Investigating the impact of the COVID-19 pandemic on the gender productivity gap in academia

Kiran Gok Lune Lee

2022-11-21

Install and load packages

```r
## 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)
```

```
## Loading required package: Matrix
```

```
## Loading required package: metadat
```

```
##
## Loading the 'metafor' package (version 3.8-1). For an
## introduction to the package please type: help(metafor)
```

```
## Loading required package: usethis
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## v purrr   0.3.5
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x tidyr::expand() masks Matrix::expand()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## x tidyr::pack()   masks Matrix::pack()
## x tidyr::unpack() masks Matrix::unpack()
##
## Loading the 'orchaRd' package (version 2.0). For an
## introduction and vignette to the package please see: https://daniel1noble.github.io/orchaRd/
```

Load data

```
##
##                                                                     NA
##                                                                     10
##               Standardised mean difference (binary proportions)
##                                                                      4
##                      Standardised mean difference (chi-square)
##                                                                      1
##          Standardised mean difference (f-test, unequal sample size)
##                                                                      1
```

```
##              Standardised mean difference (from binary proportions)
##                                                                   3
##                      Standardised mean difference (means and sds)
##                                                                   6
##                      Standardised mean difference (means and ses)
##                                                                  12
## Standardised mean difference (t-test p-value unequal sample sizes)
##                                                                   1
##           Standardised mean difference (two by two frequency table)
##                                                                   7
##                    Standardised mean difference (two by two table)
##                                                                  70

##
##       Measured Self-reported
##             92           23

##
## Biological sciences          Medicine  Multidisciplinary      Social sciences
##                  17                44                 15                   19
##                TEMCP
##                   20

##
##       Other Publications  Submissions
##            18           48           49

##
##                       Burn-out        General productivity
##                              1                          11
##                       Job loss                    Preprints
##                              1                          20
##                       Projects                 Publications
##                              1                          48
##             Submissions Submissions (self-reported)
##                             23                           5
##             Work hours              Working papers
##                      4                           1

##
##                                                          Ability to submit papers
##                                                                                 1
##                                                                 Academic job loss
##                                                                                 1
##                                                             Academic productivity
##                                                                                 1
##                                                                   Any authorship
##                                                                                13
##                                 Binary whether loss of research productivity
##                                                                                 1
##                                                            Corresponding authorship
##                                                                                12
##                                   Disruption from having to work from home
##                                                                                 1
##                                 Effect of COVID-19 pandemic on work from home
##                                                                                 1
##                                                                 First authorship
##                                                                                45
##                                                          Hours of research per day
```

```
##                                                                          1
##                                                          Last authorship
##                                                                         17
##                                                        Middle authorship
##                                                                          4
##                                          Number of new research projects
##                                                                          1
##                             Pandemic effect on academic productivity
##                                                                          1
##                                                  Research hours per week
##                                                                          1
##                                                    Research productivity
##                                                                          1
##             Research productivity decrease following COVID-19 relative to other gender
##                                                                          1
##                                    Self reported ability to submit/complete work
##                                                                          1
##                              Self reproted adverse effect of pandemics on work perfomance
##                                                                          1
## self-reported loss in productivity (unability to write/research at home compared to pre-pandemic level)
##                                                                          1
##                                    Self-reported productivity change on likert scale
##                                                                          1
##                                                          Sole authorship
##                                                                          4
##                              Whether the pandemic created low productivity (Y/N)
##                                                                          1
##                                                Work productivity from home
##                                                                          1
##    Work-productivity during pandemics: working hours, percentage of time spent on research and education
##                                                                          1
##                                                        Work-related burnout
##                                                                          1
```

Custom settings for orchaRd plots

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).

Looks like an overall negative effect: during pandemic conditions the gender gap in academic productivity has increased.

```
# Hierarchical mixed effect meta-analysis with all papers.
# Studies with multiple effect sizes are controlled for.
m <- rma.mv(yi, vi, random = ~1 | ID.article/ID.observation,
    data = all_data)
m
```

```
##
## Multivariate Meta-Analysis Model (k = 115; method: REML)
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed                        factor
## sigma^2.1  0.0022  0.0468     50     no                    ID.article
## sigma^2.2  0.0136  0.1165    115     no  ID.article/ID.observation
##
## Test for Heterogeneity:
## Q(df = 114) = 4789.5722, p-val < .0001
##
## Model Results:
##
## estimate      se     zval     pval    ci.lb    ci.ub
##  -0.0700  0.0163  -4.2910   <.0001  -0.1020  -0.0380  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Forest plot produced for visual representation.
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)
```

| **Author(s) and Year** | **Effect size [95% CI]** |

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 = 0.5, x = 1.06, label = paste("-0.070 [-0.102, -0.038]")) +
    scale_fill_manual(values = c("#DC267F"))
```

```
## Warning in k == TRUE && g == TRUE && k.pos == "right": 'length(x) = 2 > 1' in
## coercion to 'logical(1)'
```
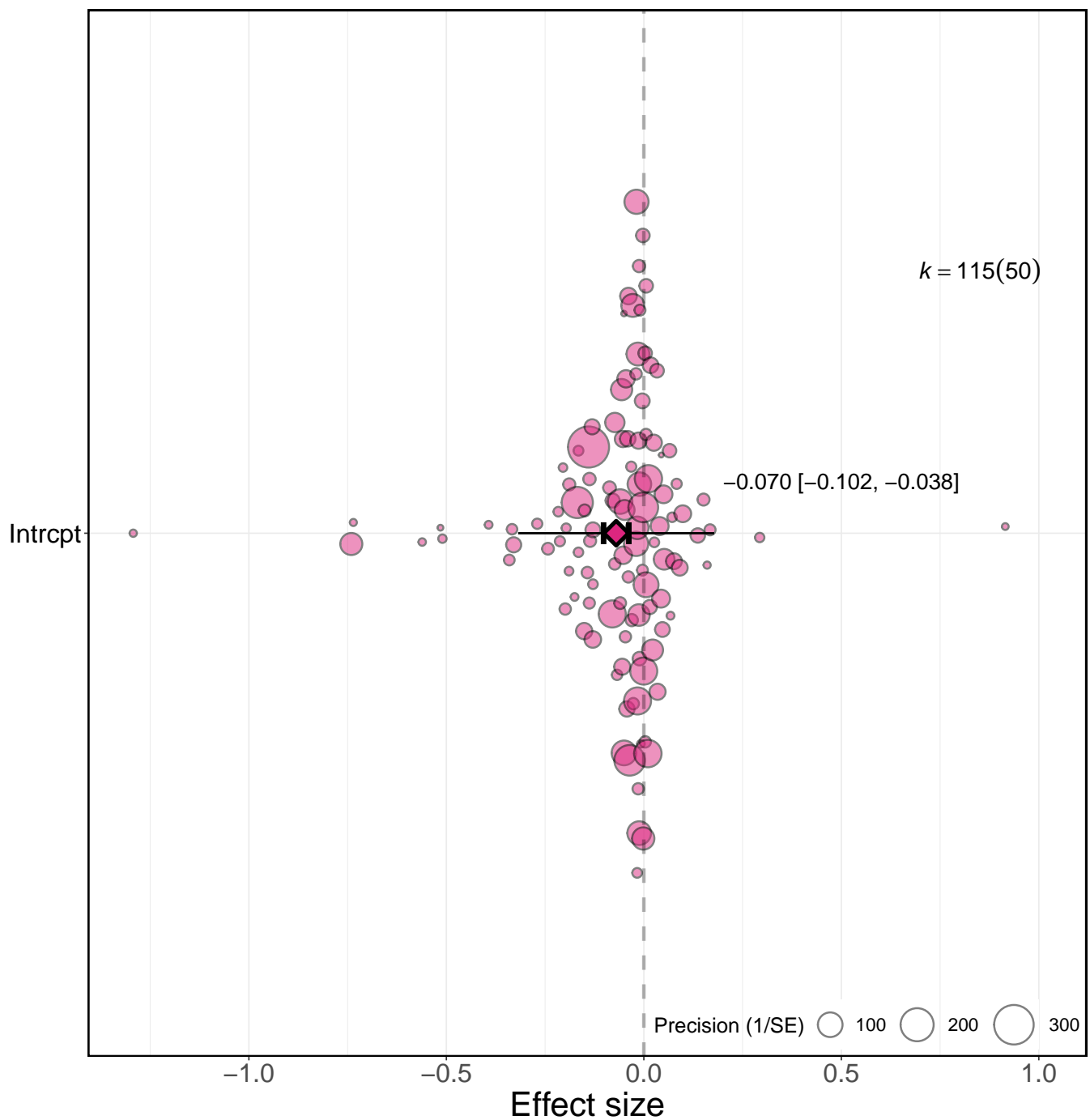
```
ggsave("figure1.png", width = 10, height = 10, dpi = 300)
```

Prediction 1b: Though studies can measure the type of research productivity as by individual surveys, numbers of submissions and numbers of publications, this does not influence the gender gap increase observed during the pandemic.

```
all_data$Broad.productivity.measure[all_data$Broad.productivity.measure ==
    "Other"] <- "Survey"

tapply(all_data$ID.observation, all_data$Broad.productivity.measure,
    length)
```

```
## Publications  Submissions       Survey
##           48           49           18
```

```
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
```

```
##
## Multivariate Meta-Analysis Model (k = 115; method: REML)
##
## Variance Components:
##
##             estim    sqrt   nlvls  fixed                  factor
## sigma^2.1  0.0000  0.0000      50     no              ID.article
## sigma^2.2  0.0142  0.1190     115     no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 112) = 4751.7560, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 12.3488, p-val = 0.0021
##
## Model Results:
##
##                                                    estimate      se     zval
## intrcpt                                             -0.0458  0.0186  -2.4599
## Broad.productivity.measure.reorderedSubmissions      0.0071  0.0268   0.2664
## Broad.productivity.measure.reorderedSurvey          -0.1469  0.0448  -3.2777
##                                                        pval   ci.lb    ci.ub
## intrcpt                                              0.0139  -0.0822  -0.0093   *
## Broad.productivity.measure.reorderedSubmissions      0.7899  -0.0453   0.0596
## Broad.productivity.measure.reorderedSurvey           0.0010  -0.2347  -0.0590   **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
##
## Multivariate Meta-Analysis Model (k = 115; method: REML)
##
## Variance Components:
##
##             estim    sqrt   nlvls  fixed                  factor
## sigma^2.1  0.0000  0.0000      50     no              ID.article
## sigma^2.2  0.0142  0.1190     115     no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 112) = 4751.7560, p-val < .0001
##
## Test of Moderators (coefficients 1:3):
## QM(df = 3) = 32.4292, p-val < .0001
##
## Model Results:
##
```

```
##                                                  estimate      se     zval
## Broad.productivity.measure.reorderedPublications  -0.0458  0.0186  -2.4599
## Broad.productivity.measure.reorderedSubmissions   -0.0386  0.0192  -2.0107
## Broad.productivity.measure.reorderedSurvey        -0.1926  0.0408  -4.7260
##                                                     pval   ci.lb    ci.ub
## Broad.productivity.measure.reorderedPublications  0.0139  -0.0822  -0.0093    *
## Broad.productivity.measure.reorderedSubmissions   0.0444  -0.0763  -0.0010    *
## Broad.productivity.measure.reorderedSurvey        <.0001  -0.2725  -0.1128  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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"))
```

```
## Warning in k == TRUE && g == TRUE && k.pos == "right": 'length(x) = 2 > 1' in
## coercion to 'logical(1)'
```

```
p1
```
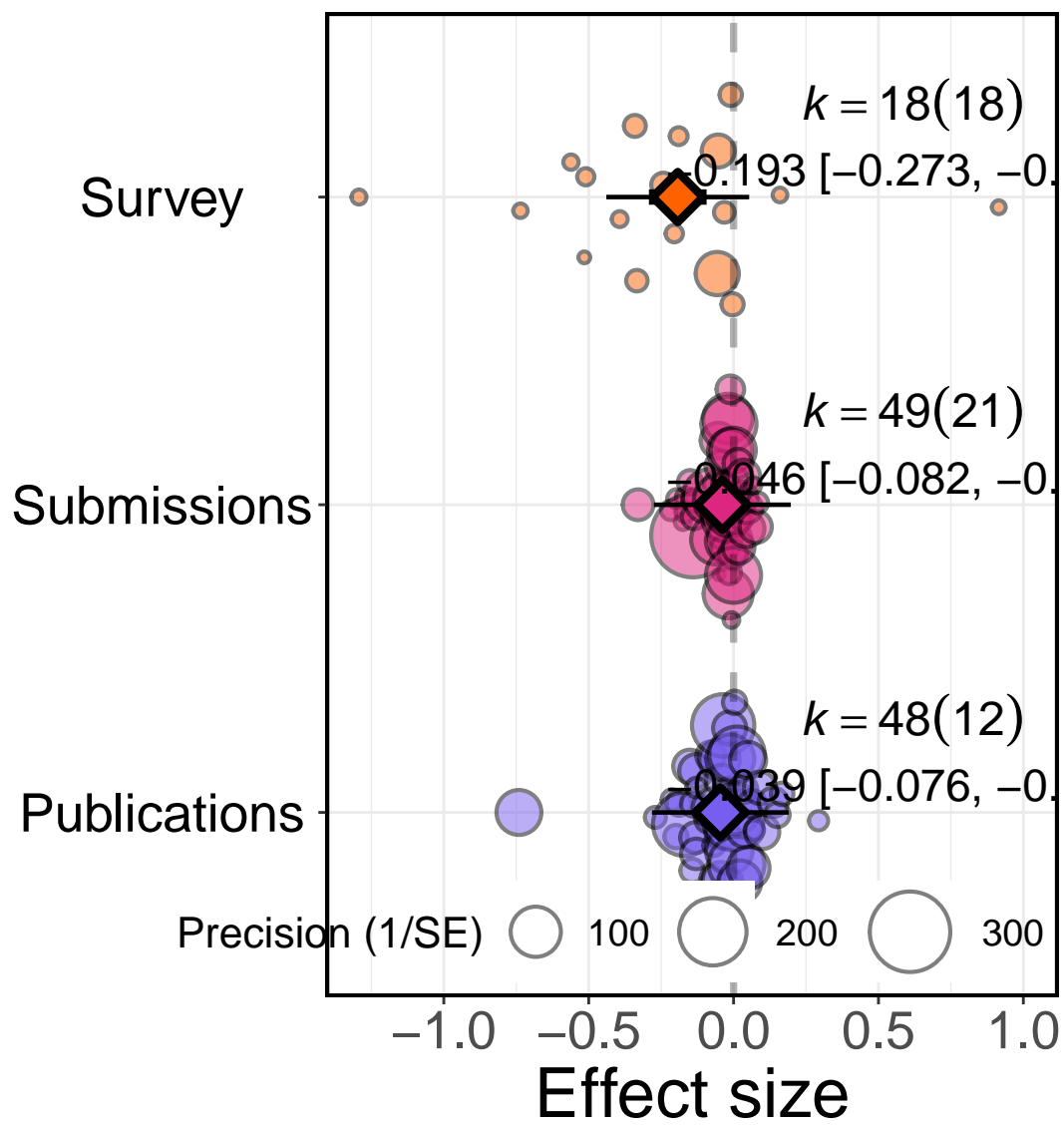
```
ggsave("figure2.png")
```

```
## Saving 6 x 6 in image
```

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 because these lacked gender-equitable support measures to prevent female academics experiencing research production setbacks.

```
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)
```
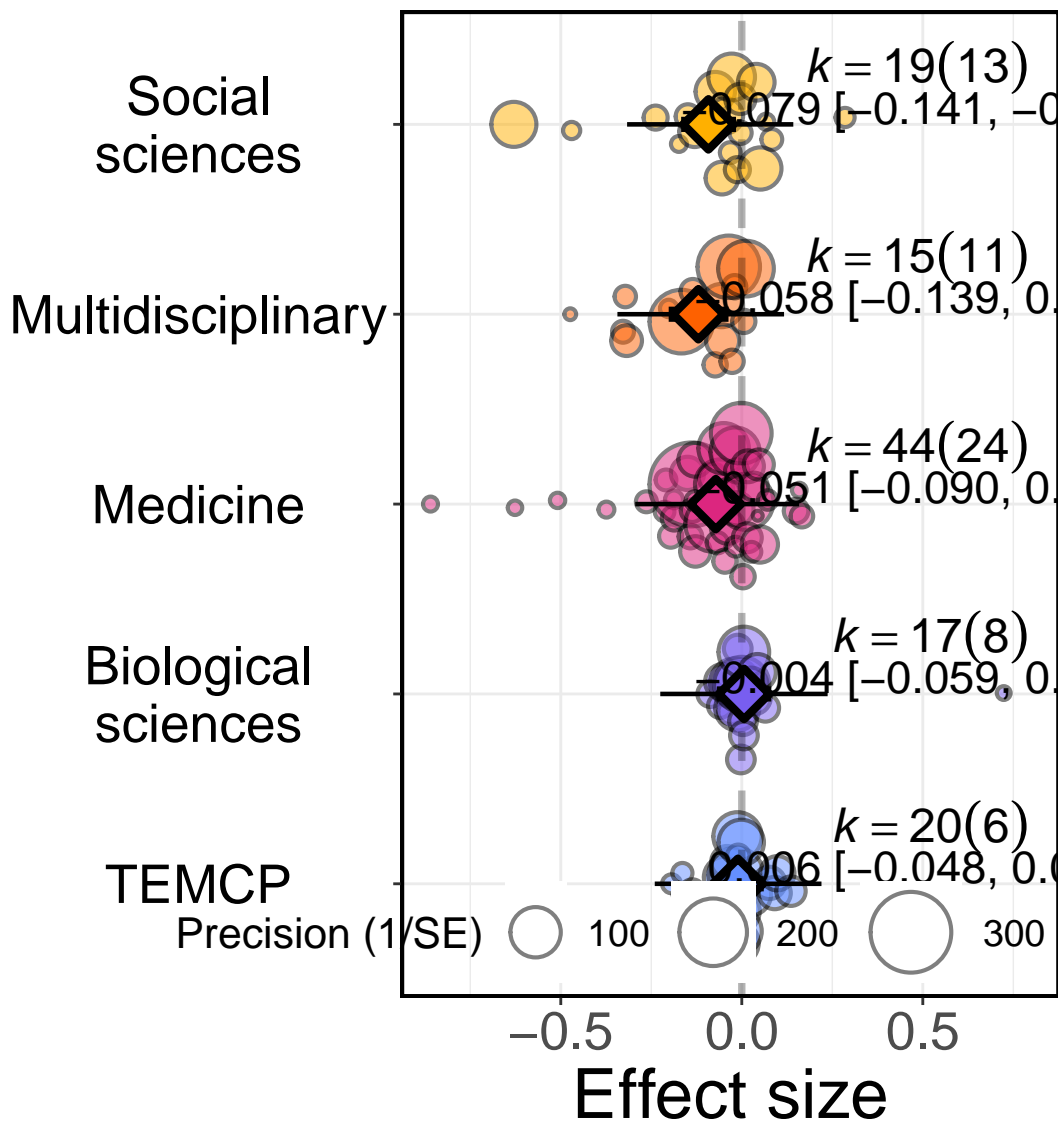
```
##
## Multivariate Meta-Analysis Model (k = 115; method: REML)
##
##    logLik  Deviance       AIC       BIC      AICc
##   44.8688  -89.7376  -75.7376  -56.8342  -74.6395
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000     50     no  Article.ID
## sigma^2.2  0.0135  0.1162    115     no          ID
##
## Test for Residual Heterogeneity:
## QE(df = 110) = 3950.1936, p-val < .0001
##
## Test of Moderators (coefficients 2:5):
## QM(df = 4) = 12.1382, p-val = 0.0164
##
## Model Results:
##
##                                                     estimate      se      zval
## intrcpt                                              -0.0108  0.0284   -0.3781
## Broad.research.field.reorderedBiological\nsciences    0.0171  0.0416    0.4099
## Broad.research.field.reorderedMedicine               -0.0612  0.0354   -1.7273
## Broad.research.field.reorderedMultidisciplinary      -0.1097  0.0449   -2.4452
## Broad.research.field.reorderedSocial\nsciences       -0.0816  0.0420   -1.9422
##                                                        pval    ci.lb     ci.ub
## intrcpt                                              0.7054  -0.0665    0.0450
## Broad.research.field.reorderedBiological\nsciences   0.6819  -0.0645    0.0986
## Broad.research.field.reorderedMedicine               0.0841  -0.1306    0.0082  .
## Broad.research.field.reorderedMultidisciplinary      0.0145  -0.1976   -0.0218  *
## Broad.research.field.reorderedSocial\nsciences       0.0521  -0.1640    0.0007  .
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
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)
```

```
##
## Multivariate Meta-Analysis Model (k = 115; method: REML)
##
##    logLik  Deviance       AIC       BIC      AICc
##   44.8688  -89.7376  -75.7376  -56.8342  -74.6395
##
## Variance Components:
##
##               estim     sqrt   nlvls   fixed      factor
## sigma^2.1    0.0000   0.0000      50      no  Article.ID
## sigma^2.2    0.0135   0.1162     115      no          ID
##
## Test for Residual Heterogeneity:
## QE(df = 110) = 3950.1936, p-val < .0001
##
## Test of Moderators (coefficients 1:5):
## QM(df = 5) = 32.8003, p-val < .0001
##
## Model Results:
##
##                                                 estimate      se     zval
## Broad.research.field.reorderedTEMCP              -0.0108  0.0284  -0.3781
## Broad.research.field.reorderedBiological\nsciences 0.0063  0.0304   0.2076
## Broad.research.field.reorderedMedicine           -0.0719  0.0211  -3.4112
## Broad.research.field.reorderedMultidisciplinary  -0.1204  0.0347  -3.4727
## Broad.research.field.reorderedSocial\nsciences   -0.0924  0.0309  -2.9864
##                                                    pval   ci.lb    ci.ub
## Broad.research.field.reorderedTEMCP              0.7054  -0.0665   0.0450
## Broad.research.field.reorderedBiological\nsciences 0.8356  -0.0532   0.0658
## Broad.research.field.reorderedMedicine           0.0006  -0.1132  -0.0306
## Broad.research.field.reorderedMultidisciplinary  0.0005  -0.1884  -0.0525
## Broad.research.field.reorderedSocial\nsciences   0.0028  -0.1530  -0.0317
##
## Broad.research.field.reorderedTEMCP
## Broad.research.field.reorderedBiological\nsciences
## Broad.research.field.reorderedMedicine                  ***
## Broad.research.field.reorderedMultidisciplinary         ***
## Broad.research.field.reorderedSocial\nsciences           **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 3

```
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 = 0.5, 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 = 0.5, 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"))
```

```
## Warning in k == TRUE && g == TRUE && k.pos == "right": 'length(x) = 2 > 1' in
## coercion to 'logical(1)'
```

11

Figure displays a bubble/forest plot with the following categories and values:

- Social sciences: $k = 19(13)$, −0.079 [−0.141, −0.
- Multidisciplinary: $k = 15(11)$, −0.058 [−0.139, 0.
- Medicine: $k = 44(24)$, −0.051 [−0.090, 0.
- Biological sciences: $k = 17(8)$, −0.004 [−0.059, 0.
- TEMCP: $k = 20(6)$, 0.006 [−0.048, 0.

Precision (1/SE): 100   200   300

Effect size

```
ggsave("figure3.png", width = 10, height = 10, dpi = 300)
```

Prediction 2b: 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 less support may have been available to women to balance the effects of the pandemic.

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

```
## Warning: NAs introduced by coercion
```

```
## Warning: NAs introduced by coercion
```

```
all_data$nmenprepandemic <- round(as.numeric(all_data$n.pre.pandemic) -
    as.numeric(all_data$nwomenprepandemic), digits = 0)
```

```
## Warning: NAs introduced by coercion
```

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

```
## Warning: NAs introduced by coercion
```

```
## Warning: NAs introduced by coercion
```

```r
all_data$nmenduringpandemic <- round(as.numeric(all_data$n.during.pandemic) -
    as.numeric(all_data$nwomenduringpandemic), digits = 0)
```

```
## Warning: NAs introduced by coercion
```

```r
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"))
```

```
## Warning: Rows with NAs omitted from model fitting.
```

```r
# 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)
```

```
##
## Multivariate Meta-Analysis Model (k = 84; method: REML)
##
##      logLik   Deviance        AIC        BIC        AICc
##     59.7103  -119.4207  -111.4207  -101.7938  -110.9012
##
## Variance Components:
##
##               estim     sqrt   nlvls  fixed     factor
## sigma^2.1   0.0000   0.0000      25     no  Article.ID
## sigma^2.2   0.0110   0.1051      84     no          ID
##
## Test for Residual Heterogeneity:
## QE(df = 82) = 3373.8831, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 11.0156, p-val = 0.0009
##
## Model Results:
##
##                                             estimate       se     zval     pval
## intrcpt                                      -0.1232   0.0301  -4.0981   <.0001
## cbind(nmenprepandemic/nwomenprepandemic)      0.0389   0.0117   3.3190   0.0009
##                                               ci.lb    ci.ub
## intrcpt                                      -0.1820  -0.0643  ***
## cbind(nmenprepandemic/nwomenprepandemic)      0.0159   0.0619  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
measured_data$X..women.authors.pre.pandemic <- as.numeric(measured_data$X..women.authors.pre.pandemic)
```

```
## Warning: NAs introduced by coercion
```

```r
measured_data$X..women.authors.during.pandemic <- as.numeric(measured_data$X..women.authors.during.pandemic
```

```
## Warning: NAs introduced by coercion
```

```r
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.pande
mean((socialsciences$X..women.authors.pre.pandemic))
```

```
## [1] 0.3765561
```

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

```
## [1] 0.3581449
```

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

```
## [1] 0.3693147
```

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

```
## [1] 0.3461209
```

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

```
## [1] 0.3916469
```

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

```
## [1] 0.3669248
```

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

```
## [1] NaN
```

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

```
## [1] NaN
```

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

```
## [1] 0.3280901
```

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

```
## [1] 0.3286446
```

```
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
```

```
## Warning: NAs introduced by coercion
```

```
all_data_long$X..women.authors.during.pandemic <- as.numeric(all_data_long$X..women.authors.during.pandemi
```

```
## Warning: NAs introduced by coercion
```

```
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 about how previous gender
# bias influence gender bias during the pandemic
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")
```
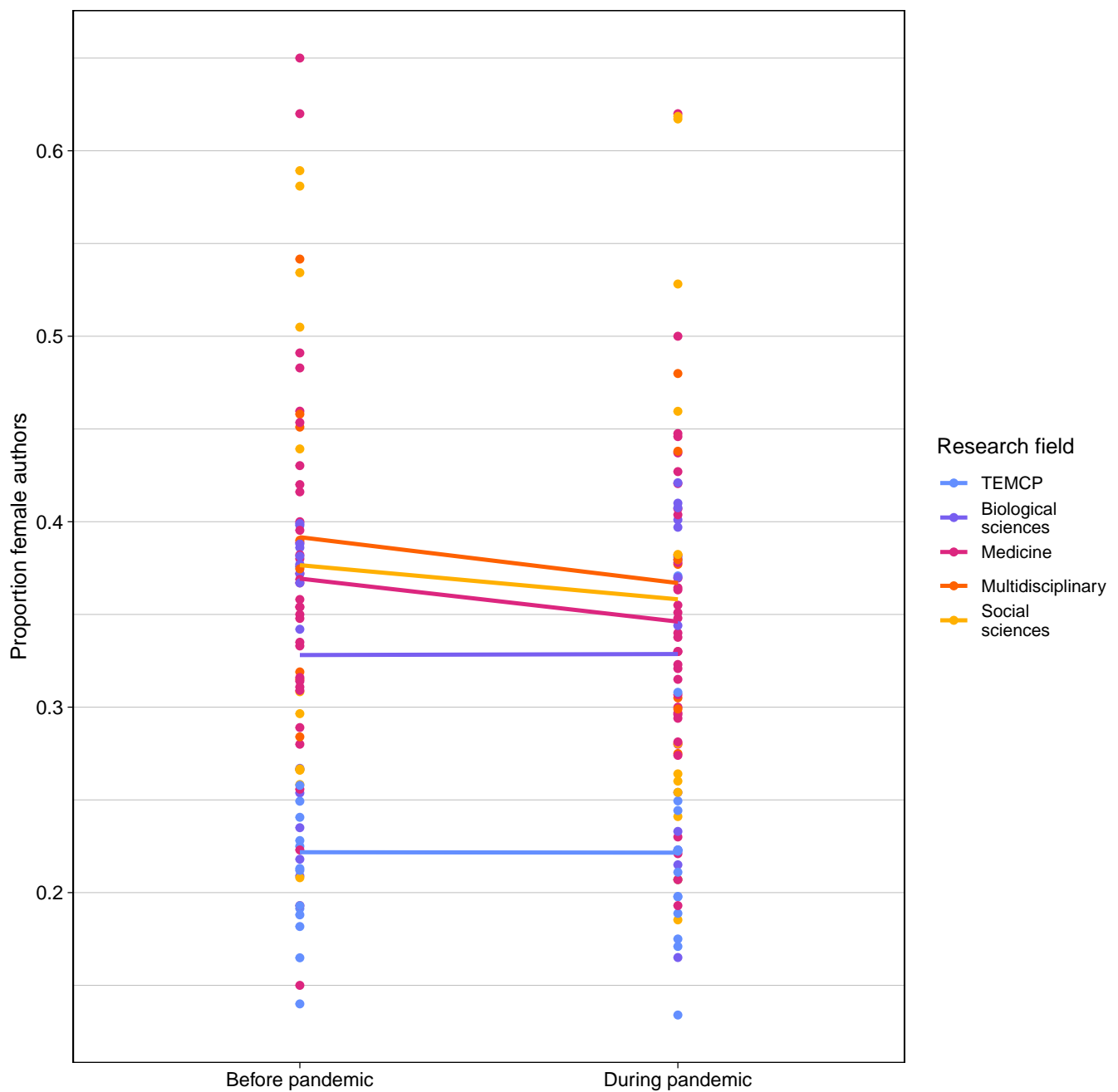
`geom_smooth()` using formula 'y ~ x'

Warning: Removed 62 rows containing non-finite values (stat_smooth).

Warning: Removed 62 rows containing missing values (geom_point).
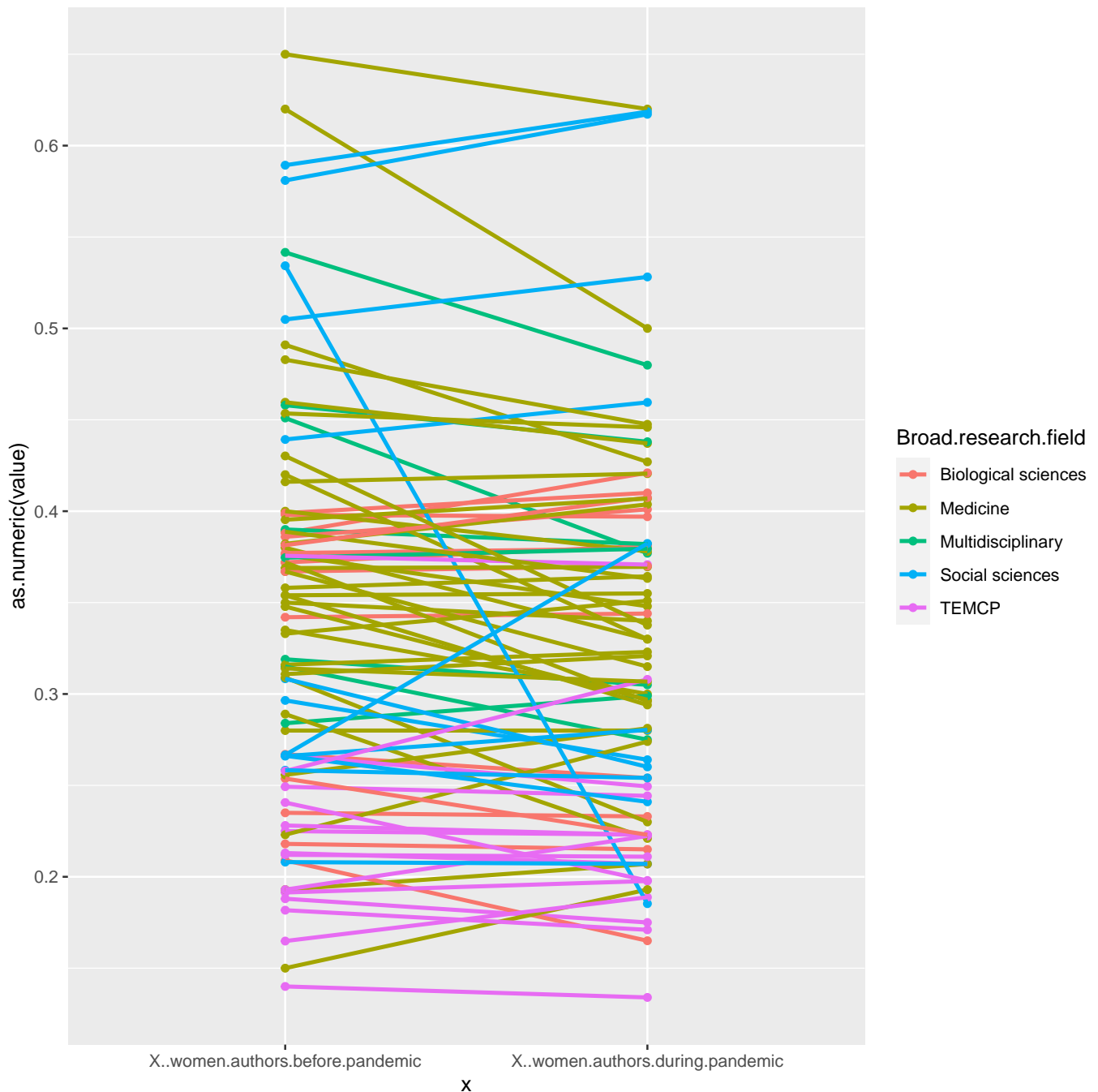
```
all_data_long$Period[all_data_long$Period == "X..women.authors.before.pandemic"] <- "Before pandemic"
all_data_long$Period[all_data_long$Period == "X..women.authors.during.pandemic"] <- "During pandemic"
all_data_long$"Broad research field" <- all_data_long$Broad.research.field.reordered
```

Figure 4

```
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"))
```

`geom_smooth()` using formula 'y ~ x'

Warning: Removed 62 rows containing non-finite values (stat_smooth).

Warning: Removed 62 rows containing missing values (geom_point).

```
ggsave("figure4.png", width = 5, height = 5, dpi = 300)
```

`geom_smooth()` using formula 'y ~ x'

Warning: Removed 62 rows containing non-finite values (stat_smooth).
Removed 62 rows containing missing values (geom_point).

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 pandemic conditions.

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

```
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))

# Models
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)
```

```
##
## Multivariate Meta-Analysis Model (k = 84; method: REML)
##
##     logLik   Deviance        AIC        BIC       AICc
##    51.1354  -102.2708   -88.2708   -71.6847   -86.6934
##
## Variance Components:
##
##             estim     sqrt  nlvls  fixed     factor
## sigma^2.1  0.0000   0.0000     25     no  Article.ID
## sigma^2.2  0.0132   0.1150     84     no          ID
##
## Test for Residual Heterogeneity:
## QE(df = 79) = 3776.9941, p-val < .0001
##
## Test of Moderators (coefficients 2:5):
## QM(df = 4) = 0.2786, p-val = 0.9912
##
## Model Results:
##
##                                                               estimate
## intrcpt                                                        -0.0343
## Specific.productivity.measure.reorderedMiddle\nauthorship      -0.0112
## Specific.productivity.measure.reorderedFirst\nauthorship        0.0033
## Specific.productivity.measure.reorderedCorresponding\nauthorship  0.0181
## Specific.productivity.measure.reorderedAny\nauthorship         -0.0123
##                                                                     se
## intrcpt                                                         0.0292
## Specific.productivity.measure.reorderedMiddle\nauthorship       0.0737
## Specific.productivity.measure.reorderedFirst\nauthorship        0.0349
## Specific.productivity.measure.reorderedCorresponding\nauthorship  0.0564
## Specific.productivity.measure.reorderedAny\nauthorship          0.0504
```

```
##                                                                   zval
## intrcpt                                                        -1.1745
## Specific.productivity.measure.reorderedMiddle\nauthorship      -0.1521
## Specific.productivity.measure.reorderedFirst\nauthorship        0.0934
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.3219
## Specific.productivity.measure.reorderedAny\nauthorship         -0.2442
##                                                                   pval
## intrcpt                                                         0.2402
## Specific.productivity.measure.reorderedMiddle\nauthorship       0.8791
## Specific.productivity.measure.reorderedFirst\nauthorship        0.9256
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.7475
## Specific.productivity.measure.reorderedAny\nauthorship          0.8071
##                                                                   ci.lb
## intrcpt                                                        -0.0916
## Specific.productivity.measure.reorderedMiddle\nauthorship      -0.1558
## Specific.productivity.measure.reorderedFirst\nauthorship       -0.0651
## Specific.productivity.measure.reorderedCorresponding\nauthorship -0.0924
## Specific.productivity.measure.reorderedAny\nauthorship         -0.1112
##                                                                   ci.ub
## intrcpt                                                         0.0230
## Specific.productivity.measure.reorderedMiddle\nauthorship       0.1333
## Specific.productivity.measure.reorderedFirst\nauthorship        0.0716
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.1287
## Specific.productivity.measure.reorderedAny\nauthorship          0.0865
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
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)
```

```
##
## Multivariate Meta-Analysis Model (k = 84; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
##   51.1354  -102.2708   -88.2708   -71.6847   -86.6934
##
## Variance Components:
##
##            estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000     25     no  Article.ID
## sigma^2.2  0.0132  0.1150     84     no          ID
##
## Test for Residual Heterogeneity:
## QE(df = 79) = 3776.9941, p-val < .0001
##
## Test of Moderators (coefficients 1:5):
## QM(df = 5) = 5.9053, p-val = 0.3155
##
## Model Results:
##
##                                                                estimate
## Specific.productivity.measure.reorderedLast\nauthorship         -0.0343
## Specific.productivity.measure.reorderedMiddle\nauthorship       -0.0456
## Specific.productivity.measure.reorderedFirst\nauthorship        -0.0311
```

```
## Specific.productivity.measure.reorderedCorresponding\nauthorship  -0.0162
## Specific.productivity.measure.reorderedAny\nauthorship            -0.0467
##                                                                         se
## Specific.productivity.measure.reorderedLast\nauthorship         0.0292
## Specific.productivity.measure.reorderedMiddle\nauthorship       0.0677
## Specific.productivity.measure.reorderedFirst\nauthorship        0.0190
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.0482
## Specific.productivity.measure.reorderedAny\nauthorship          0.0411
##                                                                       zval
## Specific.productivity.measure.reorderedLast\nauthorship        -1.1745
## Specific.productivity.measure.reorderedMiddle\nauthorship      -0.6729
## Specific.productivity.measure.reorderedFirst\nauthorship       -1.6346
## Specific.productivity.measure.reorderedCorresponding\nauthorship -0.3358
## Specific.productivity.measure.reorderedAny\nauthorship         -1.1350
##                                                                       pval
## Specific.productivity.measure.reorderedLast\nauthorship         0.2402
## Specific.productivity.measure.reorderedMiddle\nauthorship       0.5010
## Specific.productivity.measure.reorderedFirst\nauthorship        0.1021
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.7370
## Specific.productivity.measure.reorderedAny\nauthorship          0.2564
##                                                                      ci.lb
## Specific.productivity.measure.reorderedLast\nauthorship        -0.0916
## Specific.productivity.measure.reorderedMiddle\nauthorship      -0.1783
## Specific.productivity.measure.reorderedFirst\nauthorship       -0.0683
## Specific.productivity.measure.reorderedCorresponding\nauthorship -0.1107
## Specific.productivity.measure.reorderedAny\nauthorship         -0.1272
##                                                                      ci.ub
## Specific.productivity.measure.reorderedLast\nauthorship         0.0230
## Specific.productivity.measure.reorderedMiddle\nauthorship       0.0871
## Specific.productivity.measure.reorderedFirst\nauthorship        0.0062
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.0783
## Specific.productivity.measure.reorderedAny\nauthorship          0.0339
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5

```
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"))
```

```
## Warning in k == TRUE && g == TRUE && k.pos == "right": 'length(x) = 2 > 1' in
## coercion to 'logical(1)'
```
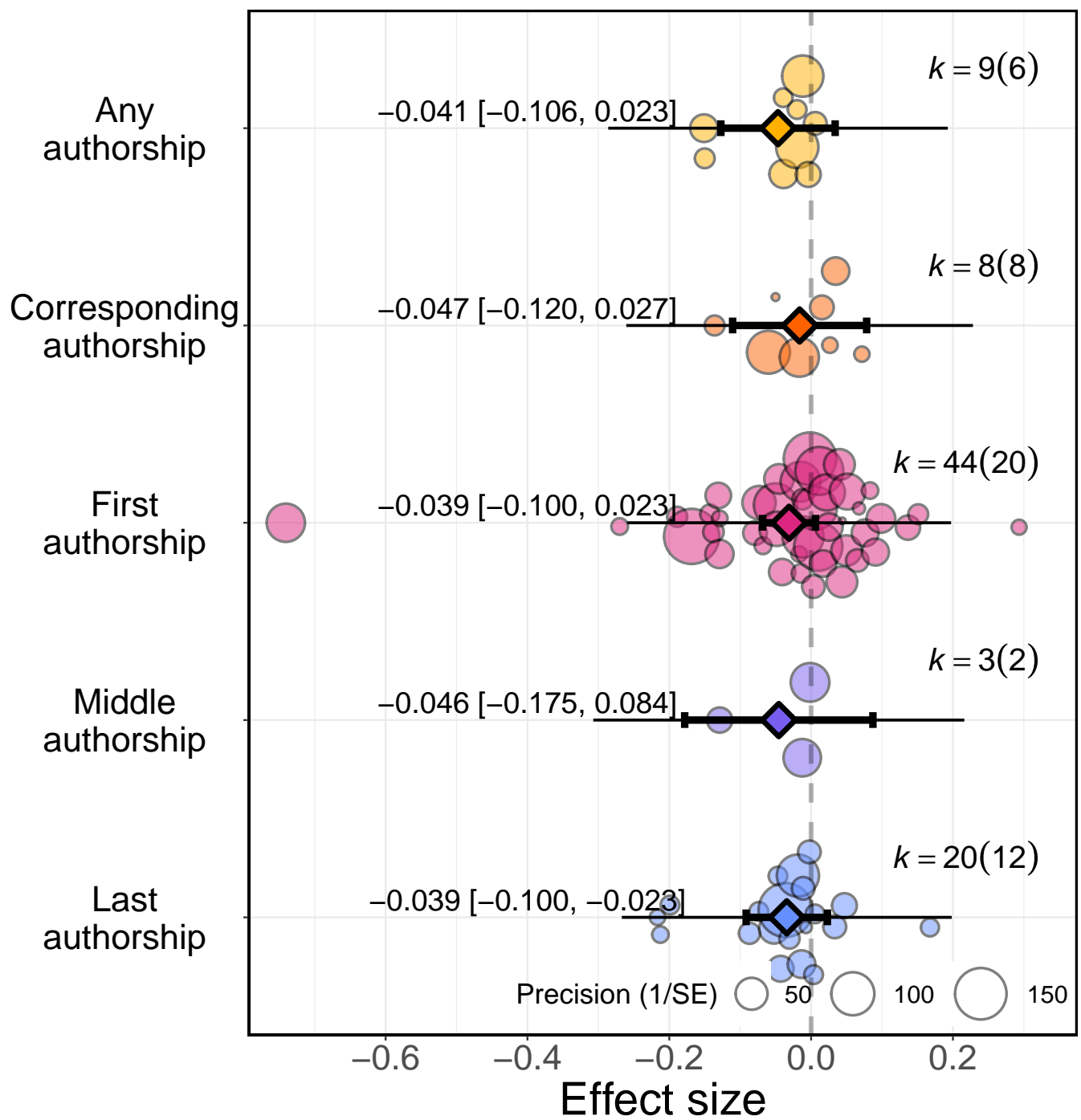
```
ggsave("figure5.png")
```

```
## Saving 9 x 9 in image
```

Heterogeneity test

```
##                     I2_Total                 I2_ID.article
##                     98.20614                      13.63583
## I2_ID.article/ID.observation
##                     84.57031

##                     I2_Total                 I2_ID.article
##                      98.206                        13.636
## I2_ID.article/ID.observation
##                      84.570

##                        I2_Total                    I2_ID.article
##                    98.891445006                      0.000000019
##               I2_ID.observation     I2_Self.reported.or.measured
##                    49.840640483                     49.050804504
## I2_Specific.productivity.measure     I2_Broad.productivity.measure
##                     0.000000000                      0.000000000
```
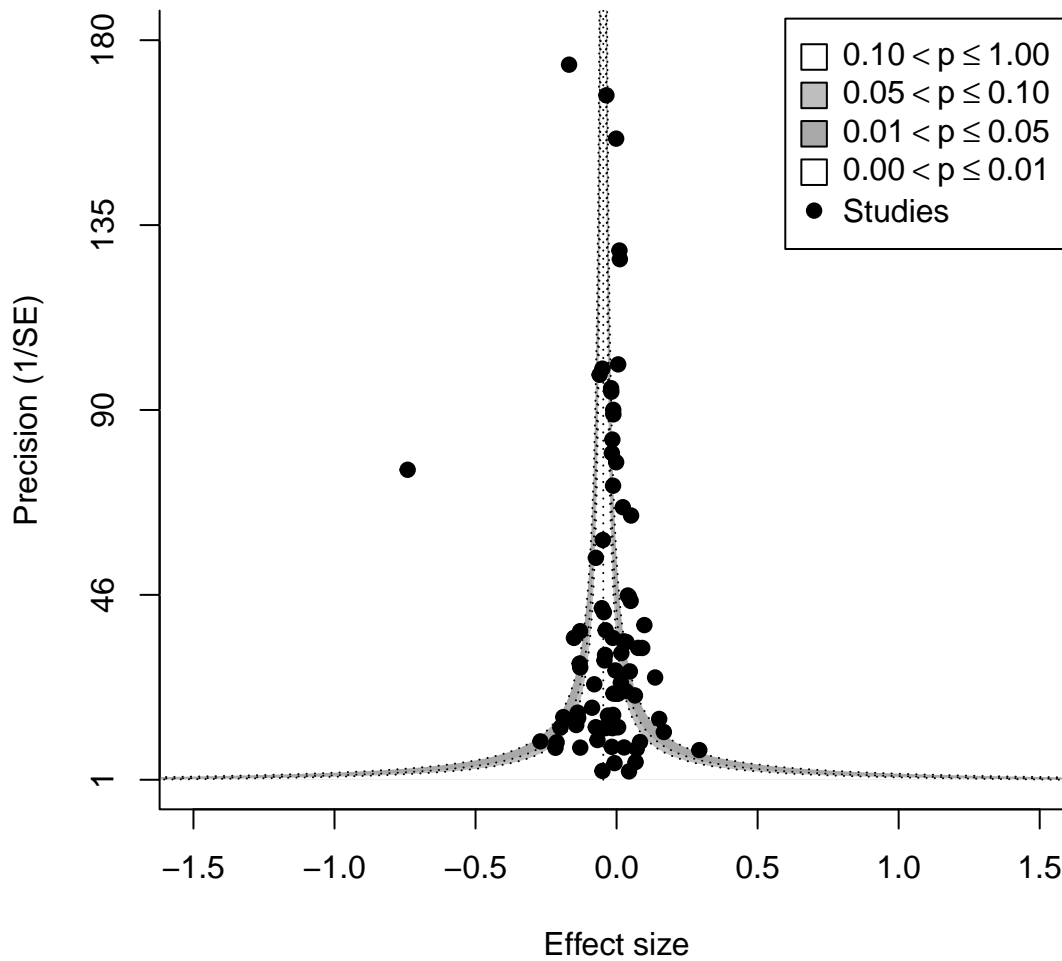
Publication bias

Figure 6

Funnel plots.

```r
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")
```

a) Survey response


```
f1
```

```
##          x          y slab
## 1  -0.7407  76.028592    1
## 2  -0.2699  10.283082    2
## 3  -0.2166   8.739376    3
## 4  -0.2125  10.051901    4
## 5  -0.1990  13.633547    5
## 6  -0.1889  16.160594    6
## 7  -0.1685 174.077656    7
## 8  -0.1511  35.333263    8
## 9  -0.1500  15.625000    9
## 10 -0.1429  14.236408   10
## 11 -0.1377  17.246507   11
## 12 -0.1361  15.971914   12
## 13 -0.1309  29.123469   13
## 14 -0.1292  36.961064   14
## 15 -0.1289  28.182994   15
## 16 -0.1288   8.772267   16
## 17 -0.0869  18.433375   17
## 18 -0.0791  24.112141   18
## 19 -0.0737  13.719241   19
## 20 -0.0734  54.717566   20
```

```
## 21 -0.0678  10.613702   21
## 22 -0.0601  99.014754   22
## 23 -0.0522  42.524326   23
## 24 -0.0502   3.132688   24
## 25 -0.0502 100.503782   25
## 26 -0.0487  59.028134   26
## 27 -0.0467  13.492586   27
## 28 -0.0452  41.558582   28
## 29 -0.0429  29.880715   29
## 30 -0.0408  31.204324   30
## 31 -0.0393  13.492586   31
## 32 -0.0389  37.190400   32
## 33 -0.0360 166.666667   33
## 34 -0.0305  16.536269   34
## 35 -0.0268  13.719241   35
## 36 -0.0197  13.719241   36
## 37 -0.0195  95.782629   37
## 38 -0.0186  94.915800   38
## 39 -0.0169   8.960056   39
## 40 -0.0166  80.064077   40
## 41 -0.0149  83.333333   41
## 42 -0.0143  13.493815   42
## 43 -0.0135  35.333263   43
## 44 -0.0125  72.168784   44
## 45 -0.0122  16.694514   45
## 46 -0.0118  90.535746   46
## 47 -0.0113  89.442719   47
## 48 -0.0112  21.853030   48
## 49 -0.0073   5.026776   49
## 50 -0.0040  27.451403   50
## 51 -0.0025  21.853030   51
## 52 -0.0015  77.849894   52
## 53 -0.0015 156.173762   53
## 54  0.0033  21.853030   54
## 55  0.0035  13.749033   55
## 56  0.0055  13.719241   56
## 57  0.0057 101.534617   57
## 58  0.0059  21.853030   58
## 59  0.0101 129.099445   59
## 60  0.0116 127.000127   60
## 61  0.0152  24.477769   61
## 62  0.0170  31.654447   62
## 63  0.0221  66.964953   63
## 64  0.0252  34.503278   64
## 65  0.0266   8.823339   65
## 66  0.0333  22.456323   66
## 67  0.0347  34.340141   67
## 68  0.0399  45.596075   68
## 69  0.0435  45.314325   69
## 70  0.0441   3.025958   70
## 71  0.0470  27.216553   71
## 72  0.0499  44.280744   72
## 73  0.0514  64.956980   73
## 74  0.0651  21.398025   74
## 75  0.0675   5.291968   75
## 76  0.0716   8.515960   76
```

```
## 77  0.0762  32.915401    77
## 78  0.0831  10.137797    78
## 79  0.0909  32.897585    79
## 80  0.0985  38.433122    80
## 81  0.1366  25.751310    81
## 82  0.1511  15.723193    82
## 83  0.1676  12.559009    83
## 84  0.2933   8.172324    84
```

```
ggsave("figure6a.png")
```

```
## Saving 6 x 6 in image
```

```
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")
```


b) Article output
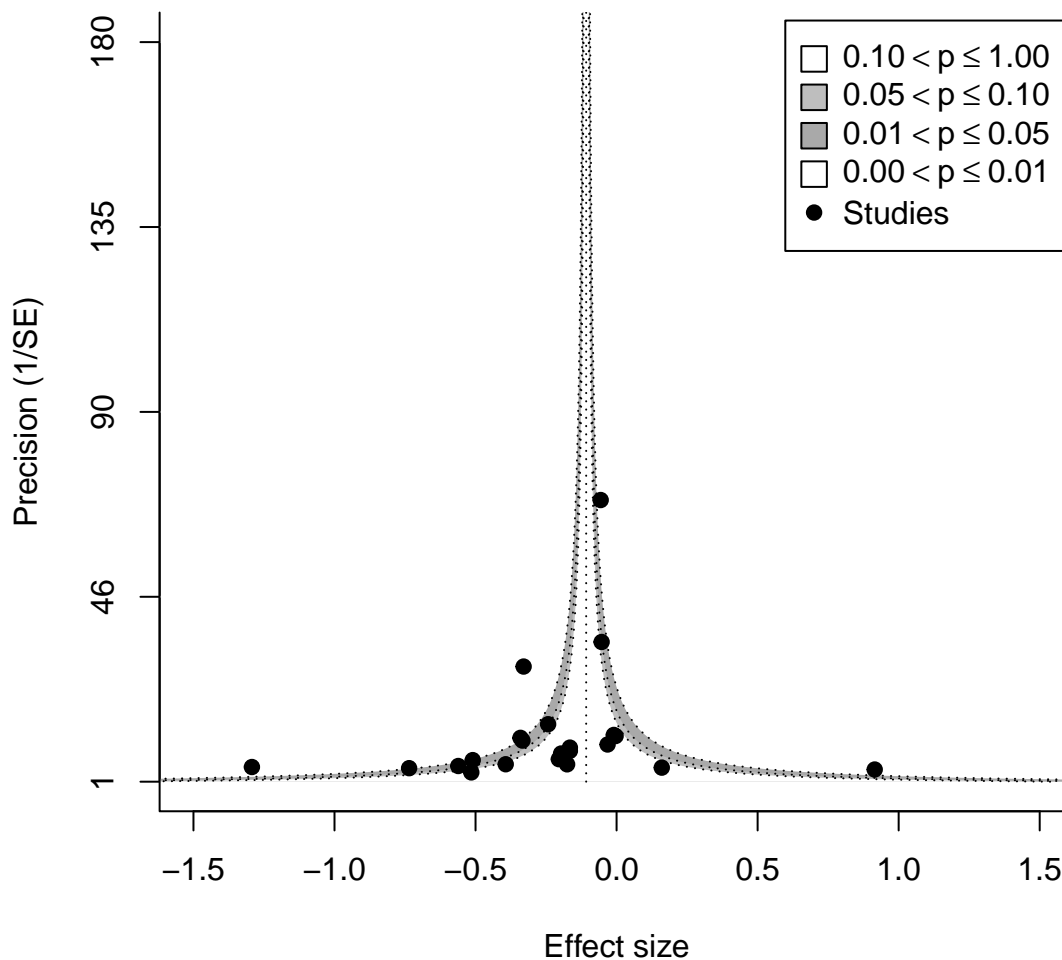
```
f2
```

```
##              x        y slab
```

```
## 1   -1.2928000   4.550676    1
## 2   -0.7355000   4.271788    2
## 3   -0.5612000   4.794633    3
## 4   -0.5150000   3.246753    4
## 5   -0.5100000   6.216579    5
## 6   -0.3930000   5.219958    6
## 7   -0.3407000  11.610651    7
## 8   -0.3337000  10.953356    8
## 9   -0.3296000  28.867513    9
## 10  -0.2427000  14.907120   10
## 11  -0.2046802   6.448692   11
## 12  -0.1966000   7.814408   12
## 13  -0.1895000   6.827410   13
## 14  -0.1754000   5.202832   14
## 15  -0.1655000   8.463642   15
## 16  -0.1655000   9.248957   16
## 17  -0.0563000  69.171446   17
## 18  -0.0530000  34.813499   18
## 19  -0.0319000  10.000000   19
## 20  -0.0098000  12.309149   20
## 21  -0.0035000  12.047318   21
## 22   0.1602000   4.400805   22
## 23   0.9151000   3.943615   23
```

```
ggsave("figure6b.png")
```

```
## Saving 6 x 6 in image
```

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

```
# 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)
```

```
##
## Multivariate Meta-Analysis Model (k = 115; method: REML)
##
##    logLik   Deviance        AIC        BIC       AICc
##   51.6769  -103.3538   -93.3538   -79.7613   -92.7877
##
## Variance Components:
##
##              estim    sqrt  nlvls  fixed                          factor
```

```
## sigma^2.1  0.0000  0.0000    50     no                    ID.article
## sigma^2.2  0.0130  0.1141   115     no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 112) = 4688.7755, p-val < .0001
##
## Test of Moderators (coefficients 1:3):
## QM(df = 3) = 41.5100, p-val < .0001
##
## Model Results:
##
##                                     estimate       se      zval      pval
## Variance.as.standard.error           -0.2643   0.3047   -0.8674    0.3857
## Self.reported.or.measuredMeasured    -0.0253   0.0178   -1.4159    0.1568
## Self.reported.or.measuredSelf-reported  -0.1731   0.0463   -3.7380    0.0002
##                                       ci.lb     ci.ub
## Variance.as.standard.error           -0.8616    0.3329
## Self.reported.or.measuredMeasured    -0.0603    0.0097
## Self.reported.or.measuredSelf-reported  -0.2638   -0.0823   ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```