang *et al.*, 2020; Romano, 2021; van der Wal, Thorogood and Horrocks, 2021; Kozlowski *et al.*, 2022). The COVID-19 pandemic has influenced social circumstances by creating novel living and working conditions, potentially worsening the research productivity, measured in these traditional ways, of many women worldwide (Anwer, 2020; Boncori, 2020; Guy and Arthur, 2020; Altan-Olcay and Bergeron, 2022). Here, we investigate the potential effect of the pandemic on publication and submission output and other self-reported measures of research productivity of female and male academics.

Multiple factors are likely to contribute to gender-biased impacts on research productivity during a pandemic. First, women, including those in academia, generally perform more unpaid caregiving and domestic work (Schiebinger, Henderson and Gilmartin, 2008; Schiebinger and Gilmartin, 2010). As many countries enforced social-distancing and closed facilities during the pandemic, this has increased caregiving and domestic burdens (Carli, 2020; Carlson, Petts and Pepin, 2020) at times when community help from nurseries, schools, care homes, house cleaners, laundrettes, nannies, babysitters and family was limited (Myers *et al.*, 2020; Barber *et al.*, 2021; Breuning *et al.*, 2021; Deryugina, Shurchkov and Stearns, 2021; Shalaby, Allam and Buttorff, 2021). These additional tasks, which disproportionately have fallen on women, likely encroached on the time and space for academic research during “work-from-home” conditions (Abdellatif and Gatto, 2020; Boncori, 2020; Guy and Arthur, 2020). Second, the distribution of work within academic institutions is often gendered. Women undertake more ‘non-promotable tasks (Babcock *et al.*, 2022) such as administrative, supportive and mentoring roles (Porter, 2007; Mitchell and Hesli, 2013; Babcock *et al.*, 2017; Guarino and Borden, 2017; O’Meara *et al.*, 2017; O’Meara, Kuvaeva and Nyunt, 2017). Changes in teaching and administration in response to the pandemic were therefore more likely to be facilitated by women (Docka-Filipek and Stone, 2021; Minello, Martucci and Manzo, 2021). Third, labour roles contributing towards publication are also gendered with women generally performing more technical work such as generating data, whilst men assume more core tasks in conceptualisation, analysis, writing and publishing (Macaluso *et al.*, 2016). Pandemic closures to research institutions would therefore likely impact women authorship stronger than men. Additionally, the surge in publications during the pandemic (Else, 2020) could have led to reductions in the quality of peer review, with evaluation being more influenced by cognitive shortcuts. These shortcuts are often associated with biases tending to operate against women (Kaatz, Gutierrez and Carnes, 2014; Reuben, Sapienza and Zingales, 2014; Carli, 2020) such that women tend to have lower success in getting submissions accepted (Chemistry, 2019; Fox and Paine, 2019; Murray *et al.*, 2019; Hagan *et al.*, 2020).

The role these factors had during the pandemic in shaping the gender gap in research production might differ across research fields. One possibility is that research fields that were already more gender-biased may have experienced the most exacerbated gender gaps during the pandemic. In fields that were already traditionally more gender-biased, less support may have been available to women to balance the effects of the pandemic. Male-dominated fields often lack viewpoints of female colleagues, and might therefore be less likely to identify and support paid care work or extended leave options (Clark, 2020; Nash and Churchill, 2020). An alternative possibility is that the pandemic might have eroded the support structures that existed in fields with higher gender balance, thereby reducing the differences in gender biases between research fields. Even in fields that appear more gender-balanced, “glass ceiling” effects remain such that imbalances can increase with higher academic rank (Addessi, Borgi and Palagi, 2012). The pandemic may also have exacerbated a gender gap in authorship position (first, middle or last) (King and Frederickson, 2021) if additional service, teaching, caregiving, and domestic roles taken up by female academics during the pandemic may limit their abilities to perform research (as first authors) or lead research (as last authors) but not supporting research (as middle authors).

Here, we use a systematic review and meta-analysis to study the impacts off the COVID-19 pandemic on research productivity across scientific disciplines, and to test hypotheses on how research field, breadth of gender gap before the pandemic, and authorship position might influence the strength of the effect. The impact of the pandemic on gender biases in research productivity has been explored in anecdotal accounts (Boncori, 2020; Excess Beth [@El\_Dritch], 2020; Fazackerley, 2020; National Academies of Sciences, 2021), surveys of potentially affected people (Myers *et al.*, 2020; Rodríguez-Rivero *et al.*, 2020; Barber *et al.*, 2021; Breuning *et al.*, 2021; Deryugina, Shurchkov and Stearns, 2021; Diaz *et al.*, 2021; Ellinas *et al.*, 2021; Gao *et al.*, 2021; Ghaffarizadeh *et al.*, 2021; Guintivano, Dick and Bulik, 2021; Hoggarth *et al.*, 2021; Krukowski, Jagsi and Cardel, 2021; Maguire *et al.*, 2021; Plaunova *et al.*, 2021; Shalaby, Allam and Buttorff, 2021; Staniscuaski *et al.*, 2021; Yildirim and Eslen-Ziya, 2021; Davis *et al.*, 2022; Stenson *et al.*, 2022), and comparisons of numbers of articles submitted or published by gender before and during the pandemic (Amano-Patiño *et al.*, 2020; Andersen *et al.*, 2020; Bell and Green, 2020; Cushman, 2020; Wehner, Li and Nead, 2020; Bell and Fong, 2021; Biondi *et al.*, 2021; Cook *et al.*, 2021; DeFilippis *et al.*, 2021; Forti, Solino and Szabo, 2021; Fox and Meyer, 2021; Gayet-Ageron *et al.*, 2021; Gerding *et al.*, 2021; Ipe *et al.*, 2021; Jemielniak, Sławska and Wilamowski, 2021; King and Frederickson, 2021; Lerchenmüller *et al.*, 2021; Muric *et al.*, 2021; Nguyen *et al.*, 2021; Quak *et al.*, 2021; Ribarovska *et al.*, 2021; Squazzoni *et al.*, 2021; Williams II *et al.*, 2021; Anabaraonye *et al.*, 2022; Ayyala and Trout, 2022; Chen and Seto, 2022; Cui, Ding and Zhu, 2022; Harris *et al.*, 2022; Wooden and Hanson, 2022). Few studies compare between research fields and authorship positions (Andersen *et al.*, 2020; Forti, Solino and Szabo, 2021; Gayet-Ageron *et al.*, 2021; Jemielniak, Sławska and Wilamowski, 2021; King and Frederickson, 2021; Muric *et al.*, 2021; Nguyen *et al.*, 2021; Squazzoni *et al.*, 2021). One review qualitatively evaluated separate findings of the pandemic effect on the gender gap in research productivity (Herman *et al.*, 2021), though the effects were not quantitatively explored. We build on this literature by calculating the overall effect of the COVID-19 pandemic on productivity measures, and further explore how research field, breadth of gender gap before the pandemic and authorship position might influence the effect.

### Objective

Our main objective was to quantitatively calculate the overall effect of the COVID-19 pandemic on gender bias in research productivity and whether the bias changed compared to the period just prior to the pandemic. Our focus is on comparing the effect of the pandemic on women relative to men. We recognize that gender extends beyond this comparison, and that biases are even more likely to target individuals whose identities are less represented and often ignored. These biases present in a lack of study of the full diversity of gender precluded our ability to include these effects in a meta-analysis. While several of the surveys we include had the option for respondents to identify beyond the binary women/men, none of these studies report on these individuals, presumably because of the respective small samples. In addition, all studies using numbers of submissions or publications (33 out of 50) to measure research productivity used automatic approaches that are more likely to mis-gender individuals as they inferred binary gender based on first names. While these approaches seemingly offer the potential to identify trends in larger samples, they themselves introduce and reinforce biases in relation to gender that are hard to assess, intersecting with biases in ethnicity as these approaches are often restricted to names common in English speaking countries (Mihaljević *et al.*, 2019). The patterns we describe should be seen as a potential indication that biases exist, but alternative approaches are needed to speculate about potential underlying causes and remedies.

Specifically, we tested three hypotheses and their corresponding predictions. First, we hypothesised that the pandemic has increased the disparity in research productivity of women versus men because of women undertaking more domestic roles at home, and more ‘non-promotable tasks’ at work. We predicted that the gender gap in research productivity has increased during the pandemic compared to the period just before. We predicted that though studies can measure the type of research productivity by individual survey responses, numbers of submissions and numbers of publications, this does not influence the gender gap increase observed during the pandemic. Second, we hypothesised that the pandemic has increased the gender gap differentially because of differences in working conditions and support structures in place. We explored differences in the gender gap increase across research fields and predicted the gender gap is exacerbated in fields that already had a previously greater gender gap because of less support available to women to balance the effects of the pandemic. Third, we hypothesised the disparity in favourable authorship positions has increased because female academics have been more limited in undertaking leading research roles in lockdown conditions. We predict that the pandemic has increased the gender gap more in first and last, rather than middle authorship positions.

# Methods

### Search process

We carried out a systematic review to identify, select and critically evaluate relevant research through data collection and analysis. We reported it following PRISMA guidelines (Moher *et al.*, 2009). We carried out the literature search process in three steps: 1) a scoping search, 2) an initial search with pre-selected author terms, and 3) a refined search using terms as recommended by the litsearchR 1.0.0 (Grames *et al.*, 2019). We initially performed a scoping search to determine if there were over ten texts with primary research investigating differences by gender in academic productivity before and during the pandemic. The scoping search was conducted on 30/06/2021 by Google searching combinations of synonyms for: 1) the COVID-19 pandemic, 2) gender, 3) academia, 4) inequality and 5) productivity. The scoping search identified 21 original research publications with quantitative metrics investigating differences in academic productivity by gender before and during the pandemic (scoped texts). Of these 21 articles, 14 were indexed by Web of Science, and 17 (including the same 14 from Web of Science) were indexed by Scopus.

Terms for the initial search were selected by scanning the title, abstract and keywords of scoped texts. We constructed an initial Boolean search string according to the PICO (Population, Intervention, Comparator, Outcome) framework (Livoreil *et al.*, 2017). Population was represented by “academia”, Intervention by “pandemic”, Comparator by “gender” and Outcome by “inequality” and “productivity”. A sixth concept group contained terms used to exclude irrelevant studies that did not investigate studies in hypothesis one. Terms within concept groups were connected by the Boolean OR operator, and the concept groups were connected by the AND or AND NOT operators, enabling searches for any combination that includes one term from each of the six concept groups (Table S1). Terms in the initial search were selected by scanning the title, abstract and keywords of scoped texts. The initial search in Scopus generated 722 texts, published from 2020 onwards, including 14/17 (82.4%) of scoped texts indexed by Scopus.

To improve the 14/722 (1.9%) efficiency of finding scoped texts from our initial search, we imported all 722 texts into R and used litsearchR. Using litsearchR, potential key terms were extracted from the title, abstract and keywords of texts using the Rapid Automatic Keyword Extraction algorithm. A ranked list of important terms was then created from building a key term co-occurrence network (Table S2). Six high-strength terms within the key term co-occurrence matrix, describing research not relevant to our study, such as those of an epidemiological or experimental nature, were added to the AND NOT operator concept group to exclude texts mentioning these terms. Table 1 describes terms of the refined Boolean search string and their respective concept groups. We performed the refined search on 27/07/2021 and generated 700 total texts combined from Scopus (126 texts, including 14/17 articles found in the scoping search), the Web of Science core collection (199 texts), EBSCO (276 texts and Proquest (99 texts) from 2020 onwards. The final search hit rate had an efficiency of 11% (14/126) on Scopus, above the 10% recommended hit rate (Foo *et al.*, 2021). After removing duplicates, 580 articles remained to enter the study screening stage.

Table 1. Final Boolean search string used in full literature search. Texts were limited to those since 2020. Terms in italics were added using litsearchR.

|  |  |  |
| --- | --- | --- |
| Concept group | PICO group | Terms |
| Academia | Population | ( academi\* OR author\* OR database\* OR journal\* OR research OR scien\* ) |
| Gender | Population | AND  ( female\* OR gender OR male\* OR men OR women ) |
| Pandemic | Intervention | AND  ( coronavirus OR covid OR pandemic ) |
| Inequality | Comparator | AND  ( bias\* OR disparit\* OR disproportion\* OR fewer OR gap OR "gender difference\*" OR imbalance\* OR inequalit\* OR inequit\* OR parity OR "sex difference\*" OR skew\* OR unequal ) |
| Productivity | Outcome | AND  ( performan\* OR publication\* OR publish\* OR productiv\* ) |
| Exclusion of biomedical studies | Population | AND NOT  (experiment OR laboratory OR mortality OR surviv\* OR *"acute respiratory"* OR *gis* OR *icu* OR *risk* OR *rna* OR *symptoms* ) |

### Study screening

To be included in our meta-analysis, a study had to quantitatively investigate gender differences in productivity within academia before and during the pandemic. Thus, we screened the titles, abstract and keywords to keep only those suggesting the study investigated: 1) academia, 2) genders, 3) pandemic and 4) some measure of productivity (table S3). To ensure repeatability of the screening process, we used Rayyan.ai (Ouzzani *et al.*, 2016) to blind the inclusion or exclusion of 420 randomly selected texts by two reviewers (K.L. and D.L.). The agreement rate between reviewers was 97%, with 49 articles that both authors agreed to include, 357 articles which both excluded, ten articles one reviewer included but the other excluded, and 4 articles only included by the other reviewer. This agreement rate resulted in a, “strong” (McHugh, 2012) to “near perfect” (Landis and Koch, 1977) Cohen’s kappa of 0.86. Of the 14 articles which were included by one but excluded the other, 3 were included after joint review. Of the remaining 160 texts not included in the screening by two reviewers, we included 18 and excluded 162 during the initial screen. Overall, out of the 580 texts, 70 were retained (Fig. S1) for the full text screening.

Full texts were then screened, including studies that had: 1) for both genders, 2) some metric of academic productivity measured, and 3) for before the pandemic compared with during the pandemic. Texts mentioning all criteria as secondary data were excluded. Thus, 25 articles that all contained necessary metrics to calculate effect sizes were retained for data extraction, excluding 45 articles (Fig. S1).

### Iterating the search

To find articles that had been published since the 27/07/21 search (Table 1), we iterated the search and screen process. The second search was repeated on 28/02/2022, generating 1646 total texts combined from Scopus (258 texts, including 14/17 articles found in the scoping search), the Web of Science core collection (413 texts), EBSCO (542 texts) and Proquest (433 texts) from 2020 onwards. We removed 438 duplicates using Rayyan.ai, leaving 1208 de-duplicated articles. To ensure our methods are repeatable, we checked and found all 580 de-duplicated articles from the previous search were also found again. Out of the 1208 texts from the final search, we included 169 after screening titles, abstracts and keywords. For these 169, we screened the full texts, excluding 116 articles and keeping 50 (including the 25 identified in the original search) that all contained the necessary information to calculate the effect sizes.

### Extracting variables

**Effect size**: We extracted values needed to calculate 115 effect sizes from 50 texts investigating the effect of the pandemic on academic research productivity between genders, before and during the pandemic. 11 measures of the effect sizes were already calculated within the articles (2 lasso regression, 2 Somers’ delta, 2 ordered logistic regression, 1 logistic regression and 4 mixed-effect models) and we recorded these as such. For the other 104 effect sizes, we entered summary data (N=101) or statistical inferences (N=3) into Campbell collaboration’s effect size calculator (Wilson, 2019) to calculate a standardised mean difference (d) effect size. For effect sizes calculated using summary data, 83 relied on the proportion of raw numbers of female and male authors before and after the pandemic, and 18 on the mean changes and standard deviations or standard errors in research productivity changes during the pandemic for female and male researchers. For effect sizes calculated from reported statistical tests, one converted the f-test statistic and sample size from a general linear model investigating the effect of gender on perceived work production, one converted the chi-square comparing proportions of female and male academics that experienced productivity changes due to the pandemic, and one converted the p-values from a t-test comparing mean changes in research time due to the pandemic. Two effect sizes (Jemielniak, Sławska and Wilamowski, 2021; Stenson *et al.*, 2022) were calculated using sample sizes obtained by personal correspondence. We calculated multiple effect sizes from one study if they were for different research fields or authorship positions. We set the sign for effect sizes as negative if the pandemic had reduced relative research productivity of women (increased gender gap) and positive if the pandemic had increased the relative research productivity of women (reduced gender gap). A subset of 59 effect sizes were double-checked by A.C., A.M. and D.L and inconsistencies were discussed to ensure repeatability. K.L. then extracted the remaining 56 effect sizes.

**Variance**: Of 9 effect sizes already calculated in the original studies, 6 provided variance as the standard error, which we squared to obtain the variance; and 3 provided the variance as 95% confidence intervals, which we divided by 1.96 and then squared (Nakagawa, Lagisz, Jennions, *et al.*, 2021). For the other 101 effect sizes, variance was estimated in the Campbell collaboration calculator (Wilson, 2019) when calculating effect sizes.

**Research productivity measure**: We first recorded whether the change in research productivity was measured from survey responses (survey response studies, N=23 effect sizes) or from the number of articles submitted or published (article output studies, N=92). Survey studies measured change in research productivity during the pandemic for each gender based on academics self-reporting their gender and change in general productivity (N=11), number of submissions (n=5), research time (N=4), number of projects (N=1), burn-out (N=1), or job loss (N=1). As 5 survey-studies measured research productivity in the number of submissions, we included these studies in the articles submitted and published category. This resulted in 18 effect sizes from surveys measuring some aspect of research productivity, 49 effect sizes measuring numbers of article submissions, and 48 effect sizes measuring numbers of publications.

**Research field**: For the article studies (N=92), we recorded the research field sampled based on the description in the original studies as either medicine (N=38), Technology, Engineering, Mathematics, Chemistry and Physics (N=16), social sciences (N=14), biological sciences (N=16), or multidisciplinary (N=8), following the classification scheme of (Astegiano, Sebastián-González and Castanho, 2019).

**Previous gender bias:** For the article output studies with available data (N=84), we recorded the number of female and male authors before the pandemic in addition to the change in gender bias.

**Authorship position**: For the article output studies (N=92), we recorded whether first (N=44), middle (N=3), last (N=20), corresponding (N=12), or the total number of (N=13) authors were studied. We classified one effect size studying submitting authors, as studying corresponding authors (Fox *et al.*, 2016) and two effect sizes studying sole authors as studying last authors (Moore and Griffin, 2006).

### Analyses

We conducted all analyses in R 3.6.2 (R Core Team, 2022). We used the ‘metafor’ package 3.0.2 to fit models, and build funnel plots (Viechtbauer, 2010). We used ‘orchaRd’ 0.0.0.9000 to build orchard plots to visualise distribution of effect sizes (points) and their precision (point size), calculated as a function of standard error (Nakagawa, Lagisz, O’Dea, *et al.*, 2021).

We fitted separate models for each prediction. All models included the identity of the article the effect size was extracted from as a random effect to control for dependency in effect sizes obtained from the same study. We tested prediction 1a in a model investigating the overall effect size and we displayed this as an orchard plot. We then tested prediction 1b in a model investigating the method of measuring research productivity (survey responses, number of submissions and number of publications) as a moderator of effect size and displayed this as an orchard plot. We included the outlier (Jemielniak, Sławska and Wilamowski, 2021) in the funnel plot of article output studies because this effect size was obtained by personal correspondence clarifying the sample sizes used in the study, which we assume was verified. We tested prediction 2a in a model investigating research field as a moderator of effect size for article studies in a model and displayed this as an orchard plot. We tested in a model how previous gender bias in research productivity before the pandemic, as measured by the proportion of female authors, influenced effect size and displayed this as a line graph, grouped by research field. To test prediction 2b, we tested in a model authorship position as a moderator on effect size for publication studies. We tested for publication bias by performing a multilevel regression model (Nakagawa, Lagisz, Jennions, *et al.*, 2021) which investigates whether small studies have large effect sizes, including research productivity measure as a moderator and display this relationship in funnel plots. We tested for total heterogeneity (*I2*) using the function ‘i2\_ml’ in ‘orchaRd’.

# 

# Results

### 1a: Has the pandemic increased the gender gap in research productivity across academia?

Across all samples, after controlling for multiple effect sizes from the same study, we found the relative productivity of women to men decreased during the pandemic at -0.070 compared to before the pandemic (95% CI= -0.1020 to -0.0380, SE= 0.0163, p-value= <0.0001, Fig. 1). There is large variation in the 115 effect sizes, with 32 indicating a clear increase in the gender gap and 43 a trend of an increase, while 12 indicate a clear decrease in the gender gap and 28 a trend of a decrease. Total heterogeneity was high (*I2* = 97.8%), with 49.1% of it explained by whether research productivity was measured by survey responses or submission/publication numbers and 49.8% explained by the individual effect sizes.

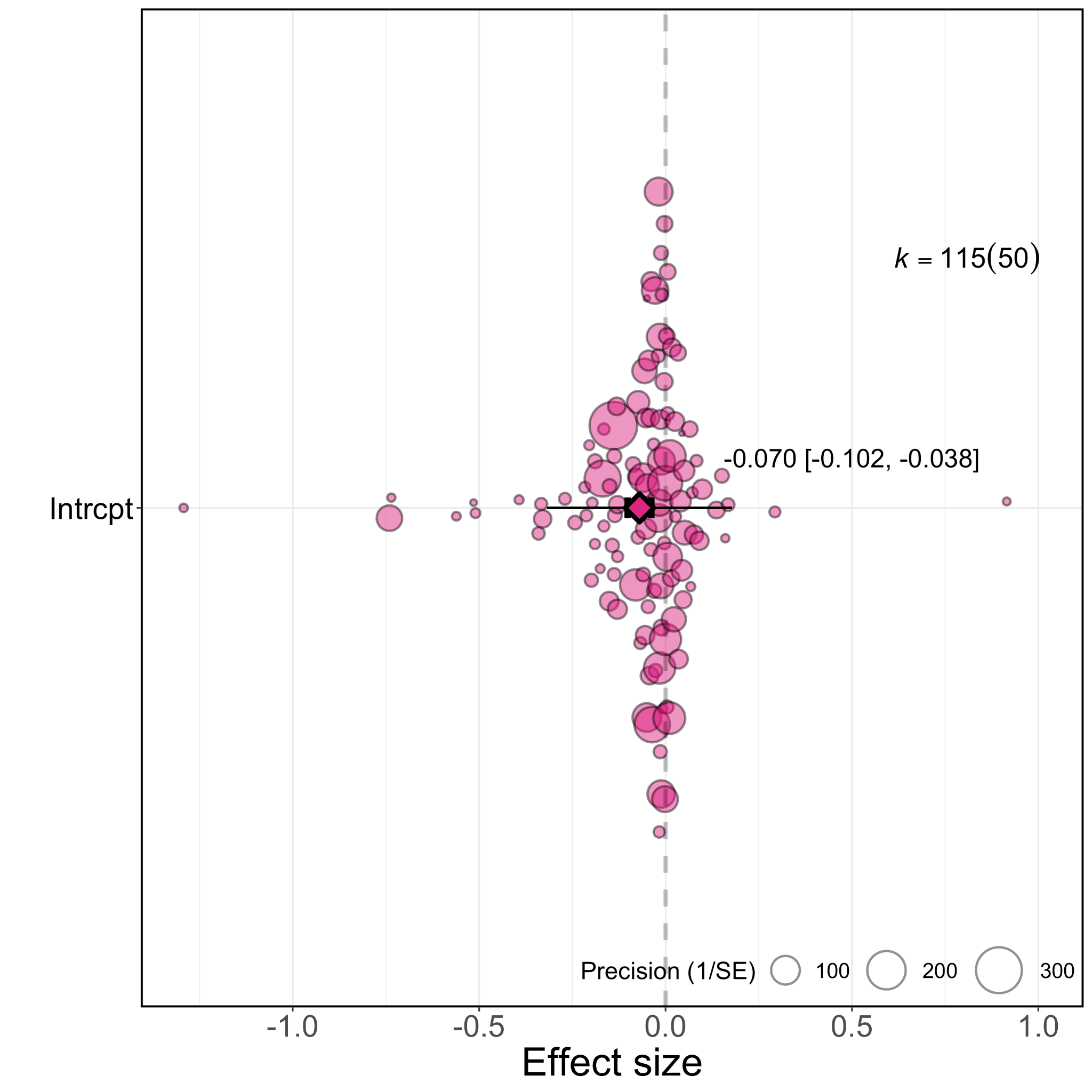


Figure 1. Orchard plot showing all 115 effect sizes (points) and their weight (point size), with the mean effect size (darker coloured point outlined black and vertically centred), the 95% confidence interval (horizontal thick black bar) and the 95% prediction interval of the expected spread of effect sizes based on between-study variance (horizontal thin black bar). k =x(y), where x is the number of effect sizes and y is the number of studies.

### 1b: Does the type of research productivity measure influence how much the pandemic has changed the gender gap in research productivity?

The degree of increase in the gender gap caused by the pandemic differed according to the type of research productivity measured (p-value= 0.0021, Fig. 2). Studies measuring changes to research productivity during the pandemic based on surveys detected a larger overall effect (-0.193, 95% CI= -0.235 to -0.125, SE= 0.041, p-value< 0.001) than studies that compared the number of articles published (-0.046, 95% CI= -0.082 to -0.009, p-value= 0.014, SE= 0.019) or submitted (-0.039, 95% CI= -0.076 to -0.001, p-value= 0.044, SE= 0.019) by authors of each gender before and during the pandemic.

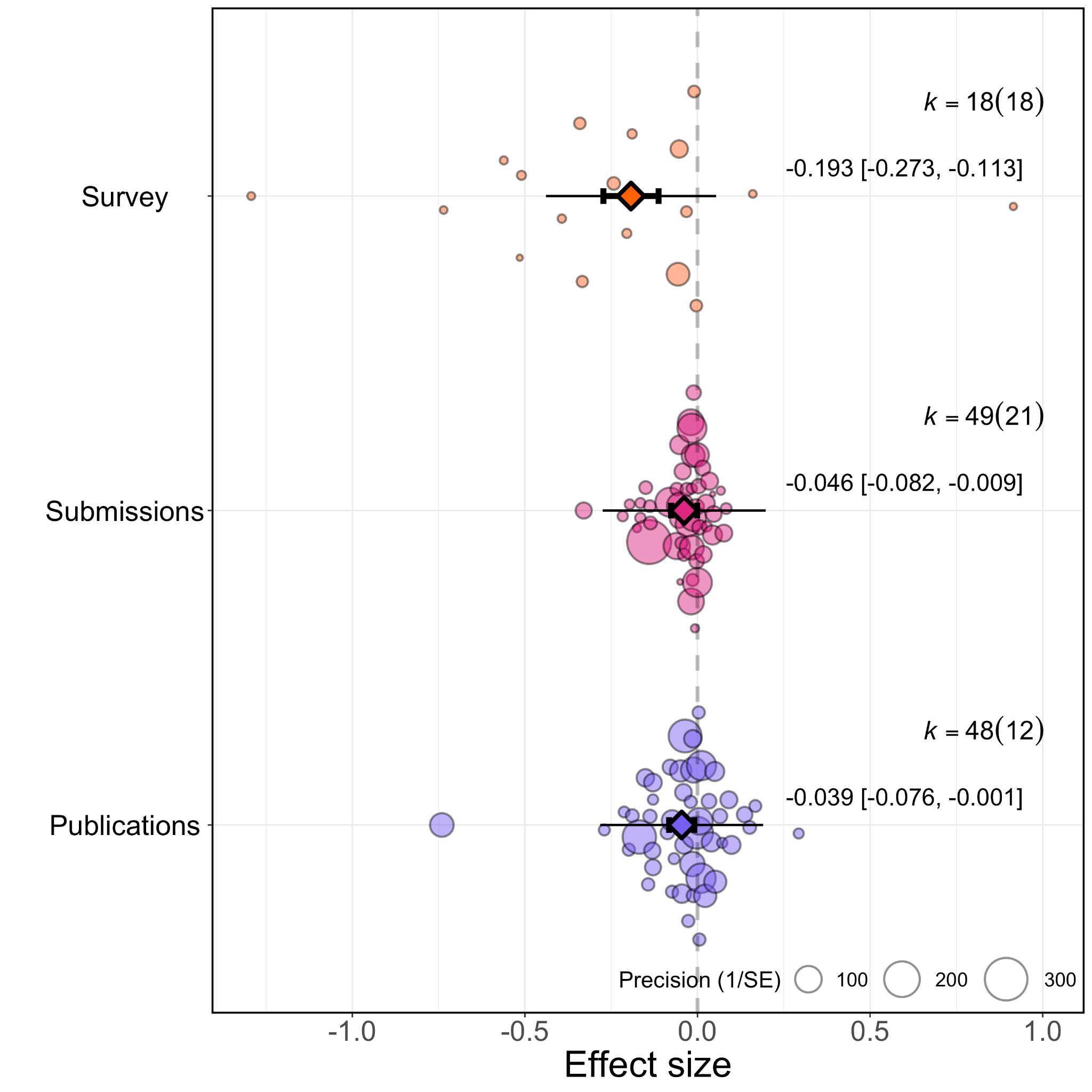


Figure 2. Orchard plots comparing the distribution of effect sizes (points) and their weight (point size) depending on the type of study (publication studies in green and self-reported studies in yellow), for which the mean effect size (darker coloured points outlined black and vertically centred), the 95% confidence interval (horizontal thick black bar) and the 95% prediction interval of the expected spread of effect sizes based on between-study variance (horizontal thin black bar) is given. k =x(y), where x is the number of effect sizes and y is the number of studies.

### 2a: Has the pandemic affected women differently across research fields?

We found little evidence of a significant differential impact of research fields on the reported effect sizes (QM (df=4)= 6.341, p-value= 0.175, Fig. 3). When considering research fields individually, social sciences showed the greatest increases in the academic productivity gender gap during the pandemic (-0.079, 95% CI= -0.141 to -0.018, SE= 0.031, p-value= 0.011), followed by multidisciplinary fields (-0.058, 95% CI= -0.139 to 0.023, SE= 0.041, p-value= 0.164), and then medicine (-0.051, 95% CI= -0.090 to -0.012, SE= 0.020, p-value= 0.011). The pandemic showed little effect in biological sciences (-0.004, 95% CI= -0.-0.059 to -0.050, SE= 0.028, p-value= 0.874) or Technology, Engineering, Mathematics, Chemistry and Physics fields (0.006, 95% CI= -0.048 to 0.060, SE= 0.028, p-value= 0.827).

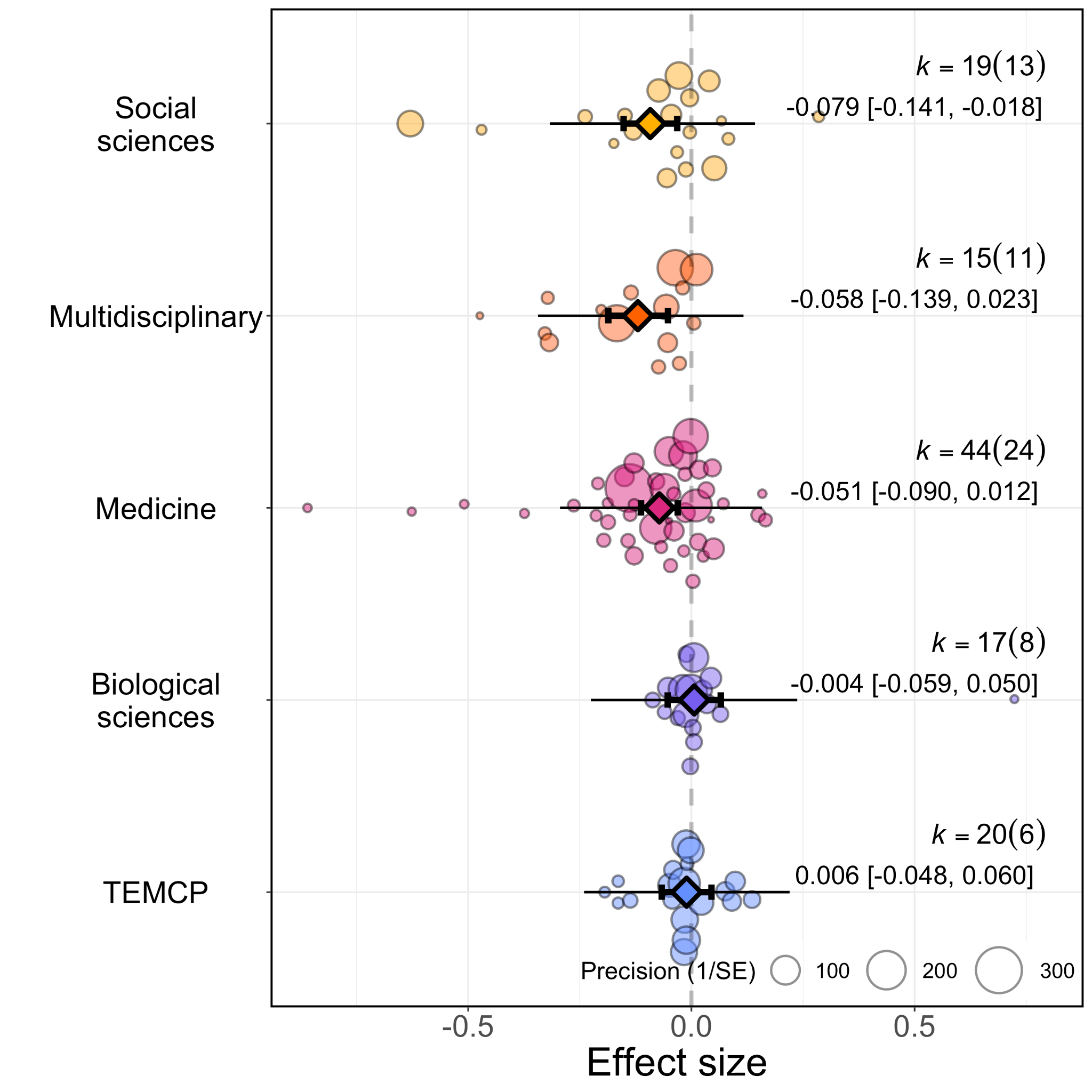


Figure 3. Orchard plot comparing the distribution of effect sizes (points) and their weights (point sizes) depending on the research fields sampled, for which the mean effect size (darker coloured points outlined black and vertically centred), the 95% confidence interval (horizontal thick black bar) and the 95% prediction interval of the expected spread of effect sizes based on between-study variance (horizontal thin black bar) is given. TEMCP corresponds to Technology, Engineering, Mathematics, Chemistry and Physics. k =x(y), where x is the number of effect sizes and y is the number of studies.

### 2b: Has the pandemic exacerbated existing differences in gender biases across research fields?

The pandemic has increased the gender gap in article output more in journals/repositories/pre-print servers that were previously less gender-biased (QM(df = 1) = 11.0156, p-value= 0.0009). When grouping studies by research fields (Fig. 4), those with a smaller gender gap prior to the pandemic experienced greater increases in the gender gap in academic productivity during the pandemic compared with fields where the gender gap was already large to start with (Social sciences: 0.377% to 0.358%, medicine: 0.369% to 0.346%, multidisciplinary: 0.392% to 0.367%, biological sciences: 0.328% to 0.328%, Technology, Engineering, Mathematics, Chemistry and Physics: 0.222% to 0.222%).

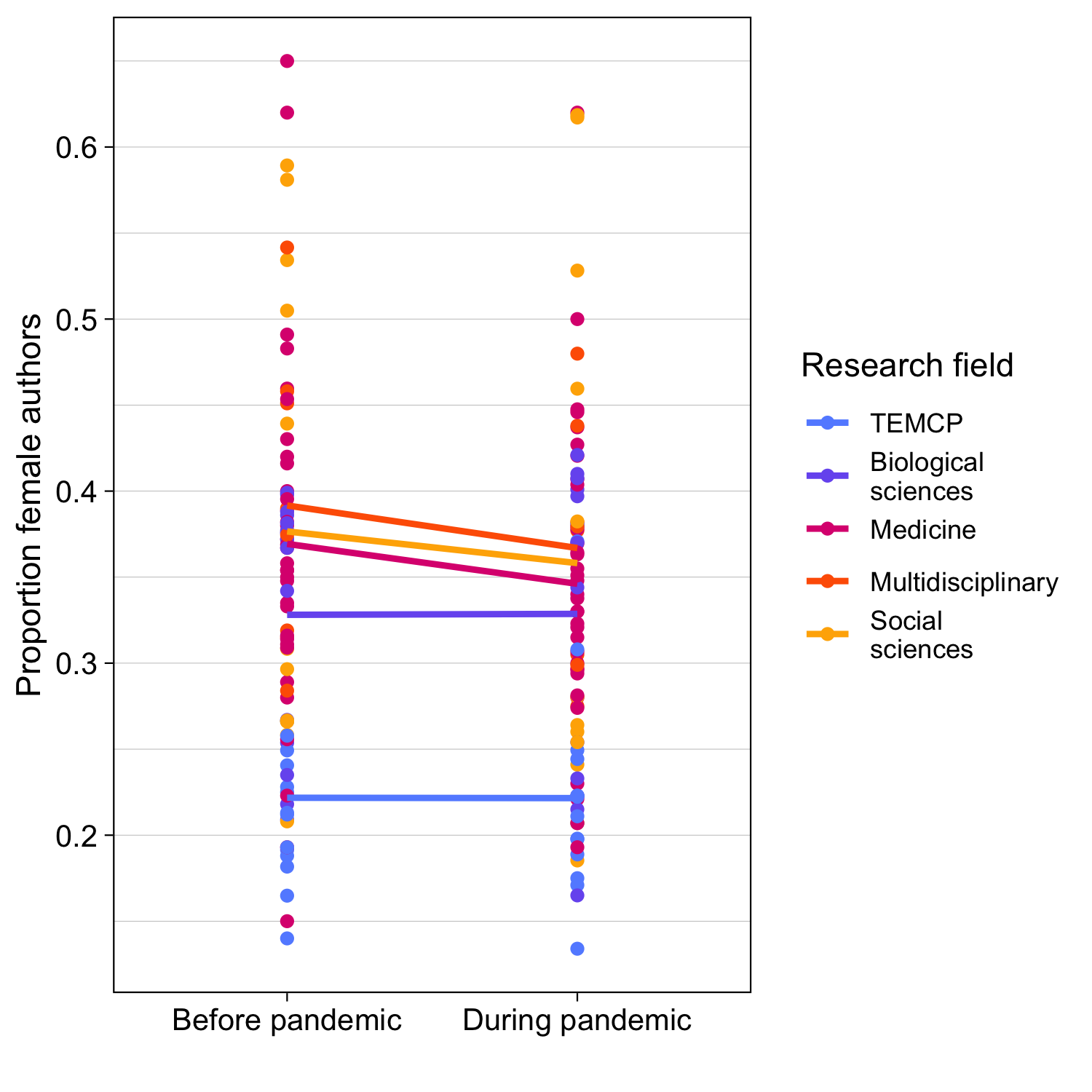


Figure 4. Line plot of gender gap in authorship before and during the pandemic, grouped by research field. Points show gender gap as the proportion of female authors publishing and submitting before or during the pandemic and are coloured according to research field. Lines take the mean value of these points according to research field. TEMCP corresponds to Technology, Engineering, Mathematics, Chemistry and Physics.

### 

### 3a: Has the pandemic affected women more in their ability to lead rather than support research?

We found no evidence of a significant differential impact of authorship position on effect sizes (QM(df = 5) = 8.133, p-value = 0.149, Fig. 5).

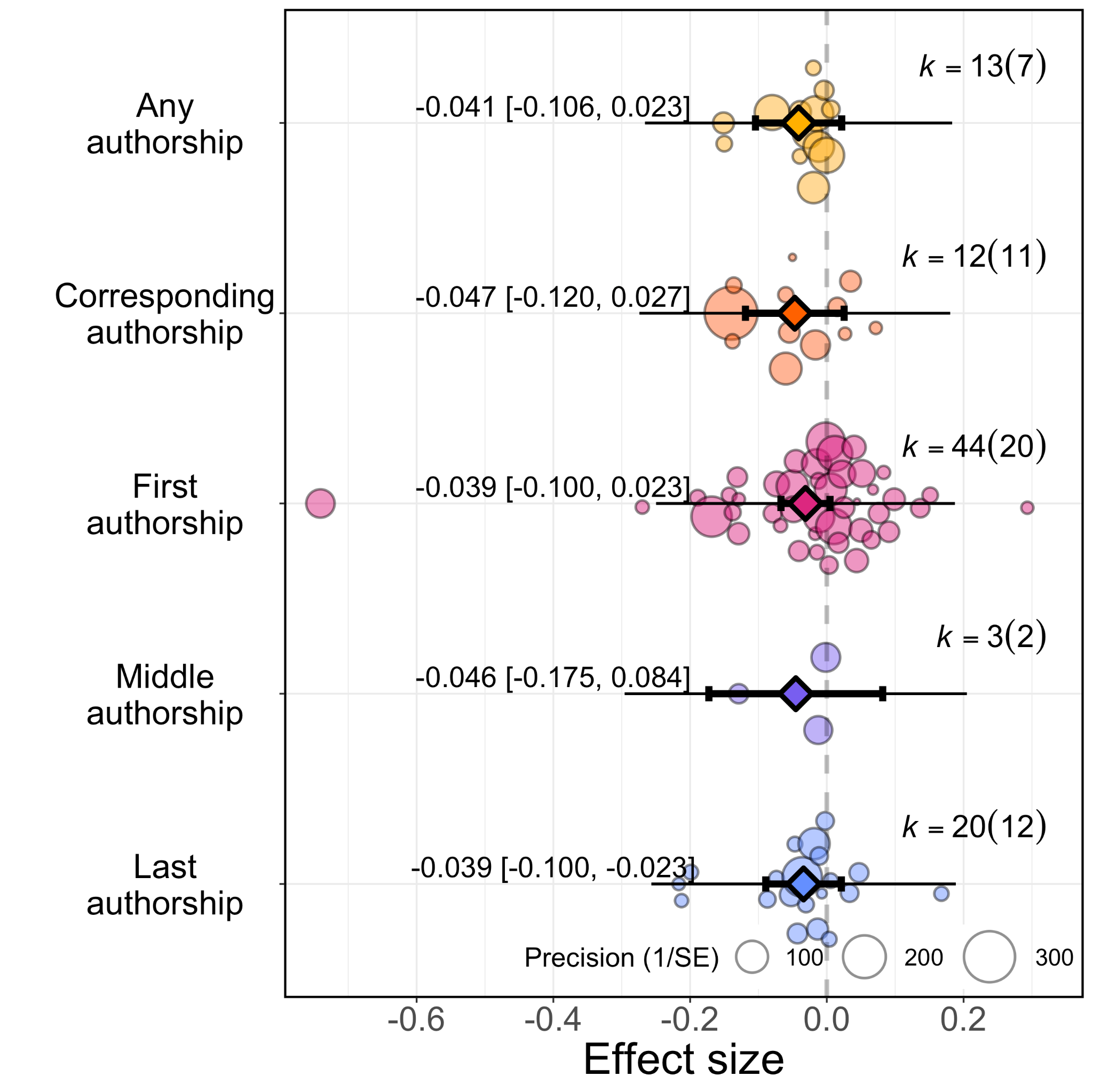
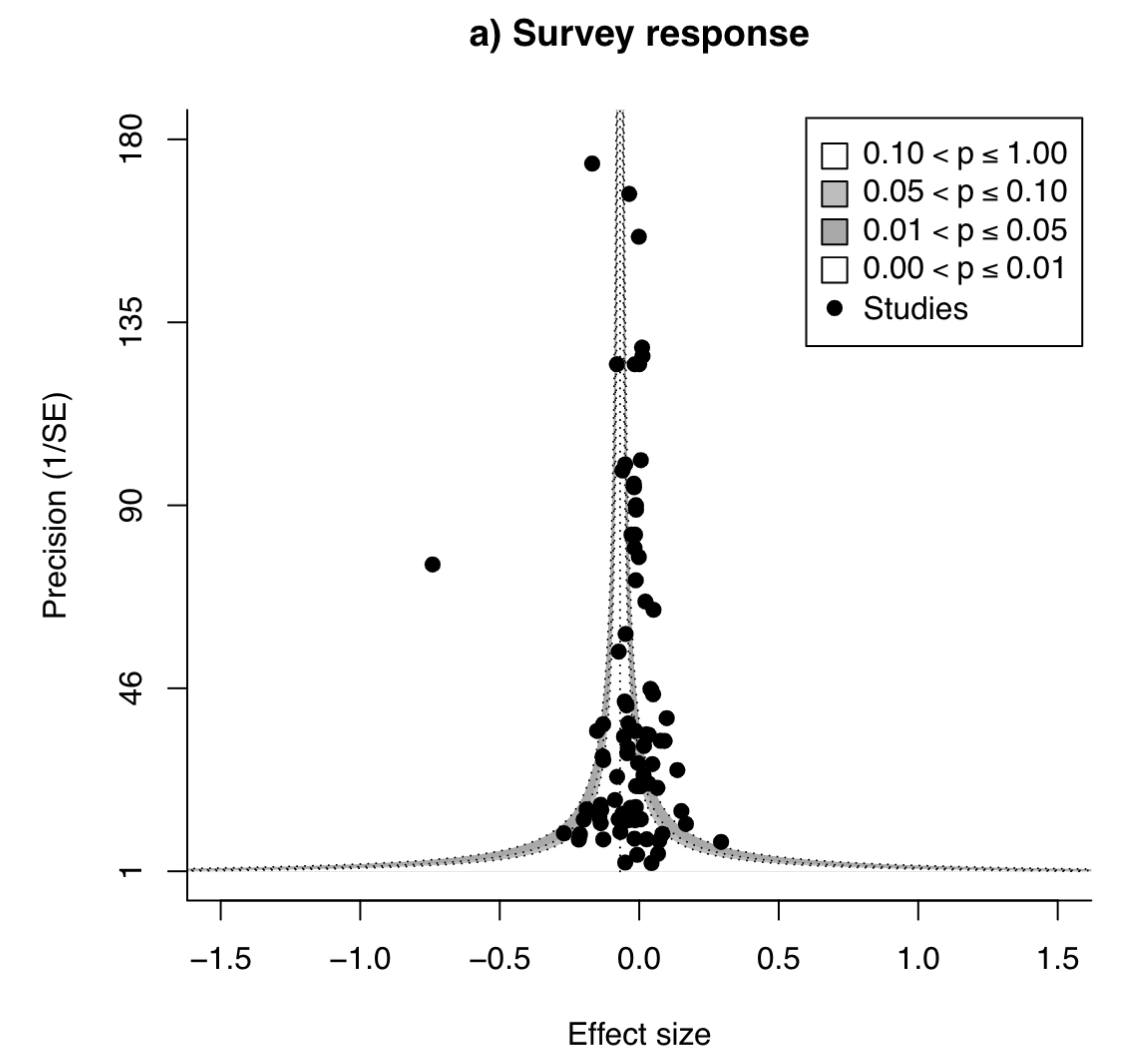


Figure 5. Orchard plot comparing the distribution of effect sizes (points) and their weight (point sizes), depending on the authorship position sampled in publication studies, for which the mean effect size (darker coloured points outlined black and vertically centred), the 95% confidence interval (horizontal thick black bar) and the 95% prediction interval of the expected spread of effect sizes based on between-study variance (horizontal thin black bar) is given. k =x(y), where x is the number of effect sizes and y is the number of studies.

### Is there evidence of publication bias?

We found no evidence of publication bias based on our multilevel meta-regression, suggesting small studies with large effect sizes did not skew our model (Article-output studies: slope= -0.026, 95% CI= [-0.061- 0.009], SE=0.018, p=0.140; Survey-reponse studies: slope= -0.190, 95% CI= [-0.286- -0.095], SE=0.049, p<0.001). A visual inspection of the funnel plots (Fig. 6) similarly did not indicate any suggestion of publication bias.



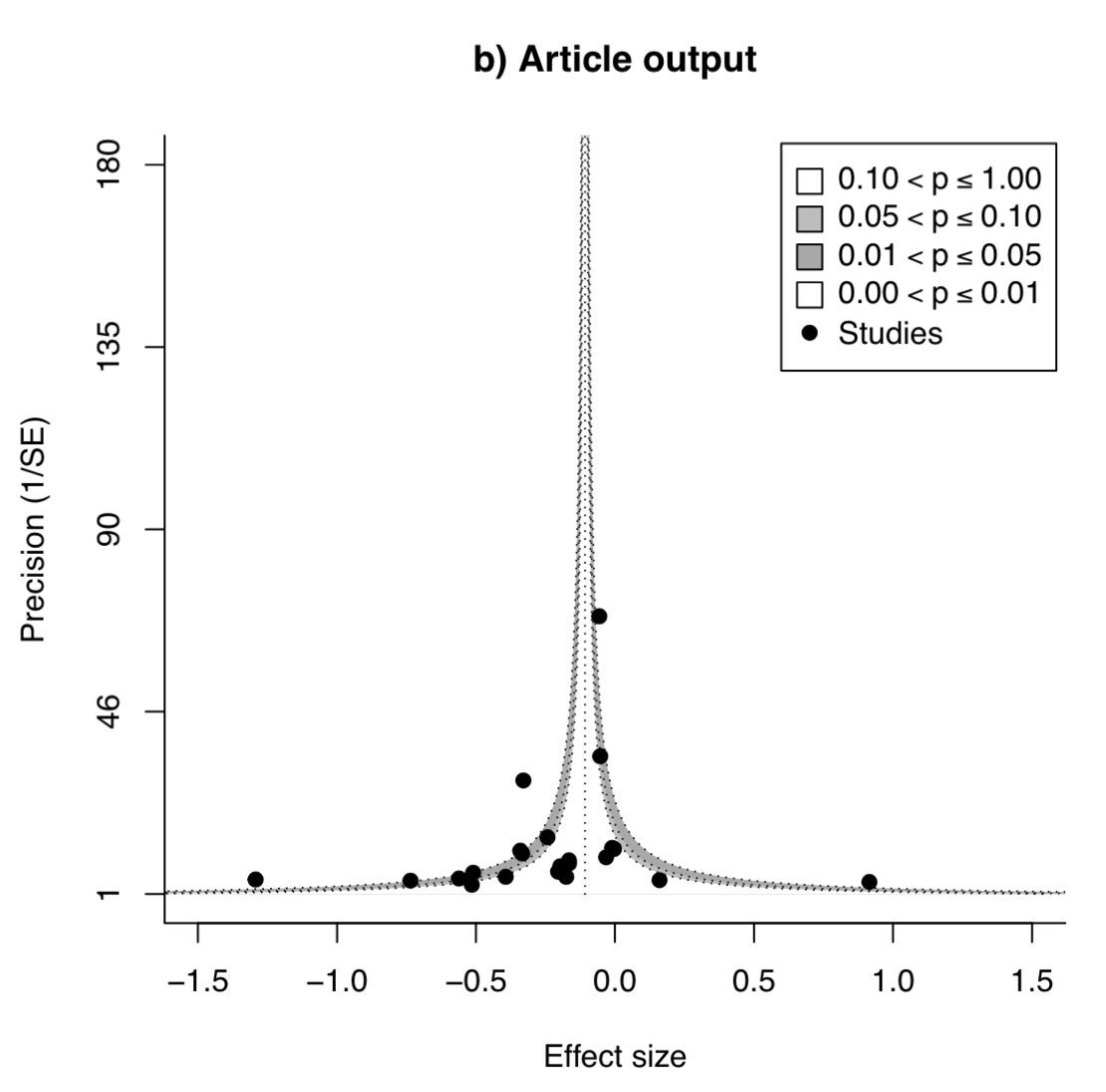


Figure 6. Funnel plot of effect sizes and their precision as a function of standard error from studies that measure research productivity by (a) survey-response (b) article output. The vertical dashed line is the summary effect size. The legend outlines levels of statistical significance for effect sizes based on their precision.

# Discussion

Our study finds quantitative evidence, based on 50 studies and 115 effect sizes, to support the hypothesis that the COVID-19 pandemic has exacerbated gender gaps in academic productivity. These findings are consistent with the notion that novel social conditions induced by the pandemic have disadvantaged women in academia even more than before. Overall, the studies summarised in our meta-analysis suggest that the gender gap in research productivity within academia has increased during the pandemic by 7%.We found no evidence of a publication bias.There is high heterogeneity in the effect sizes reported from different studies, arising from the type of research productivity measured. When measuring research productivity as the number of published or submitted articles, we find a slightly smaller increase in the gap of around 4%. This corresponds to the proportion of authors on submitted/published articles who are women declining from an average of 34.0% pre-pandemic to 32.6% during the pandemic (-0.04\*34.0%=-1.4%). Such a change might reflect lower submission and acceptance rate of articles by women compared to their male colleagues or an increased drop-out of woman from academia caused by the pandemic.

Our study likely underestimates the pandemic effect on article productivity in women because writing and publishing can take a long time (Powell, 2016). Many of the articles submitted or published during the pandemic were likely started and at least partially completed prior to the pandemic, given that most research grants span multiple years. With restricted access to laboratories, field sites and collaborators, many new projects have been delayed (Corbera *et al.*, 2020), making it likely that the article-output studies we could include by the time our study started in 2021 underestimates the true effect of the pandemic, which might span over many years. In support of this view we find some indication for a larger, real-time effect from the effect sizes based on survey responses, which indicate a much stronger negative effect of the pandemic on women’s productivity compared to men’s (effect size = -0.19). This signals that women are one fifth more likely than men to indicate that the pandemic has negatively affected their academic activities, which may stem from a combination of women on average feeling a larger strain, and a larger proportion of women being severely affected by the pandemic. In the literature used within our meta-analysis, five of six survey studies report evidence of a negative interaction effect of being both female and a parent on research productivity during the pandemic, presumably because of increased caregiving demands. Effect sizes are highly varied in survey response studies, which may reflect subtle differences in the measure of research productivity asked in the survey or some populations of survey respondents having strong opinions of the pandemic.

Our analysis suggests the pandemic may have differentially impacted female researchers across research fields, with increases in gender gaps particularly visible in research fields that were nearest gender-equality before the pandemic. Social sciences and medicine were two fields closest to gender equality that experienced the most significant decrease in female authors. Female researchers working in fields with previously gender-equitable environments may have experienced new, difficult research conditions induced by the pandemic, whereas in gender-biased fields, these difficulties might already have been present. For example, financial structures or childcare arrangements that could previously assist parents with caregiving responsibilities may have broken down during the pandemic (Fortin and Taylor, 2020; National Academies of Sciences, 2021). Alternatively, social sciences and medicine are fields that could have had the greatest surge in COVID-19 and pandemic-related research. Women in social sciences and medicine potentially had less opportunities to pursue this new pandemic-related research because of extra work performed in gender roles, or because women already had relatively smaller collaborative networks, fewer senior positions and less funding. Additionally, many medical journals sped up the publication process (Horbach, 2020), so the real-time effect of the pandemic on research productivity in women versus men may be reflected more in the biases in papers submitted and published in medicine than in other fields.

We did not find a clear signal that biases in research productivity differed according to authorship positions on submitted/published articles. We were limited in addressing this prediction because the samples used for calculating effect sizes for particular authorship positions were too small to infer whether there was a differential change in the gender gap.

Our study cannot identify the causes of the increased biases in research productivity during the pandemic. It seems unlikely that this increase in the gender gap simply represents a normal temporal fluctuation. The survey results, which report the strongest effects, specifically focused on the influence of the pandemic above and beyond the pressures researchers might already normally experience. The 4% decline in the proportion of authors who are women also likely indicates the extraordinary circumstances of the pandemic. This decline is remarkable given that a study comparing the change in the proportion of female authors between 1945 and 2005 showed a steady increase from 14% of all authors being women to 35%, with no apparent year-on-year decline since at least 1990 (Huang et al. 2020). As discussed above, the patterns likely reflect a combination of many individuals reducing their productivity rate as well as particularly affected individuals dropping out of academia, with the potential for longer-term effects in the coming years.

# Conclusion

Overall, our study highlights an exacerbated gender gap in academic research productivity during the COVID-19 pandemic. This gender gap was exacerbated more in social sciences and medicine, which are fields that were previously less gender-biased and may represent regression in progress made towards gender equality. Academic institutions should acknowledge and carefully accommodate the pandemic period when using research productivity to evaluate female academics for career progression in the coming years. Measures designed to reduce the gender gap in research may inadvertently exacerbate the gap by extending the period that advantaged individuals can outperform. At the broader level, the pandemic presents one social circumstance of many, including class, ethnicity, nationality, religion, disabilities that can interactively compound individual research productivity. More support should be given to academics disadvantaged by social circumstance, and those historically under-represented in academia. Simultaneously, more emphasis could be placed evaluating academic merit using more holistic measures and on an individual basis.

# Acknowledgements

We express gratitude for papers not behind paywalls.

# Additional information and declarations

### Funding

The authors received no funding for this work.

### Competing interests

The authors declare there are no competing interests.

### Author contributions

Adele Mennerat (AM) was involved in conceptualization, data extraction, methodology, analysis and reviewing and editing the manuscript draft.

Alecia J Carter (AJC) was involved in conceptualization, data extraction, methodology, analysis and reviewing and editing the manuscript draft.

Antica Culina (AC) was involved in conceptualization, data extraction, methodology, analysis and reviewing and editing the manuscript draft.

Dieter Lukas (DL) was involved in conceptualization, data extraction, methodology, analysis and reviewing and editing the manuscript draft.

Hannah Dugdale (HD) was involved in conceptualization, methodology, analysis and reviewing and editing the manuscript draft.

Kiran Lee (KL) was involved in conceptualization, data extraction, methodology, analysis, writing the first draft and reviewing and editing the draft.

### Data availability

Data and materials to reproduce the meta-analysis can be found at Zenodo.

### Supplementary information

Supplemental information is currently at the end of this document but can be separated if requested.

# References

Abdellatif, A. and Gatto, M. (2020) ‘It’s OK not to be OK: Shared reflections from two PhD parents in a time of pandemic’, *Gender, Work & Organization*, 27(5), pp. 723–733.

Addessi, E., Borgi, M. and Palagi, E. (2012) ‘Is Primatology an Equal-Opportunity Discipline?’, *PLOS ONE*, 7(1), p. e30458. Available at: https://doi.org/10.1371/journal.pone.0030458.

Altan-Olcay, Ö. and Bergeron, S. (2022) ‘Care in times of the pandemic: Rethinking work meanings of work in the university’, *Gender, Work & Organization* [Preprint].

Amano-Patiño, N. *et al.* (2020) ‘Who is doing new research in the time of COVID-19? Not the female economists’, *Publishing and Measuring Success in Economics*, 13.

Anabaraonye, N. *et al.* (2022) ‘Impact of the early COVID-19 pandemic on gender participation in academic publishing in radiation oncology’, *Advances in Radiation Oncology*, 7(2), p. 100845.

Andersen, J.P. *et al.* (2020) ‘COVID-19 medical papers have fewer women first authors than expected’, *elife*, 9.

Anwer, M. (2020) ‘Academic labor and the global pandemic: Revisiting life-work balance under COVID-19’, *Susan Bulkeley Butler Center for leadership excellence and advance working paper series*, 3(1), pp. 5–13.

Armstrong, M.A. and Jovanovic, J. (2015) ‘Starting at the crossroads: Intersectional approaches to institutionally supporting underrepresented minority women STEM faculty’, *Journal of Women and Minorities in Science and Engineering*, 21(2).

Astegiano, J., Sebastián-González, E. and Castanho, C.D.T. (2019) ‘Unravelling the gender productivity gap in science: a meta-analytical review’, *Royal Society Open Science*, 6(6). Available at: https://doi.org/10.1098/RSOS.181566.

Ayyala, R.S. and Trout, A.T. (2022) ‘Gender trends in authorship of Pediatric Radiology publications and impact of the COVID-19 pandemic’, *Pediatric Radiology*, 52(5), pp. 868–873.

Babcock, L. *et al.* (2017) ‘Gender differences in accepting and receiving requests for tasks with low promotability’, *American Economic Review*, 107(3), pp. 714–47.

Babcock, L. *et al.* (2022) ‘Saying “no” in science isn’t enough’, *Nature* [Preprint]. Available at: https://doi.org/10.1038/d41586-022-03677-6.

Barber, B.M. *et al.* (2021) ‘What explains differences in finance research productivity during the pandemic?’, *The Journal of Finance*, 76(4), pp. 1655–1697.

Bell, K. and Green, J. (2020) ‘Premature evaluation? Some cautionary thoughts on global pandemics and scholarly publishing’, *Critical Public Health*. Taylor & Francis.

Bell, M.L. and 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), pp. 159–163.

Biondi, B. *et al.* (2021) ‘Journal submissions, review and editorial decision patterns during initial COVID-19 restrictions’, *Food Policy*, 105, p. 102167.

Bol, T., Vaan, M.D. and Rijt, A.V.D. (2018) ‘The Matthew effect in science funding’, *Proceedings of the National Academy of Sciences of the United States of America*, 115(19), pp. 4887–4890. Available at: https://doi.org/10.1073/PNAS.1719557115/-/DCSUPPLEMENTAL.

Boncori, I. (2020) ‘The Never-ending Shift: A feminist reflection on living and organizing academic lives during the coronavirus pandemic’, *Gender, Work & Organization*, 27(5), pp. 677–682.

Breuning, M. *et al.* (2021) ‘The great equalizer? Gender, parenting, and scholarly productivity during the global pandemic’, *PS: Political Science & Politics*, 54(3), pp. 427–431.

Carli, L.L. (2020) ‘Women, Gender equality and COVID-19’, *Gender in Management: An International Journal* [Preprint].

Carlson, D.L., Petts, R.J. and Pepin, J.R. (2020) ‘US Couples’ Divisions of Housework and Childcare during COVID-19 Pandemic’. Available at: https://doi.org/10.31235/osf.io/jy8fn.

Chemistry, R.S. of (2019) ‘Is publishing in the chemical sciences gender biased?’ Royal Society of Chemistry Cambridge, UK.

Chen, T.-H.K. and Seto, K.C. (2022) ‘Gender and authorship patterns in urban land science’, *Journal of Land Use Science*, 17(1), pp. 245–261.

Clark, D. (2020) ‘Reflections on institutional equity for faculty in response to COVID-19’, *Susan Bulkeley Butler Center for Leadership Excellence and ADVANCE Working Paper Series*, 3(2).

Cook, J. *et al.* (2021) ‘Gender differences in authorship of obstetrics and gynecology publications during the coronavirus disease 2019 pandemic’, *American journal of obstetrics & gynecology MFM*, 3(1).

Corbera, E. *et al.* (2020) ‘Academia in the Time of COVID-19: Towards an Ethics of Care’, *Planning Theory and Practice*, 21(2), pp. 191–199. Available at: https://doi.org/10.1080/14649357.2020.1757891.

Cui, R., Ding, H. and Zhu, F. (2022) ‘Gender inequality in research productivity during the COVID-19 pandemic’, *Manufacturing & Service Operations Management*, 24(2), pp. 707–726.

Cushman, M. (2020) ‘Gender gap in women authors is not worse during COVID-19 pandemic: Results from Research and Practice in Thrombosis and Haemostasis’, *Research and Practice in Thrombosis and Haemostasis*, 4(5), p. 672.

Davis, J.C. *et al.* (2022) ‘Are we failing female and racialized academics? A Canadian national survey examining the impacts of the COVID-19 pandemic on tenure and tenure-track faculty’, *Gender, Work & Organization*, 29(3), pp. 703–722.

DeFilippis, E.M. *et al.* (2021) ‘Gender differences in publication authorship during COVID-19: a bibliometric analysis of high-impact cardiology journals’, *Journal of the American Heart Association*, 10(5), p. e019005.

Deryugina, T., Shurchkov, O. and Stearns, J.E. (2021) ‘COVID-19 Disruptions Disproportionately Affect Female Academics’.

Diaz, J.O. *et al.* (2021) ‘Burnout syndrome in pediatric urology: A perspective during the COVID-19 pandemic—Ibero-American survey’, *Journal of Pediatric Urology*, 17(3), pp. 402-e1.

Docka-Filipek, D. and Stone, L.B. (2021) ‘Twice a “housewife”: On academic precarity,“hysterical” women, faculty mental health, and service as gendered care work for the “university family” in pandemic times’, *Gender, Work & Organization*, 28(6), pp. 2158–2179.

Ellinas, E.H. *et al.* (2021) ‘Winners and Losers in Academic Productivity During the COVID-19 Pandemic: Is the Gender Gap Widening for Faculty?’, *Journal of women’s health (2002)* [Preprint]. Available at: https://doi.org/10.1089/jwh.2021.0321.

Else, H. (2020) ‘How a torrent of COVID science changed research publishing–in seven charts’, *Nature*, 588(7839), pp. 553–554.

Excess Beth [@El\_Dritch] (2020) ‘Negligible number of submissions to the journal from women in the last month. Never seen anything like it.’, *Twitter*. Available at: https://twitter.com/El\_Dritch/status/1251469394582089731 (Accessed: 28 August 2022).

Fazackerley, A. (2020) ‘Women’s research plummets during lockdown - but articles from men increase’, *The Guardian*, 12 May. Available at: https://www.theguardian.com/education/2020/may/12/womens-research-plummets-during-lockdown-but-articles-from-men-increase (Accessed: 28 August 2022).

Foo, Y.Z. *et al.* (2021) ‘A practical guide to question formation, systematic searching and study screening for literature reviews in ecology and evolution’, *Methods in Ecology and Evolution*, 12(9), pp. 1705–1720. Available at: https://doi.org/10.1111/2041-210X.13654.

Forti, L.R., Solino, L.A. and Szabo, J.K. (2021) ‘Trade-off between urgency and reduced editorial capacity affect publication speed in ecological and medical journals during 2020’, *Humanities and Social Sciences Communications*, 8(1), pp. 1–9.

Fortin, J. and Taylor, D.B. (2020) ‘Florida State University Child Care Policy Draws Backlash’, *The New York Times*, 2 July. Available at: https://www.nytimes.com/2020/07/02/us/fsu-telecommute-remote.html (Accessed: 6 November 2022).

Fox, C.W. *et al.* (2016) ‘Gender differences in patterns of authorship do not affect peer review outcomes at an ecology journal’, *Functional Ecology*, 30(1), pp. 126–139. Available at: https://doi.org/10.1111/1365-2435.12587.

Fox, C.W. and Meyer, J. (2021) ‘The influence of the global COVID-19 pandemic on manuscript submissions and editor and reviewer performance at six ecology journals’, *Functional Ecology*. Wiley Online Library.

Fox, C.W. and Paine, C.T. (2019) ‘Gender differences in peer review outcomes and manuscript impact at six journals of ecology and evolution’, *Ecology and Evolution*, 9(6), pp. 3599–3619.

Gao, J. *et al.* (2021) ‘Potentially long-lasting effects of the pandemic on scientists’, *Nature communications*, 12(1), pp. 1–6.

Gayet-Ageron, A. *et al.* (2021) ‘Female authorship of covid-19 research in manuscripts submitted to 11 biomedical journals: cross sectional study’, *bmj*, 375.

Gerding, A.B. *et al.* (2021) ‘Scholarly productivity in clinical pharmacology amid pandemic-related workforce disruptions: are men and women affected equally?’, *Clinical Pharmacology and Therapeutics*, 110(4), p. 841.

Ghaffarizadeh, S Aryan *et al.* (2021) ‘Life and work of researchers trapped in the COVID-19 pandemic vicious cycle’, *bioRxiv* [Preprint].

Grames, E.M. *et al.* (2019) ‘An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks’, *Methods in Ecology and Evolution*, 10(10), pp. 1645–1654. Available at: https://doi.org/10.1111/2041-210X.13268.

Guarino, C.M. and Borden, V.M.H. (2017) ‘Faculty Service Loads and Gender: Are Women Taking Care of the Academic Family?’, *Research in Higher Education*, 58(6), pp. 672–694. Available at: https://doi.org/10.1007/s11162-017-9454-2.

Guintivano, J., Dick, D. and Bulik, C.M. (2021) ‘Psychiatric genomics research during the COVID-19 pandemic: A survey of Psychiatric Genomics Consortium researchers’, *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 186(1), pp. 40–49.

Guy, B. and Arthur, B. (2020) ‘Academic motherhood during COVID-19: Navigating our dual roles as educators and mothers’, *Gender, Work & Organization*, 27(5), pp. 887–899. Available at: https://doi.org/10.1111/gwao.12493.

Hagan, A.K. *et al.* (2020) ‘Women Are Underrepresented and Receive Differential Outcomes at ASM Journals: a Six-Year Retrospective Analysis’, *mBio*, 11(6), pp. e01680-20. Available at: https://doi.org/10.1128/mBio.01680-20.

Harris, B. *et al.* (2022) ‘Exploratory Investigation of Gender Differences in School Psychology Publishing Before and During the Initial Phase of COVID-19’, *Canadian Journal of School Psychology*, 37(2), pp. 204–211.

Herman, E. *et al.* (2021) ‘The impact of the pandemic on early career researchers: what we already know from the internationally published literature’, *Profesional de la Información*, 30(2).

Hoggarth, J.A. *et al.* (2021) ‘Impacts of the COVID-19 pandemic on women and early career archaeologists’, *Heritage*, 4(3), pp. 1681–1702.

Horbach, S.P.J.M. (2020) ‘Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19’, *Quantitative Science Studies*, 1(3), pp. 1056–1067. Available at: https://doi.org/10.1162/QSS\_A\_00076.

Huang, J. *et al.* (2020) ‘Historical comparison of gender inequality in scientific careers across countries and disciplines’, *Proceedings of the National Academy of Sciences*, 117(9), pp. 4609–4616. Available at: https://doi.org/10.1073/PNAS.1914221117.

Ipe, T.S. *et al.* (2021) ‘The impact of COVID-19 on academic productivity by female physicians and researchers in transfusion medicine’, *Transfusion*, 61(6), pp. 1690–1693.

James, A., Chisnall, R. and Plank, M.J. (2019) ‘Gender and societies: a grassroots approach to women in science’, *Royal Society open science*, 6(9). Available at: https://doi.org/10.1098/RSOS.190633.

Jemielniak, D., Sławska, A. and Wilamowski, M. (2021) ‘COVID-19 effect on the gender gap in academic publishing’, *Journal of Information Science*, p. 01655515211068168.

Kaatz, A., Gutierrez, B. and Carnes, M. (2014) ‘Threats to objectivity in peer review: the case of gender’, *Trends in pharmacological sciences*, 35(8), pp. 371–373.

King, M.M. and Frederickson, M.E. (2021) ‘The Pandemic Penalty: The Gendered Effects of COVID-19 on Scientific Productivity’:, *Socius* [Preprint]. Available at: https://doi.org/10.1177/23780231211006977.

Kozlowski, D. *et al.* (2022) ‘Intersectional inequalities in science’, *Proceedings of the National Academy of Sciences*, 119(2), p. e2113067119. Available at: https://doi.org/10.1073/pnas.2113067119.

Krukowski, R.A., Jagsi, R. and 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), pp. 341–347. Available at: https://doi.org/10.1089/jwh.2020.8710.

Landis, J. and Koch, G. (1977) ‘The measurement of observer agreement for categorical data.’, *Biometrics*, 33(1), pp. 159–174. Available at: https://doi.org/10.2307/2529310.

Langin, K. (2019) ‘Women of color face double dose of bias’, *Science (New York, N.Y.)*, 364(6444), pp. 921–922. Available at: https://doi.org/10.1126/science.364.6444.921.

Larivière, V. *et al.* (2013) ‘Bibliometrics: Global gender disparities in science’, *Nature*, 504(7479), pp. 211–213. Available at: https://doi.org/10.1038/504211a.

Lerchenmüller, C. *et al.* (2021) ‘Longitudinal analyses of gender differences in first authorship publications related to COVID-19’, *BMJ open*, 11(4), p. e045176.

Livoreil, B. *et al.* (2017) ‘Systematic searching for environmental evidence using multiple tools and sources’, *Environmental Evidence 2017 6:1*, 6(1), pp. 1–14. Available at: https://doi.org/10.1186/S13750-017-0099-6.

Macaluso, B. *et al.* (2016) ‘Is Science Built on the Shoulders of Women? A Study of Gender Differences in Contributorship’, *Academic Medicine*, 91(8), pp. 1136–1142. Available at: https://doi.org/10.1097/ACM.0000000000001261.

Maguire, R. *et al.* (2021) ‘Research interrupted: The impact of the COVID-19 pandemic on multiple sclerosis research in the field of rehabilitation and quality of life’, *Multiple Sclerosis Journal–Experimental, Translational and Clinical*, 7(3), p. 20552173211038030.

McHugh, M.L. (2012) ‘Interrater reliability: the kappa statistic’, *Biochemia Medica*, 22(3), p. 276.

Mihaljević, H. *et al.* (2019) ‘Reflections on Gender Analyses of Bibliographic Corpora’, *Frontiers in Big Data*, 2. Available at: https://www.frontiersin.org/articles/10.3389/fdata.2019.00029 (Accessed: 20 October 2022).

Minello, A., Martucci, S. and Manzo, L.K. (2021) ‘The pandemic and the academic mothers: present hardships and future perspectives’, *European Societies*, 23(sup1), pp. S82–S94.

Mitchell, S.M. and Hesli, V.L. (2013) ‘Women don’t ask? Women don’t say no? Bargaining and service in the political science profession’, *PS: Political Science & Politics*, 46(2), pp. 355–369.

Moher, D. *et al.* (2009) ‘Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement’, *PLOS Medicine*, 6(7), p. e1000097. Available at: https://doi.org/10.1371/JOURNAL.PMED.1000097.

Moore, M.T. and Griffin, B.W. (2006) ‘Identification of factors that influence authorship name placement and decisions to collaborate in peer-reviewed, education-related publications’, *Studies in Educational Evaluation*, 32(2), pp. 125–135. Available at: https://doi.org/10.1016/j.stueduc.2006.04.004.

Morgan, A.C. *et al.* (2022) ‘Socioeconomic roots of academic faculty’, *Nature Human Behaviour*, pp. 1–9. Available at: https://doi.org/10.1038/s41562-022-01425-4.

Muric, G. *et al.* (2021) ‘Gender disparity in the authorship of biomedical research publications during the COVID-19 pandemic: retrospective observational study’, *Journal of medical Internet research*, 23(4), p. e25379.

Murray, D. *et al.* (2019) ‘Author-reviewer homophily in peer review’, *BioRxiv*, p. 400515.

Myers, K.R. *et al.* (2020) ‘Unequal effects of the COVID-19 pandemic on scientists’, *Nature human behaviour*, 4(9), pp. 880–883.

Nakagawa, S., Lagisz, M., Jennions, M.D., *et al.* (2021) ‘Methods for testing publication bias in ecological and evolutionary meta‐analyses’, *Methods in Ecology and Evolution*, 13(1), pp. 4–21. Available at: https://doi.org/10.1111/2041-210x.13724.

Nakagawa, S., Lagisz, M., O’Dea, R.E., *et al.* (2021) ‘The orchard plot: Cultivating a forest plot for use in ecology, evolution, and beyond’, in *Research Synthesis Methods*. John Wiley and Sons Ltd, pp. 4–12. Available at: https://doi.org/10.1002/jrsm.1424.

Nash, M. and Churchill, B. (2020) ‘Caring during COVID-19: A gendered analysis of Australian university responses to managing remote working and caring responsibilities’, *Gender, Work & Organization*, 27(5), pp. 833–846.

National Academies of Sciences, E. (2021) *The Impact of COVID-19 on the Careers of Women in Academic Sciences, Engineering, and Medicine*. Available at: https://doi.org/10.17226/26061.

Nguyen, A.X. *et al.* (2021) ‘Impact of COVID-19 on longitudinal ophthalmology authorship gender trends’, *Graefe’s Archive for Clinical and Experimental Ophthalmology*, 259(3), pp. 733–744.

O’Meara, K., Kuvaeva, A. and Nyunt, G. (2017) ‘Constrained choices: A view of campus service inequality from annual faculty reports’, *The Journal of Higher Education*, 88(5), pp. 672–700.

O’Meara, K.A. *et al.* (2017) ‘Asked More Often: Gender Differences in Faculty Workload in Research Universities and the Work Interactions That Shape Them’, *American Educational Research Journal*, 54(6), pp. 1154–1186. Available at: https://doi.org/10.3102/0002831217716767.

Ouzzani, M. *et al.* (2016) ‘Rayyan—a web and mobile app for systematic reviews’, *Systematic Reviews 2016 5:1*, 5(1), pp. 1–10. Available at: https://doi.org/10.1186/S13643-016-0384-4.

Plaunova, A. *et al.* (2021) ‘Impact of COVID-19 on radiology faculty-an exacerbation of gender differences in unpaid home duties and professional productivity’, *Academic Radiology*, 28(9), pp. 1185–1190.

Porter, S.R. (2007) ‘A closer look at faculty service: What affects participation on committees?’, *The Journal of Higher Education*, 78(5), pp. 523–541.

Powell, K. (2016) ‘Does it take too long to publish research?’, *Nature*, 530(7589), pp. 148–151. Available at: https://doi.org/10.1038/530148a.

Quak, E. *et al.* (2021) ‘Author Gender Inequality in Medical Imaging Journals and the COVID-19 Pandemic’, *Radiology*, 300(1), pp. E301–E307. Available at: https://doi.org/10.1148/radiol.2021204417.

R Core Team (2022) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Reuben, E., Sapienza, P. and Zingales, L. (2014) ‘How stereotypes impair women’s careers in science’, *Proceedings of the National Academy of Sciences*, 111(12), pp. 4403–4408.

Ribarovska, A.K. *et al.* (2021) ‘Gender inequality in publishing during the COVID-19 pandemic’, *Brain, behavior, and immunity*, 91, p. 1.

Rodríguez-Rivero, R. *et al.* (2020) ‘Is it time for a revolution in work–life balance? Reflections from Spain’, *Sustainability*, 12(22), p. 9563.

Romano, N. (2021) ‘Assessing the gender gap in academia’, *Nature Italy 2021* [Preprint]. Available at: https://www.nature.com/articles/d43978-021-00037-2 (Accessed: 21 August 2022).

Safdar, B. *et al.* (2021) ‘Gender Disparity in Grants and Awards at the National Institute of Health’, *Cureus*, 13(4). Available at: https://doi.org/10.7759/CUREUS.14644.

Schiebinger, L. and Gilmartin, S.K. (2010) ‘Housework is an academic issue’, *Academe*, 96(1), pp. 39–44.

Schiebinger, L.L., Henderson, A.D. and Gilmartin, S.K. (2008) *Dual-career academic couples: What universities need to know*. Michelle R. Clayman institute for gender research, Stanford University.

Shalaby, M., Allam, N. and Buttorff, G.J. (2021) ‘Leveling the field: Gender inequity in academia during COVID-19’, *PS: Political Science & Politics*, 54(4), pp. 661–667.

Shen, H. (2013) ‘Inequality quantified: Mind the gender gap’, *Nature*, 495(7439), pp. 22–24. Available at: https://doi.org/10.1038/495022A.

Squazzoni, F. *et al.* (2021) ‘Gender gap in journal submissions and peer review during the first wave of the COVID-19 pandemic. A study on 2329 Elsevier journals’, *PloS one*, 16(10), p. e0257919.

Staniscuaski, F. *et al.* (2021) ‘Gender, race and parenthood impact academic productivity during the COVID-19 pandemic: from survey to action’, *Frontiers in psychology*, 12, p. 663252.

Stenson, M.C. *et al.* (2022) ‘Impact of COVID-19 on access to laboratories and human participants: exercise science faculty perspectives’, *Advances in Physiology Education*, 46(2), pp. 211–218.

Valentova, J.V. *et al.* (2017) ‘Underrepresentation of women in the senior levels of Brazilian science’, *PeerJ*, 5(12). Available at: https://doi.org/10.7717/PEERJ.4000.

Viechtbauer, W. (2010) ‘Conducting Meta-Analyses in R with the metafor Package’, *Journal of Statistical Software*, 36(3), pp. 1–48. Available at: https://doi.org/10.18637/JSS.V036.I03.

van der Wal, J.E.M., Thorogood, R. and Horrocks, N.P.C. (2021) ‘Collaboration enhances career progression in academic science, especially for female researchers’, *Proceedings of the Royal Society B: Biological Sciences*, 288(1958), p. 20210219. Available at: https://doi.org/10.1098/rspb.2021.0219.

Wehner, M.R., Li, Y. and 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), pp. e2020335–e2020335.

West, J.D. *et al.* (2013) ‘The Role of Gender in Scholarly Authorship’, *PLOS ONE*, 8(7), p. e66212. Available at: https://doi.org/10.1371/journal.pone.0066212.

Williams II, W.A. *et al.* (2021) ‘Impact of the coronavirus disease 2019 pandemic on authorship gender in the Journal of Pediatrics: disproportionate productivity by international male researchers’, *The Journal of Pediatrics*, 231, pp. 50–54.

Wilson, D.B. (2019) ‘Practical Meta-Analysis Effect Size Calculator [Online calculator’. Available at: https://campbellcollaboration.org/research-resources/effect-size-calculator.html (Accessed: 28 August 2022).

Wooden, P. and Hanson, B. (2022) ‘Effects of the COVID-19 Pandemic on Authors and Reviewers of American Geophysical Union Journals’, *Earth and Space Science*, 9(2), p. e2021EA002050.

Yildirim, T.M. and Eslen-Ziya, H. (2021) ‘The differential impact of COVID-19 on the work conditions of women and men academics during the lockdown’, *Gender, Work & Organization*, 28, pp. 243–249.

Zeng, X.H.T. *et al.* (2016) ‘Differences in Collaboration Patterns across Discipline, Career Stage, and Gender’, *PLOS Biology*, 14(11), p. e1002573. Available at: https://doi.org/10.1371/journal.pbio.1002573.

# Supplementary materials

Table S1. Naive Boolean search string used in initial literature search broken down into terms and their respective concept groups. Wildcards (\*) were used to return results containing different word endings. Texts were limited to those since 2020. Search was conducted on 10/07/2021.

|  |  |
| --- | --- |
| Concept group (PICO) | Terms |
| Academia (population) | ( academi\* OR author\* OR database\* OR journal\* OR research OR scien\* ) |
| Gender (population) | AND    ( female\* OR gender OR male\* OR men OR women ) |
| Pandemic (intervention) | AND    ( coronavirus OR covid OR pandemic ) |
| Inequality (comparator) | AND    ( bias\* OR disparit\* OR disproportion\* OR fewer OR gap OR "gender difference\*" OR imbalance\* OR inequalit\* OR inequit\* OR parity OR "sex difference\*" OR skew\* OR unequal ) |
| Productivity (outcome) | AND  ( performan\* OR publication\* OR publish\* OR productiv\* ) |
| Exclusion of biomedical studies (population) | AND NOT  ( surviv\* OR experiment OR laboratory ) |

Table S2. Top 60 strongest terms in the term co-occurrence matrix and their rank (ascending order). Terms in bold were included in the AND NOT operator concept group excluding biomedical studies.

|  |  |  |
| --- | --- | --- |
| Term | Strength | Rank ascending |
| child | 4257 | 2063 |
| regression | 4335 | 2064 |
| survey | 4367 | 2065 |
| ***acute respiratory*** | 4427 | 2066 |
| embase | 4637 | 2067 |
| databases | 4646 | 2068 |
| death | 4664 | 2069 |
| knowledge | 4764 | 2070 |
| hip | 4918 | 2071 |
| mortality | 5000 | 2072 |
| female | 5065 | 2073 |
| rights reserved | 5257 | 2074 |
| city | 5504 | 2075 |
| information | 5700 | 2076 |
| symptom | 6108 | 2077 |
| sars-cov-2 | 6112 | 2078 |
| sars-cov | 6185 | 2079 |
| coronavirus disease 2019 | 6397 | 2080 |
| systematic review | 6506 | 2081 |
| rest | 6578 | 2082 |
| science | 6659 | 2083 |
| distance | 6679 | 2084 |
| literature | 6816 | 2085 |
| database | 6916 | 2086 |
| medical | 7023 | 2087 |
| ***icu*** | 7170 | 2088 |
| affect | 7297 | 2089 |
| control | 7386 | 2090 |
| population | 7418 | 2091 |
| women | 7422 | 2092 |
| bias | 7665 | 2093 |
| gender | 7693 | 2094 |
| male | 7977 | 2095 |
| coronavirus disease | 8009 | 2096 |
| article | 8135 | 2097 |
| virus disease | 8230 | 2098 |
| ***iga*** | 8235 | 2099 |
| hospital | 8318 | 2100 |
| work | 8319 | 2101 |
| covid-19 pandemic | 8377 | 2102 |
| ***gis*** | 8489 | 2103 |
| severe | 8518 | 2104 |
| time | 8831 | 2105 |
| infection | 8914 | 2106 |
| ***rna*** | 10303 | 2107 |
| outcome | 10461 | 2108 |
| review | 11869 | 2109 |
| coronavirus | 12243 | 2110 |
| stem | 12316 | 2111 |
| ***risk*** | 12776 | 2112 |
| analysis | 13610 | 2113 |
| research | 13737 | 2114 |
| virus | 13835 | 2115 |
| pandemic | 14212 | 2116 |
| author | 14333 | 2117 |
| car | 15893 | 2118 |
| over | 15964 | 2119 |
| health | 16176 | 2120 |
| covid-19 | 21088 | 2121 |
| age | 21194 | 2122 |
| covid | 21330 | 2123 |

Table S3. Flowchart of questions for initial screen of title, abstract and key-words. Any 'No' answer meant that the article was excluded. 'Yes' AND/OR 'Maybe' answers to all of the below meant that the articles were included. Reports such as reviews or comments that may contain secondary data investigating our PICO framework were included. Articles that lack a formal abstract, but have a title or informal abstract suggesting at least 3 of the above questions can be answered with a ‘yes’.

|  |  |
| --- | --- |
| 1) Population: Is there suggestion the article investigates academia/ research/academics/journals? | Note: If the population is not narrowed to academia/research/academics/journals, but the study considers the effect on ‘work’ or ‘work productivity’ at a broad level, and looks like it has enough power to subdivide effects on different workplaces such as academia, then include.  Some academic fields such as those within medicine and engineering are dominated by a practical/applied/professional/labour component. Explicit reference to only investigating the professional/practical/labour context of these fields are excluded. If no reference is made to whether the academic field is investigated in a research or professional capacity, we include. |
| 2) Intervention: Is there suggestion the article investigates the effect of the pandemic? |  |
| 3) Comparison: Is there suggestion the article investigates pre-pandemic? | Note: Often titles and abstracts do not have dates and will not refer explicitly to pre-pandemic data being included. If there is suggestion the ‘effect of the pandemic’ (or similar) is considered, we are generous and assume a comparison with pre-pandemic data is included. |
| 4) Outcome: Is there suggestion the article investigates the effect on gender in productivity/publication/submission/authorship numbers? | Note for gender: If there is no reference to investigating the effect of gender or sex, but does refer to investigating broader sociodemographic or socioeconomic factors, then include because gender may be investigated in the full text.    Note for productivity: If there is no direct reference to publication or submission numbers, but does refer to, ‘productivity’ or ‘output’ in a research context, then include because publication or submission metrics may be included in the full text. |

Table S4. PRISMA flow diagram for number of records included at each stage in iterated search.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 10/07/2021 search | Web of Science | | Scopus | | EBSCO | | ProQuest |
|  | n=199 | | n=126 | | n=276 | | n=99 |
|  | |  | | **↓** | |  | |
|  | |  | | Total | |  | |
|  | |  | | n=700 | |  | |
|  | |  | | **↓** | |  | |
|  | |  | | De-duplicated | |  | |
|  | |  | | n=580 | |  | |
|  | |  | | **↓** | |  | |
|  | |  | | Title, abstract, keyword screened | |  | |
|  | |  | | n=81 | |  | |
|  | |  | | **↓** | |  | |
|  | |  | | Full text screened | |  | |
|  | |  | | n=25 | |  | |
|  |  | |  | |  | |  |
| 28/02/2022 search | Web of Science | | Scopus | | EBSCO | | ProQuest |
|  | n=413 | | n=258 | | n=542 | | n=433 |
|  | |  | | **↓** | |  | |
|  | |  | | Total | |  | |
|  | |  | | n=1646 | |  | |
|  | |  | | **↓** | |  | |
|  | |  | | De-duplicated | |  | |
|  | |  | | n=1208 | |  | |
|  | |  | | **↓** | |  | |
|  | |  | | Title, abstract, keyword screened | |  | |
|  | |  | | n=169 | |  | |
|  | |  | | **↓** | |  | |
|  | |  | | Full text screened | |  | |
|  | |  | | n=50 | |  | |

**R script:**

---

title: Investigating the impact of the COVID-19 pandemic on the gender productivity

gap in academia

author: "Kiran Gok Lune Lee"

date: "`r Sys.Date()`"

geometry: margin=2cm

output:

pdf\_document:

latex\_engine: xelatex

fig\_width: 8

fig\_height: 6

html\_document:

toc: yes

toc\_float:

collapsed: no

fig\_width: 12

fig\_height: 10

---

```{r setup, include=FALSE}

library(knitr)

knitr::opts\_chunk$set(echo = TRUE)

knitr::opts\_chunk$set(tidy.opts = list(width.cutoff = 60), tidy = TRUE)

```

## Load stuff

```{r eval = FALSE, echo = TRUE}

## install packages like so

install.packages("pacman")

rm(list = ls())

devtools::install\_github("daniel1noble/orchaRd", force = TRUE)

pacman::p\_load(devtools, tidyverse, metafor, patchwork, R.rsp, orchaRd, emmeans, ape, phytools, flextable)

```

```{r eval = TRUE, echo = TRUE}

library(metafor)

library(readxl)

library(ggplot2)

library(formatR)

library(devtools)

library(tidyverse)

library(patchwork)

library(orchaRd)

## Load data with all papers

all\_data<-read\_xlsx("all\_data.xlsx")

colnames(all\_data) <- make.names(colnames(all\_data), unique = TRUE)

### set effect sizes

all\_data$Effect.size.used.in.MA<-as.numeric(all\_data$Effect.size.used.in.MA)

all\_data$Variance.used<-as.numeric(all\_data$Variance.used)

all\_data$vi <- all\_data$Variance.used

all\_data$yi <- all\_data$Effect.size.used.in.MA

### data used for writing methods

table(all\_data$Effect.size.kiran.use..from.Campbell.collaboration.)

table(all\_data$Self.reported.or.measured)

table(all\_data$Broad.research.field)

table(all\_data$Broad.productivity.measure)

table(all\_data$Broad.specific.productivity)

table(all\_data$Specific.productivity.measure)

all\_data<-arrange(all\_data, Effect.size.used.in.MA) # Order the data by effect size ID for plots

all\_data$Effect.size.used.in.MA<-as.factor(all\_data$Effect.size.used.in.MA)

#custom orchard plots borrowed from https://github.com/p-pottier/Dev\_plasticity\_thermal\_tolerance/blob/main/Data\_analysis/R/Statistical\_analyses.Rmd

my.orchard<- function (object, mod = "1", group, data, xlab, N = "none",

alpha = 0.5, angle = 0, cb = FALSE, k = TRUE, g = TRUE,

trunk.size = 7, branch.size = 2, twig.size = 0.8, whisker, transfm = c("none", # increased point size, branch size, and added a whisker argument

"tanh"), condition.lab = "Condition", legend.pos = "bottom.right", k.pos = c("right",

"left"))

{

transfm <- match.arg(transfm)

if (any(class(object) %in% c("rma.mv", "rma"))) {

if (mod != "1") {

results <- orchaRd::mod\_results(object, mod, group,

data)

}

else {

results <- orchaRd::mod\_results(object, mod = "1",

group, data)

}

}

if (any(class(object) %in% c("orchard"))) {

results <- object

}

mod\_table <- results$mod\_table

data <- results$data

data$moderator <- factor(data$moderator, levels = mod\_table$name,

labels = mod\_table$name)

data$scale <- (1/sqrt(data[, "vi"]))

legend <- "Precision (1/SE)"

if (any(N != "none")) {

data$scale <- N

legend <- "Sample Size (N)"

}

if (transfm == "tanh") {

cols <- sapply(mod\_table, is.numeric)

mod\_table[, cols] <- Zr\_to\_r(mod\_table[, cols])

data$yi <- Zr\_to\_r(data$yi)

label <- xlab

}

else {

label <- xlab

}

mod\_table$K <- as.vector(by(data, data[, "moderator"],

function(x) length(x[, "yi"])))

mod\_table$g <- as.vector(num\_studies(data, moderator, stdy)[,

2])

group\_no <- length(unique(mod\_table[, "name"]))

cbpl <- c("#88CCEE", "#CC6677", "#DDCC77",

"#117733", "#332288", "#AA4499", "#44AA99",

"#999933", "#882255", "#661100", "#6699CC",

"#888888", "#E69F00", "#56B4E9", "#009E73",

"#F0E442", "#0072B2", "#D55E00", "#CC79A7",

"#999999")

if (names(mod\_table)[2] == "condition") {

condition\_no <- length(unique(mod\_table[, "condition"]))

plot <- ggplot2::ggplot() + ggbeeswarm::geom\_quasirandom(data = data,

ggplot2::aes(y = yi, x = moderator, size = scale,

colour = moderator), alpha = alpha) + ggplot2::geom\_hline(yintercept = 0,

linetype = 2, colour = "black", alpha = alpha) +

ggplot2::geom\_linerange(data = mod\_table, ggplot2::aes(x = name,

ymin = lowerCL, ymax = upperCL), size = branch.size,

position = ggplot2::position\_dodge2(width = 0.3)) +

ggplot2::geom\_pointrange(data = mod\_table, ggplot2::aes(y = estimate,

x = name, ymin = lowerPR, ymax = upperPR, shape = as.factor(condition),

fill = name), size = twig.size, position = ggplot2::position\_dodge2(width = 0.3),

fatten = trunk.size) + ggplot2::scale\_shape\_manual(values = 20 +

(1:condition\_no)) + ggplot2::coord\_flip() + ggplot2::theme\_bw() +

ggplot2::guides(fill = "none", colour = "none") +

ggplot2::theme(legend.position = c(0, 1), legend.justification = c(0,

1)) + ggplot2::theme(legend.title = ggplot2::element\_text(size = 9)) +

ggplot2::theme(legend.direction = "horizontal") +

ggplot2::theme(legend.background = ggplot2::element\_blank()) +

ggplot2::labs(y = label, x = "", size = legend) +

ggplot2::labs(shape = condition.lab) + ggplot2::theme(axis.text.y = ggplot2::element\_text(size = 10,

colour = "black", hjust = 0.5, angle = angle))

}

else {

plot <- ggplot2::ggplot() + ggbeeswarm::geom\_quasirandom(data = data,

ggplot2::aes(y = yi, x = moderator, size = scale,

fill = moderator), alpha = alpha, width=0.4, pch=21, stroke=1.1, col="black") + # Change point shape (21, with black borders)

ggplot2::geom\_hline(yintercept = 0,

linetype = 2, colour = "black", alpha = 0.3, lwd=1.3) + # Change thickness 0 line

ggplot2::geom\_errorbar(data = mod\_table, ggplot2::aes(x = name,

ymin = lowerCL, ymax = upperCL), size = branch.size, width= whisker) + # Added variable whisker size

ggplot2::geom\_pointrange(data = mod\_table, ggplot2::aes(y = estimate,

x = name, ymin = lowerPR, ymax = upperPR, fill = name),

size = twig.size, fatten = trunk.size, shape = 23, stroke=2.2) + # Change point shape

scale\_size\_continuous(range = c(1, 14))+ # change point scaling

ggplot2::coord\_flip() +

ggplot2::theme\_bw() +

ggplot2::guides(fill = "none", colour = "none") +

ggplot2::theme(text=element\_text(size=26, colour="black"))+ # Change font size

ggplot2::theme(legend.title = ggplot2::element\_text(size = 16)) + # Increased font legend title

ggplot2::theme(legend.text = ggplot2::element\_text(size = 14)) +

ggplot2::theme(legend.direction = "horizontal") +

ggplot2::theme(legend.background = ggplot2::element\_blank()) +

ggplot2::labs(y = label, x = "", size = legend) +

ggplot2::theme(axis.text.y = ggplot2::element\_text(size = 20,

colour = "black", hjust = 0.5, angle = angle)) + # Increased size title axis label

ggplot2::theme(axis.text.x = ggplot2::element\_text(size = 20)) + # Increase size axis ticks

ggplot2::theme(panel.border = element\_rect(colour = "black", fill=NA, size=1.3))

}

if (legend.pos == "bottom.right") {

plot <- plot + ggplot2::theme(legend.position = c(1,

0), legend.justification = c(1, 0))

}

else if (legend.pos == "bottom.left") {

plot <- plot + ggplot2::theme(legend.position = c(0,

0), legend.justification = c(0, 0))

}

else if (legend.pos == "top.right") {

plot <- plot + ggplot2::theme(legend.position = c(1,

1), legend.justification = c(1, 1))

}

else if (legend.pos == "top.left") {

plot <- plot + ggplot2::theme(legend.position = c(0,

1), legend.justification = c(0, 1))

}

else if (legend.pos == "top.out") {

plot <- plot + ggplot2::theme(legend.position = "top")

}

else if (legend.pos == "bottom.out") {

plot <- plot + ggplot2::theme(legend.position = "bottom")

}

if (cb == TRUE) {

plot <- plot + ggplot2::scale\_fill\_manual(values = cbpl) +

ggplot2::scale\_colour\_manual(values = cbpl)

}

if (k == TRUE && g == FALSE && k.pos == "right") {

plot <- plot + ggplot2::annotate("text", y = (max(data$yi) +

(max(data$yi) \* 0.1)), x = (seq(1, group\_no, 1) +

0.3), label = paste("italic(k)==", mod\_table$K[1:group\_no]), # Size changed to 5.5

parse = TRUE, hjust = "right", size = 6.5)

}

else if (k == TRUE && g == FALSE && k.pos == "left") {

plot <- plot + ggplot2::annotate("text", y = (min(data$yi) +

(min(data$yi) \* 0.1)), x = (seq(1, group\_no, 1) +

0.3), label = paste("italic(k)==", mod\_table$K[1:group\_no]),

parse = TRUE, hjust = "left", size = 6.5) # Size changed to 5.5

}

else if (k == TRUE && g == TRUE && k.pos == "right") {

plot <- plot + ggplot2::annotate("text", y = (max(data$yi) +

(max(data$yi) \* 0.1)), x = (seq(1, group\_no, 1) +

0.3), label = paste("italic(k)==", mod\_table$K[1:group\_no],

" (", mod\_table$g[1:group\_no], ")"),

parse = TRUE, hjust = "right", size = 6.5) # Size changed to 5.5

}

else if (k == TRUE && g == TRUE && k.pos == "left") {

plot <- plot + ggplot2::annotate("text", y = (min(data$yi) +

(min(data$yi) \* 0.1)), x = (seq(1, group\_no, 1) +

0.3), label = paste("italic(k)==", mod\_table$K[1:group\_no],

" (", mod\_table$g[1:group\_no], ")"),

parse = TRUE, hjust = "left", size = 6.5) # Size changed to 5.5

}

return(plot)

}

```

\newpage

## Hypothesis 1: During pandemic conditions the gender gap in academic productivity has increased.

Prediction 1a: The pandemic has increased the gender gap in productivity (as indicated by an overall negative effect size).

Hierarchical mixed effect meta-analysis with all papers.

Studies with multiple effect sizes are controlled for. Looks like an overall negative effect: during pandemic conditions the gender gap in academic productivity has increased. Forest plot produced for visual represenation.

However, meta-analysing all papers is probably not so useful considering productivity is measured in different ways (self-reported/submissions/publications)- see later section.

```{r eval = TRUE, echo = TRUE}

#Article ID refers to the article ID number the effect size comes from

all\_data$ID.article<-as.factor(all\_data$Article.ID)

#ID refers to the unique effect size ID number

all\_data$ID.observation<-as.factor(all\_data$ID)

m<-rma.mv(yi, vi, random=~1|ID.article/ID.observation,data=all\_data)

m

forest (m, slab=all\_data$Author, xlim=c(-2,2), ylim=c(-1, 120), digits= 2, xlab="Raw proportion", mlab="Overall effect (46)")

text(-1,122, "Author(s) and Year", pos=2, font=2, cex=0.8)

text(2,122, "Effect size [95% CI]", pos=2, font=2, cex=0.8)

```

```{r, fig.width=12, fig.height=12}

#Figure 1

my.orchard(m, mod="1", alpha=0.5,data = all\_data, whisker=0.025

, group = "Article.ID", xlab = "Effect size")+

annotate("text", size=6, y=0.50, x=1.06, label=paste("-0.070 [-0.102, -0.038]"))+

scale\_fill\_manual(values=c("#DC267F"))

ggsave("figure1.png", width=10, height=10, dpi=300)

```

### Investigating productivity measure as a moderator.

Prediction 1b: Studies that measure productivity as self reported detect stronger increase in gender gap than the studies that measure it as number of submissions, while the studies based on number of publications detect the weakest increase. This is because self-reporting productivty measured via questionnaire can measure the pandemics effects in real time. It can take a while to submit research and even longer to publish research, which therefore may not immediately be detected in numbers of submissions or publications.

```{r eval = TRUE, echo = TRUE}

all\_data$Broad.productivity.measure[all\_data$Broad.productivity.measure=="Other"]<-"Survey"

tapply(all\_data$ID.observation, all\_data$Broad.productivity.measure, length)

all\_data$Broad.productivity.measure.reordered<-factor(all\_data$Broad.productivity.measure, c("Publications", "Submissions", "Survey"))

m.area <- rma.mv(yi, vi, mods= ~ Broad.productivity.measure.reordered, random=~1|ID.article/ID.observation, data=all\_data)

m.area

m.area1 <- rma.mv(yi, vi, mods= ~ Broad.productivity.measure.reordered-1, random=~1|ID.article/ID.observation, data=all\_data)

m.area1

```

Orchard plot: Aggregated effect sizes for individual-based studies and group-based studies. Each dot is an effect size. Size of dot is precision in relation to variance (a function of sample size). Mean effect size for each group is the outlined dot with 95% confidence intervals as the thick horizontal black bar and "prediction intervals" (a range of plausible effect size values fo ra new study assuming an average sample size) as the thin horizontal bars. k=49 (21) means 49 effect sizes from 21 articles.

```{r eval = TRUE, echo = TRUE, fig.width=6, fig.height=6}

# Figure 2

p1 <- my.orchard(m.area1, mod="Broad.productivity.measure.reordered", group = "Article.ID", data=all\_data, xlab = "Effect size", whisker=0.05,transfm = "none")+

annotate("text", size=6, y=0.6, x=3.09, label=paste("-0.193 [-0.273, -0.113]"))+

annotate("text", size=6, y=0.6, x=2.09, label=paste("-0.046 [-0.082, -0.009]"))+

annotate("text", size=6, y=0.6, x=1.09, label=paste("-0.039 [-0.076, -0.001]"))+

scale\_fill\_manual(values=c("#785EF0","#DC267F","#FE6100"))

p1

ggsave("figure2.png")

```

## Funnel plot

Prediction 1b: We do not predict to find evidence of publication bias.

```{r eval = TRUE, echo = TRUE, fig.width=6, fig.height=6}

measured\_data<-subset(all\_data, all\_data$Self.reported.or.measured == "Measured")

meta\_measured <- rma.mv(yi=yi, V=vi, data=as.data.frame(measured\_data))

f1<-funnel(meta\_measured, level=c(90, 95, 99), shade=c("white", "gray", "darkgray"),

yaxis="seinv",xlab="Effect size",ylab="Precision (1/SE)",digits=c(1,0),xlim=c(-1.5,1.5),ylim=c(1,180), legend=TRUE,back="white", hlines = "white", main="a) Survey response")

f1

ggsave("figure3a.png")

selfreported\_data<-subset(all\_data, all\_data$Self.reported.or.measured == "Self-reported")

meta\_selfreported <- rma.mv(yi=yi, V=vi, data=as.data.frame(selfreported\_data))

f2<-funnel(meta\_selfreported, level=c(90, 95, 99), shade=c("white", "gray", "darkgray"),

yaxis="seinv",xlab="Effect size",ylab="Precision (1/SE)",digits=c(1,0),xlim=c(-1.5,1.5),ylim=c(1,180), legend=TRUE,back="white", hlines = "white", main="b) Article output")

f2

ggsave("figure3b.png")

```

An attempt to test for publication bias with multilevel meta-regression. Significant positive slope would suggest small-study effects (small-studies with larger effect sizes being published that skew my meta-analysis)

```{r eval = TRUE, echo = TRUE, fig.width=9, fig.height=9}

# Application of Equation 24 from the main manuscript of Nakagawa et. al 2021

# Note that we are removing the intercept in this model (using '-1') because the intercept is not of interest at this point. Instead we prefer to see the slopes of the continues moderators and the estimates of all the levels of the categorical moderator 'CaptivityC'. Note that whether or not the intercept is removed, it does not change the conclusions of the small-study effects test or the decline effects test.

publication.bias.model.r.all.se <- rma.mv(yi, vi,

mods=~Variance.as.standard.error+Self.reported.or.measured -1,

random=list(~1|ID.article/ID.observation),

data=all\_data)

#

summary(publication.bias.model.r.all.se)

```

\newpage

## Hypothesis 2: During pandemic conditions the gender gap in academic productivity has increased differentially across research fields.

Prediction 2a: The pandemic has increased the gender gap in academic productivity more in fields that [already had a previously greater gender gap](https://royalsocietypublishing.org/doi/10.1098/rsos.181566) because these lacked gender-equitable support measures to prevent female academics experiencing research production setbacks.

```{r eval = TRUE, echo = TRUE, fig.width=6, fig.height=6}

all\_data$Broad.research.field.reordered<-factor(all\_data$Broad.research.field, c("TEMCP","Biological sciences", "Medicine", "Multidisciplinary", "Social sciences"))

levels(all\_data$Broad.research.field.reordered) <- gsub(" ", "\n", levels(all\_data$Broad.research.field.reordered))

research\_field <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field.reordered, random = list(~1 |

Article.ID, ~1 | ID), data = all\_data)

summary(research\_field)

research\_field1 <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field.reordered - 1, random = list(~1 |

Article.ID, ~1 | ID), data = all\_data)

summary(research\_field1)

my.orchard(research\_field1, mod = "Broad.research.field.reordered", group = "Article.ID", data = all\_data, whisker=0.08, xlab = "Effect size",

alpha = 0.5, transfm = "tanh", cb = FALSE)+

annotate("text", size=6, y=0.5, x=5.09, label=paste("-0.079 [-0.141, -0.018]"))+

annotate("text", size=6, y=0.50, x=4.09, label=paste("-0.058 [-0.139, 0.023]"))+

annotate("text", size=6, y=0.5, x=3.09, label=paste("-0.051 [-0.090, 0.012]"))+

annotate("text", size=6, y=0.50, x=2.09, label=paste("-0.004 [-0.059, 0.050]"))+

annotate("text", size=6, y=0.5, x=1.09, label=paste("0.006 [-0.048, 0.060]"))+

scale\_fill\_manual(values=c("#648FFF","#785EF0","#DC267F","#FE6100","#FFB000"))

ggsave("figure4.png", width=10, height=10, dpi=300)

# same analysis but only used measured studies

research\_field2a <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field, random = list(~1 |

Article.ID, ~1 | ID), data = measured\_data)

summary(research\_field2a)

research\_field2b <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field - 1, random = list(~1 |

Article.ID, ~1 | ID), data = measured\_data)

summary(research\_field2b)

orchaRd::orchard\_plot(research\_field2b, mod = "Broad.research.field", group = "Article.ID", data = measured\_data, xlab = "Effect size",

alpha = 0.5, transfm = "tanh", angle = 45, cb = FALSE)

```

Prediction 2b: The pandemic has increased the gender gap in academic productivity more in fields that had the greatest article output during the pandemic (medicine/social sciences). This could be tricky. The way I would do this is to divide the number of publications or submissions outputted during the pandemic by the number of months sampled during the pandemic (range=1-17, mean=7.55, sd=4.64), to estimate a rate of article output per month and then subtract this from the number of publications/submissions outputted prior to the pandemic divided by the number of months sampled prior to the pandemic (range=1-50, mean=12.32, sd=10.20). This would give a change in the rate of publications per month before and during the pandemic for each paper. But, rates for each paper are calculated using a good range of timescales. 4 papers (12 effect sizes) use only COVID-19-related articles during the pandemic- so I would exclude these?

```{r eval = FALSE, echo = TRUE, fig.width=9, fig.height=9}

## code i dont use

all\_data$surge<-(as.numeric(all\_data$n.during.pandemic) / as.numeric(all\_data$Timeframe.during.pandemic))-(as.numeric(all\_data$n.pre.pandemic) / as.numeric(all\_data$Timeframe.pre.pandemic))

surge\_data<- all\_data[!is.na(all\_data$surge), ]

surge\_model <- metafor::rma.mv(yi = yi, V = vi, mods = ~surge, random = list(~1 |

Article.ID, ~1 | ID), data = all\_data[!is.na(all\_data$surge), ])

summary(surge\_model)

surge1 <- metafor::rma.mv(yi = yi, V = vi, mods = ~surge - 1, random = list(~1 |

Article.ID, ~1 | ID), data = all\_data[!is.na(all\_data$surge), ])

summary(surge1)

orchaRd::orchard\_plot(research\_field1, mod = "Broad.research.field", group = "Article.ID", data = all\_data, xlab = "Effect size",

alpha = 0.5, transfm = "tanh", angle = 45, cb = FALSE)

# same analysis but only used measured studies

research\_field2a <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field, random = list(~1 |

Article.ID, ~1 | ID), data = measured\_data)

summary(research\_field2a)

research\_field2b <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field - 1, random = list(~1 |

Article.ID, ~1 | ID), data = measured\_data)

summary(research\_field2b)

orchaRd::orchard\_plot(research\_field2b, mod = "Broad.research.field", group = "Article.ID", data = measured\_data, xlab = "Effect size", alpha = 0.5, transfm = "tanh", angle = 45, cb = FALSE)

```

\newpage

## Hypothesis 3: Pandemic conditions made it difficult for women to lead research, rather than support research.

Prediction 3: We predict the pandemic has increased the gender gap more in first and last, rather than middle authorship positions as female academics have been more limited in undertaking leading research roles, but not supportive research roles in lockdown conditions.

Note: Only 3 effect sizes from 2 studies that use middle authorship makes it difficult to test this.

```{r eval = TRUE, echo = TRUE, fig.width=9, fig.height=9}

measured\_data$Specific.productivity.measure[measured\_data$Specific.productivity.measure=="Sole authorship"] <- "Last authorship"

measured\_data$Specific.productivity.measure.reordered<-factor(measured\_data$Specific.productivity.measure, c("Last authorship", "Middle authorship", "First authorship", "Corresponding authorship", "Any authorship"))

levels(measured\_data$Specific.productivity.measure.reordered) <- gsub(" ", "\n", levels(measured\_data$Specific.productivity.measure.reordered))

authorship\_position <- metafor::rma.mv(yi = yi, V = vi, mods = ~Specific.productivity.measure.reordered, random = list(~1 | Article.ID, ~1 | ID), data = measured\_data)

summary(authorship\_position)

authorship\_position1 <- metafor::rma.mv(yi = yi, V = vi, mods = ~Specific.productivity.measure.reordered - 1, random = list(~1 |

Article.ID, ~1 | ID), data = measured\_data)

summary(authorship\_position1)

my.orchard(authorship\_position, mod = "Specific.productivity.measure.reordered", group = "Article.ID", data = measured\_data, whisker = 0.08, xlab = "Effect size",

alpha = 0.5)+

annotate("text", size=6, y=-0.4, x=1.09, label=paste("-0.039 [-0.100, -0.023]"))+

annotate("text", size=6, y=-0.4, x=2.09, label=paste("-0.046 [-0.175, 0.084]"))+

annotate("text", size=6, y=-0.4, x=3.09, label=paste("-0.039 [-0.100, 0.023]"))+

annotate("text", size=6, y=-0.4, x=4.09, label=paste("-0.047 [-0.120, 0.027]"))+

annotate("text", size=6, y=-0.4, x=5.09, label=paste("-0.041 [-0.106, 0.023]"))+

scale\_fill\_manual(values=c("#648FFF","#785EF0","#DC267F","#FE6100","#FFB000"))

ggsave("figure6.png")

```

Prediction 3b: We predict the pandemic has increased the gender gap more for research fields of a given authorship position that already had a previously greater gender gap because these lacked gender-equitable support measures to prevent female academics experiencing research production setbacks.

```{r eval = TRUE, echo = TRUE, fig.width=8, fig.height=8}

all\_data$nwomenprepandemic<-round(as.numeric(all\_data$X..women.authors.pre.pandemic)\*as.numeric(all\_data$n.pre.pandemic), digits=0)

all\_data$nmenprepandemic<-round(as.numeric(all\_data$n.pre.pandemic)-as.numeric(all\_data$nwomenprepandemic), digits=0)

all\_data$nwomenduringpandemic<-round(as.numeric(all\_data$X..women.authors.during.pandemic)\*as.numeric(all\_data$n.during.pandemic), digits=0)

all\_data$nmenduringpandemic<-round(as.numeric(all\_data$n.during.pandemic)-as.numeric(all\_data$nwomenduringpandemic), digits=0)

previous\_bias <- metafor::rma.mv(yi = yi, V = vi, mods = ~ cbind(nmenprepandemic/nwomenprepandemic), random = list(~1 |

Article.ID, ~1 | ID), data = subset(all\_data, all\_data$Self.reported.or.measured=="Measured"))

#model suggests that contrary to our prediction, the pandemic has increased the gender gap more for research fields of a given authorship position that were previously less biased.

summary(previous\_bias)

measured\_data$X..women.authors.pre.pandemic<-as.numeric(measured\_data$X..women.authors.pre.pandemic)

measured\_data$X..women.authors.during.pandemic<-as.numeric(measured\_data$X..women.authors.during.pandemic)

measured\_data<-subset(measured\_data, !is.na(measured\_data$X..women.authors.pre.pandemic))

measured\_data<-subset(measured\_data, !is.na(measured\_data$X..women.authors.during.pandemic))

socialsciences<-subset(measured\_data, measured\_data$Broad.research.field=="Social sciences")

medicine<-subset(measured\_data, measured\_data$Broad.research.field=="Medicine")

multi<-subset(measured\_data, measured\_data$Broad.research.field=="Multidisciplinary")

temcp<-subset(measured\_data, measured\_data$Broad.research.field=="TEMCP technology, engineering, mathematic, chemistry and physics")

bio<-subset(measured\_data, measured\_data$Broad.research.field=="Biological sciences")

socialsciences$X..women.authors.pre.pandemic<-as.numeric(socialsciences$X..women.authors.pre.pandemic)

socialsciences$X..women.authors.during.pandemic<-as.numeric(socialsciences$X..women.authors.during.pandemic)

mean((socialsciences$X..women.authors.pre.pandemic))

mean((socialsciences$X..women.authors.during.pandemic))

mean((medicine$X..women.authors.pre.pandemic))

mean((medicine$X..women.authors.during.pandemic))

mean((multi$X..women.authors.pre.pandemic))

mean((multi$X..women.authors.during.pandemic))

mean((temcp$X..women.authors.pre.pandemic))

mean((temcp$X..women.authors.during.pandemic))

mean((bio$X..women.authors.pre.pandemic))

mean((bio$X..women.authors.during.pandemic))

all\_data$Broad.research.field.reordered.opposite<-factor(all\_data$Broad.research.field, c( "Social sciences","Multidisciplinary","Medicine","Biological sciences","TEMCP" ))

all\_data\_long<-all\_data

all\_data\_long$X..women.authors.before.pandemic<-all\_data\_long$X..women.authors.pre.pandemic

all\_data\_long$X..women.authors.before.pandemic<-as.numeric(all\_data\_long$X..women.authors.before.pandemic)

all\_data\_long$X..women.authors.during.pandemic<-as.numeric(all\_data\_long$X..women.authors.during.pandemic)

all\_data\_long <- pivot\_longer(all\_data\_long, c(X..women.authors.before.pandemic, X..women.authors.during.pandemic), names\_to = "Period")

#this plot shows the above model somewhat

ggplot(all\_data\_long, aes(x=factor(Period), y=as.numeric(value), color=factor(Broad.research.field), group = ID)) +

geom\_point(position = position\_jitter(width = .0)) +

geom\_smooth(method = 'lm', se = FALSE) +

labs(

x = "x",

color = "Broad.research.field"

)

all\_data\_long$Period[all\_data\_long$Period == 'X..women.authors.before.pandemic']<-'Before pandemic'

all\_data\_long$Period[all\_data\_long$Period == 'X..women.authors.during.pandemic']<-'During pandemic'

all\_data\_long$'Broad research field'<-all\_data\_long$Broad.research.field.reordered

#this plot shows perhaps more clearly the change in gender gap according to previous gender gap and research field.

ggplot(all\_data\_long, aes(x=factor(Period), y=as.numeric(value), color=factor(Broad.research.field.reordered), group = Broad.research.field.reordered)) +

geom\_point(position = position\_jitter(width = .0)) +

geom\_smooth(method = 'lm', se = FALSE) +

labs(

color = "Research field"

) +

labs( x = "", y = "Proportion female authors")+ theme\_linedraw()+

theme( panel.grid.major.x = element\_blank() , panel.grid.major = element\_line(size = 0.1, linetype = 'solid',

colour = "gray"),

panel.grid.minor = element\_line(size = 0.1, linetype = 'solid',

colour = "gray"))+theme(axis.text=element\_text(size=10))+scale\_color\_manual(values=c("#648FFF","#785EF0","#DC267F","#FE6100","#FFB000"))

ggsave("figure7.png", width=5, height=5, dpi=300)

```

Prediction 3b: We predict the pandemic has increased the gender gap more for research fields of a given authorship position that already had a previously greater gender gap because these lacked gender-equitable support measures to prevent female academics experiencing research production setbacks.

```{r eval = TRUE, echo = TRUE, fig.width=9, fig.height=9}

all\_data$nwomenprepandemic<-round(as.numeric(all\_data$X..women.authors.pre.pandemic)\*as.numeric(all\_data$n.pre.pandemic), digits=0)

all\_data$nmenprepandemic<-round(as.numeric(all\_data$n.pre.pandemic)-as.numeric(all\_data$nwomenprepandemic), digits=0)

all\_data$nwomenduringpandemic<-round(as.numeric(all\_data$X..women.authors.during.pandemic)\*as.numeric(all\_data$n.during.pandemic), digits=0)

all\_data$nmenduringpandemic<-round(as.numeric(all\_data$n.during.pandemic)-as.numeric(all\_data$nwomenduringpandemic), digits=0)

previous\_bias <- metafor::rma.mv(yi = yi, V = vi, mods = ~ cbind(nmenprepandemic/nwomenprepandemic), random = list(~1 |

Article.ID, ~1 | ID), data = subset(all\_data, all\_data$Self.reported.or.measured=="Measured"))

summary(previous\_bias)

ggplot(all\_data) +

geom\_segment(aes(x = 1, xend = 2, y = as.numeric(X..women.authors.pre.pandemic), yend = as.numeric(X..women.authors.during.pandemic), color=Broad.research.field), size=0.6) +

scale\_x\_discrete(name = "Time period", breaks = c("1", "2"), labels = c("Pre-pandemic", "During pandemic"), limits = c(1, 2)) +

scale\_y\_continuous(name = "Proportion female authors")+

theme(legend.position="right")+

theme(legend.title = element\_text(size = 4),

legend.text = element\_text(size = 4))

```

### Heterogeneity for all papers

```{r eval = TRUE, echo = FALSE}

I2\_all\_data<-i2\_ml(m)

I2\_all\_data

round(I2\_all\_data, digits = 3)

m\_het <- rma.mv(yi = yi,

V = vi,

mods = ~ 1,

random = list(~1|ID.article,

~1|ID.observation,

~1|Self.reported.or.measured,

~1|Specific.productivity.measure,

~1|Broad.productivity.measure),

data=all\_data)

I2\_m\_het<-i2\_ml(m\_het)

round(I2\_m\_het, digits = 10)

```