

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

Kiran Gok Lune Lee

2023-06-12

Install and load packages

```
## install packages like so
install.packages("pacman")
rm(list = ls())
devtools::install_github("daniellnoble/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 required package: numDeriv
##
## Loading the 'metafor' package (version 4.2-0). For an
## introduction to the package please type: help(metafor)
## Loading required package: usethis
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.1      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v lubridate  1.9.2      v tibble    3.2.1
## v purrr      1.0.1      v tidyr     1.3.0
## -- 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()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
##
## Loading the 'orchaRd' package (version 2.0). For an
## introduction and vignette to the package please see: https://daniellnoble.github.io/orchaRd/
##
##
##
## Attaching package: 'gridExtra'
##
##
## The following object is masked from 'package:dplyr':
##
##   combine
```

Load data

```
## Warning: Unknown or uninitialised column:
```

```
## `Effect.size.kiran.use..from.Campbell.collaboration.`.
```

```
## < table of extent 0 >
```

```
##
```

```
##      Measured Self-reported
```

```
##      107          23
```

```
##
```

```
## Biological sciences      Medicine      Multidisciplinary      Social sciences
```

```
##          18          50          17          21
```

```
##          TEMCP
```

```
##          24
```

```
##
```

```
##      Other Publications      Submissions
```

```
##      18          48          64
```

```
##
```

```
##          Burn-out      General productivity
```

```
##          1          11
```

```
##          Job loss      Preprints
```

```
##          1          20
```

```
##          Projects      Publications
```

```
##          1          48
```

```
##          Research time      Submissions
```

```
##          4          38
```

```
## Submissions (self-reported)      Working papers
```

```
##          5          1
```

```
##
```

```
##          Ability to submit papers
```

```
##          1
```

```
##          Academic job loss
```

```
##          1
```

```
##          Academic productivity
```

```
##          1
```

```
##          Any authorship
```

```
##          14
```

```
##          Binary whether loss of research productivity
```

```
##          1
```

```
##          Corresponding authorship
```

```
##          15
```

```
##          Disruption from having to work from home
```

```
##          1
```

```
##          Effect of COVID-19 pandemic on work from home
```

```
##          1
```

```
##          First authorship
```

```
##          55
```

```
##          Hours of research per day
```

```
##          1
```

```
##          Last authorship
```

```
##          18
```

```
##          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

## Warning: NAs introduced by coercion
## [1] 10.80097
## [1] 9.884217
## [1] 1 50
## Warning: NAs introduced by coercion
## [1] 6.635922
## [1] 4.576793
## [1] 1 17
## Warning: NAs introduced by coercion
## [1] "2020-08-21"
## [1] 98.75268
## [1] "2020-04-20" "2021-02-28"
## Time difference of 314 days

##
## Europe Iberico- America North America
## 3 3 18
## North America and Europe Worldwide
## 7 99

##
## Europe North America North America and Europe
## 1 8 1
## Worldwide
## 97

```

##			
##	Europe	Iberico- America	North America
##	2	3	10
##	North America and Europe	Worldwide	
##	6	2	

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 = 130; method: REML)
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0013  0.0366    55    no      ID.article
## sigma^2.2  0.0135  0.1163   130    no ID.article/ID.observation
##
## Test for Heterogeneity:
## Q(df = 129) = 4939.2102, p-val < .0001
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## -0.0706  0.0144  -4.9170  <.0001  -0.0988  -0.0425  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(m)
```

```
##
## Multivariate Meta-Analysis Model (k = 130; method: REML)
##
##      logLik      Deviance      AIC      BIC      AICc
##  52.3984 -104.7968  -98.7968  -90.2174  -98.6048
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0013  0.0366    55    no      ID.article
## sigma^2.2  0.0135  0.1163   130    no ID.article/ID.observation
##
## Test for Heterogeneity:
## Q(df = 129) = 4939.2102, p-val < .0001
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## -0.0706  0.0144  -4.9170  <.0001  -0.0988  -0.0425  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Forest plot produced for visual representation.
png(file = "forest.png")
forest(m, slab = all_data$Author, xlim = c(-2, 2), ylim = c(-1,
  140), digits = 2, xlab = "Raw proportion", mlab = "Overall effect (46)",
  cex = 0.4)
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)
while (!is.null(dev.list())) dev.off()
```

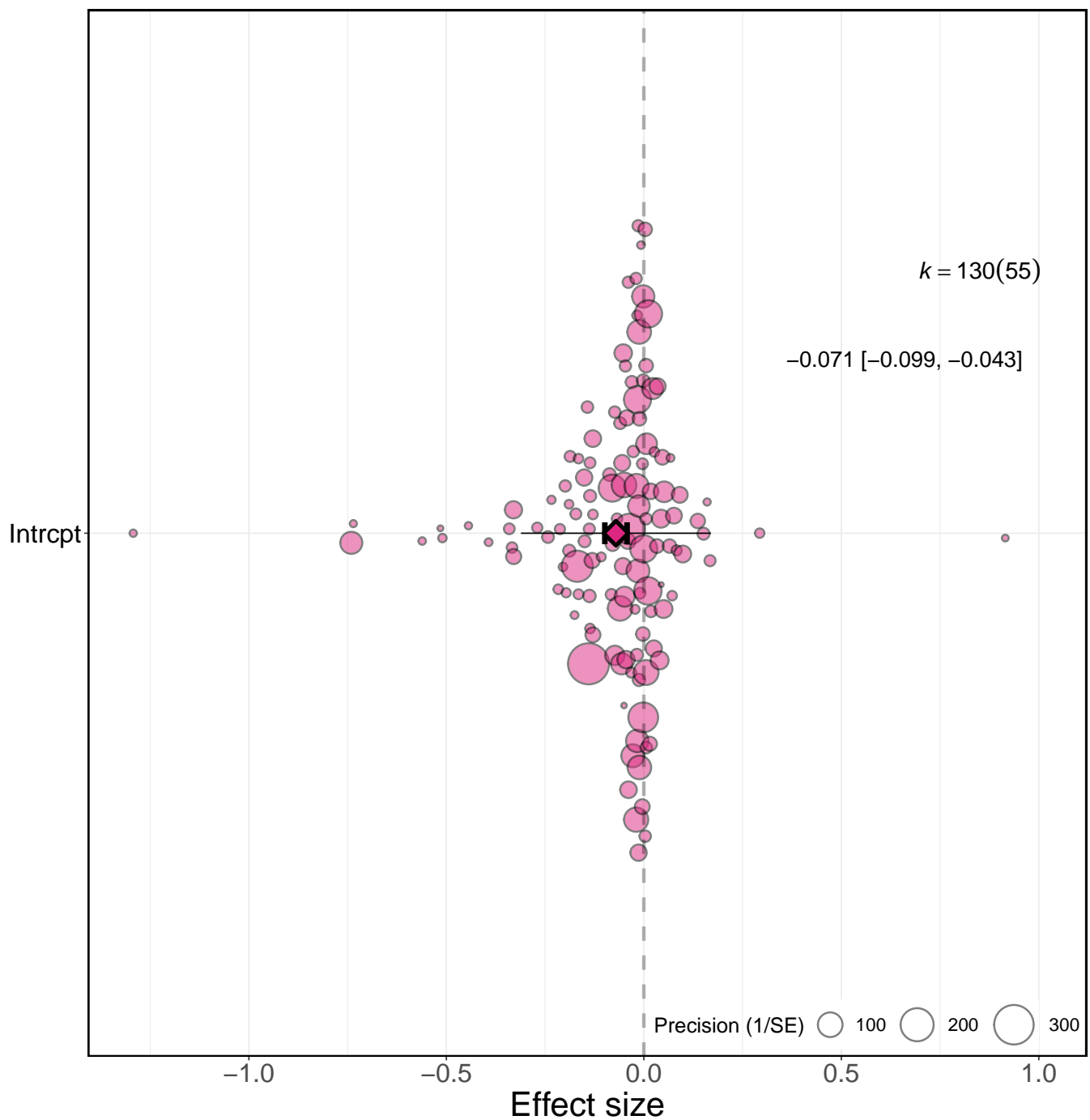
Figure 1

```
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 = 0.66, x = 1.2, label = paste("-0.071 [-0.099, -0.043]")) +
  scale_fill_manual(values = c("#DC267F"))

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## Warning: The `size` argument of `element_rect()` is deprecated as of ggplot2 3.4.0.
## i Please use the `linewidth` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## 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	64	18

```
all_data$Broad.productivity.measure.reordered <- factor(all_data$Broad.productivity.measure,
  c("Submissions", "Publications", "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 = 130; method: REML)
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0002  0.0135    55    no      ID.article
## sigma^2.2  0.0138  0.1176   130    no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 127) = 4900.4414, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 11.1700, p-val = 0.0038
##
## Model Results:
##
##              estimate      se      zval
## intrcpt          -0.0525  0.0174  -3.0142
## Broad.productivity.measure.reorderedPublications    0.0057  0.0262   0.2183
## Broad.productivity.measure.reorderedSurvey          -0.1398  0.0442  -3.1623
##
##              pval      ci.lb      ci.ub
## intrcpt          0.0026  -0.0867  -0.0184  **
## Broad.productivity.measure.reorderedPublications    0.8272  -0.0456   0.0570
## Broad.productivity.measure.reorderedSurvey          0.0016  -0.2265  -0.0532  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(m.area)
```

```
##
## Multivariate Meta-Analysis Model (k = 130; method: REML)
##
##      logLik  Deviance      AIC      BIC      AICc
##  55.8510  -111.7020  -101.7020  -87.4811  -101.2062
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0002  0.0135    55    no      ID.article
## sigma^2.2  0.0138  0.1176   130    no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 127) = 4900.4414, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 11.1700, p-val = 0.0038
##
## Model Results:
```



```
##
##
##               estimate      se      zval
## intrcpt      -0.0525  0.0174  -3.0142
## Broad.productivity.measure.reorderedPublications  0.0057  0.0262   0.2183
## Broad.productivity.measure.reorderedSurvey      -0.1398  0.0442  -3.1623
##
##               pval      ci.lb      ci.ub
## intrcpt      0.0026  -0.0867  -0.0184  **
## Broad.productivity.measure.reorderedPublications  0.8272  -0.0456   0.0570
## Broad.productivity.measure.reorderedSurvey      0.0016  -0.2265  -0.0532  **
##
## ---
## 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

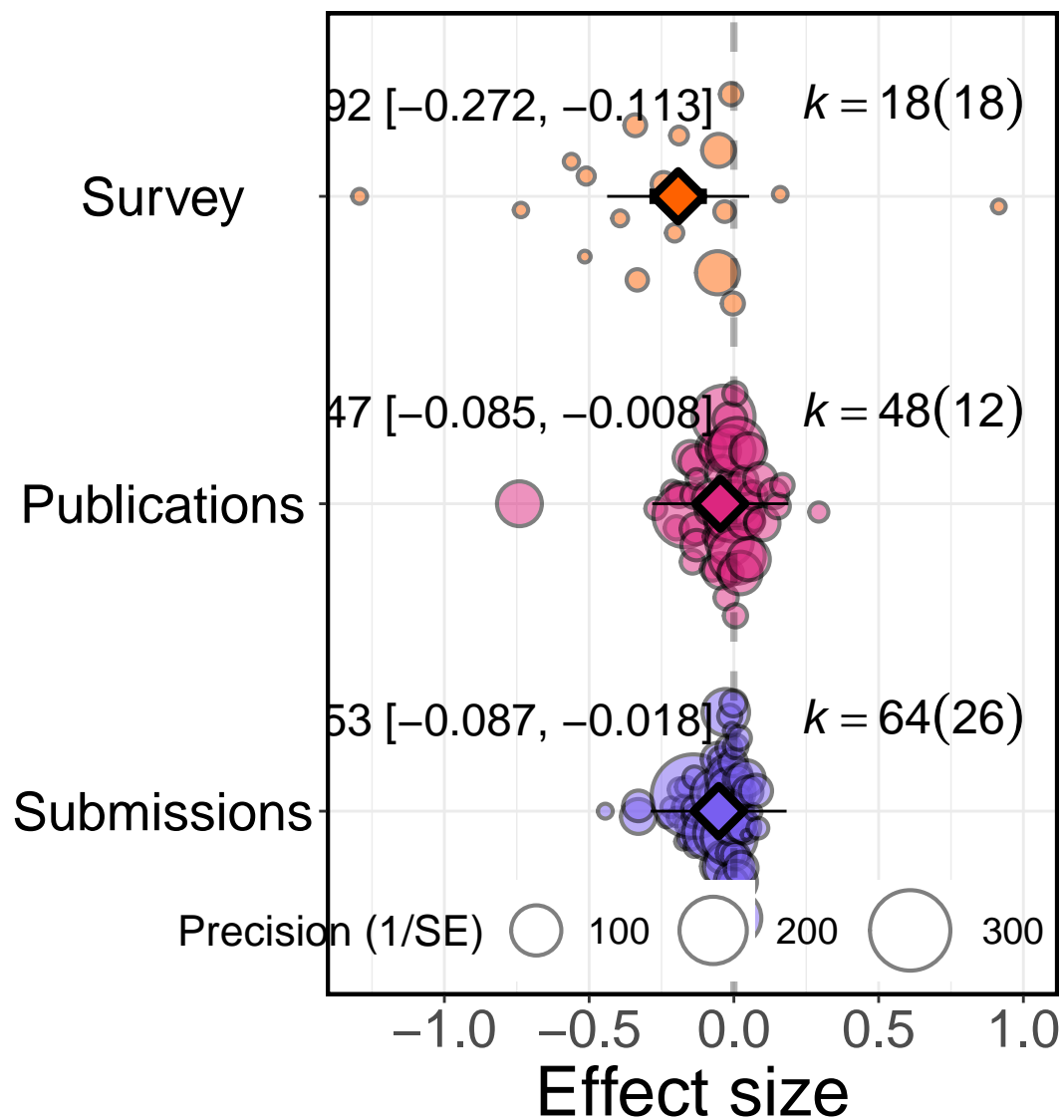
##
## Multivariate Meta-Analysis Model (k = 130; method: REML)
##
## Variance Components:
##
##               estim      sqrt  nlvls  fixed              factor
## sigma^2.1  0.0002  0.0135    55    no              ID.article
## sigma^2.2  0.0138  0.1176   130    no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 127) = 4900.4414, p-val < .0001
##
## Test of Moderators (coefficients 1:3):
## QM(df = 3) = 37.1300, p-val < .0001
##
## Model Results:
##
##               estimate      se      zval
## Broad.productivity.measure.reorderedSubmissions  -0.0525  0.0174  -3.0142
## Broad.productivity.measure.reorderedPublications  -0.0468  0.0196  -2.3879
## Broad.productivity.measure.reorderedSurvey      -0.1924  0.0406  -4.7339
##
##               pval      ci.lb      ci.ub
## Broad.productivity.measure.reorderedSubmissions  0.0026  -0.0867  -0.0184  **
## Broad.productivity.measure.reorderedPublications  0.0169  -0.0853  -0.0084   *
## Broad.productivity.measure.reorderedSurvey      <.0001  -0.2721  -0.1127  ***
##
## ---
## 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 = -0.9, x = 3.3, label = paste("-0.192 [-0.272, -0.113]")) +
  annotate("text", size = 6, y = -0.9, x = 2.3, label = paste("-0.047 [-0.085, -0.008]")) +
  annotate("text", size = 6, y = -0.9, x = 1.3, label = paste("-0.053 [-0.087, -0.018]")) +
  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", width = 10, height = 10, dpi = 300)
```

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("Biological sciences", "TEMCP", "Medicine", "Social sciences",
    "Multidisciplinary"))

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 = 130; method: REML)
##
##      logLik   Deviance      AIC      BIC      AICc
##  54.0441  -108.0883   -94.0883   -74.2901   -93.1310
##
## Variance Components:
##
##           estim      sqrt nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000    55    no  Article.ID
## sigma^2.2  0.0131  0.1146   130    no           ID
##
## Test for Residual Heterogeneity:
## QE(df = 125) = 4058.9132, p-val < .0001
##
## Test of Moderators (coefficients 2:5):
## QM(df = 4) = 12.9827, p-val = 0.0114
##
## Model Results:
##
##                                     estimate      se      zval
## intrcpt                          0.0067  0.0293  0.2282
## Broad.research.field.reorderedTEMCP    -0.0238  0.0393 -0.6050
## Broad.research.field.reorderedMedicine  -0.0897  0.0352 -2.5512
## Broad.research.field.reorderedSocial\nsciences -0.1027  0.0418 -2.4582
## Broad.research.field.reorderedMultidisciplinary -0.1158  0.0438 -2.6446
##                                     pval      ci.lb      ci.ub
## intrcpt                          0.8195  -0.0507  0.0641
## Broad.research.field.reorderedTEMCP    0.5452  -0.1009  0.0533
## Broad.research.field.reorderedMedicine  0.0107  -0.1586 -0.0208  *
## Broad.research.field.reorderedSocial\nsciences 0.0140  -0.1846 -0.0208  *
## Broad.research.field.reorderedMultidisciplinary 0.0082  -0.2016 -0.0300  **
##
## ---
## 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 = 130; method: REML)
##
##      logLik    Deviance      AIC      BIC      AICc
##    54.0441  -108.0883   -94.0883   -74.2901   -93.1310
##
## Variance Components:
##
##           estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000    55    no  Article.ID
## sigma^2.2  0.0131  0.1146   130    no          ID
##
## Test for Residual Heterogeneity:
## QE(df = 125) = 4058.9132, p-val < .0001
##
## Test of Moderators (coefficients 1:5):
## QM(df = 5) = 40.2623, p-val < .0001
##
## Model Results:
##
##                                     estimate      se      zval
## Broad.research.field.reorderedBiological\nsciences    0.0067  0.0293  0.2282
## Broad.research.field.reorderedTEMCP                   -0.0171  0.0263 -0.6518
## Broad.research.field.reorderedMedicine                 -0.0830  0.0195 -4.2628
## Broad.research.field.reorderedSocial\nsciences         -0.0960  0.0298 -3.2213
## Broad.research.field.reorderedMultidisciplinary        -0.1091  0.0325 -3.3522
##
##                                     pval      ci.lb      ci.ub
## Broad.research.field.reorderedBiological\nsciences    0.8195  -0.0507  0.0641
## Broad.research.field.reorderedTEMCP                   0.5145  -0.0686  0.0343
## Broad.research.field.reorderedMedicine                 <.0001  -0.1212 -0.0449
## Broad.research.field.reorderedSocial\nsciences         0.0013  -0.1545 -0.0376
## Broad.research.field.reorderedMultidisciplinary        0.0008  -0.1729 -0.0453
##
## Broad.research.field.reorderedBiological\nsciences
## Broad.research.field.reorderedTEMCP
## Broad.research.field.reorderedMedicine                ***
## Broad.research.field.reorderedSocial\nsciences         **
## Broad.research.field.reorderedMultidisciplinary        ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Prediction 2a Measured studies: 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.

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

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

research_fieldm <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field.reordered,
  random = list(~1 | Article.ID, ~1 | ID), data = measured_data)
summary(research_fieldm)
```

```
##
```

```
## Multivariate Meta-Analysis Model (k = 107; method: REML)
##
##      logLik    Deviance      AIC      BIC      AICc
##    71.7173  -143.4347  -129.4347  -111.0599  -128.2432
##
## Variance Components:
##
##           estim    sqrt  nlvls  fixed    factor
## sigma^2.1  0.0000  0.0000    34    no  Article.ID
## sigma^2.2  0.0109  0.1046   107    no      ID
##
## Test for Residual Heterogeneity:
## QE(df = 102) = 3906.0102, p-val < .0001
##
## Test of Moderators (coefficients 2:5):
## QM(df = 4) = 7.9479, p-val = 0.0935
##
## Model Results:
##
##                                     estimate      se      zval
## intrcpt                           -0.0027  0.0256  -0.1052
## Broad.research.field.reorderedBiological\nciences -0.0007  0.0373  -0.0176
## Broad.research.field.reorderedMultidisciplinary  -0.0473  0.0458  -1.0329
## Broad.research.field.reorderedMedicine           -0.0629  0.0316  -1.9893
## Broad.research.field.reorderedSocial\nciences    -0.0812  0.0397  -2.0444
##
##                                     pval      ci.lb      ci.ub
## intrcpt                           0.9162  -0.0529  0.0475
## Broad.research.field.reorderedBiological\nciences 0.9859  -0.0738  0.0725
## Broad.research.field.reorderedMultidisciplinary  0.3017  -0.1372  0.0425
## Broad.research.field.reorderedMedicine           0.0467  -0.1248  -0.0009  *
## Broad.research.field.reorderedSocial\nciences    0.0409  -0.1591  -0.0034  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

research_field1m <- metafor::rma.mv(yi = yi, V = vi, mods = ~Broad.research.field.reordered -
  1, random = list(~1 | Article.ID, ~1 | ID), data = measured_data)

summary(research_field1m)

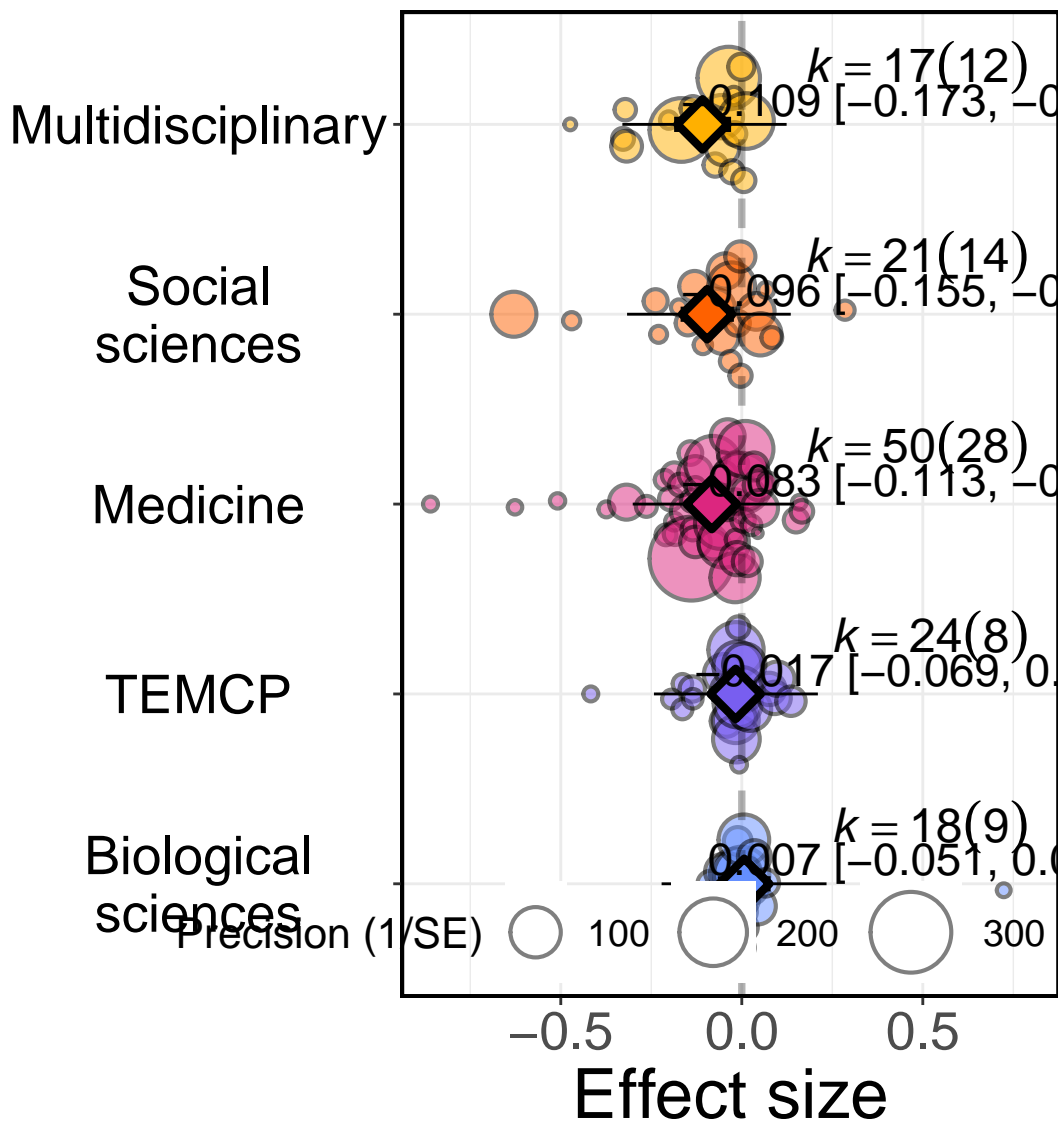
##
## Multivariate Meta-Analysis Model (k = 107; method: REML)
##
##      logLik    Deviance      AIC      BIC      AICc
##    71.7173  -143.4347  -129.4347  -111.0599  -128.2432
##
## Variance Components:
##
##           estim    sqrt  nlvls  fixed    factor
## sigma^2.1  0.0000  0.0000    34    no  Article.ID
## sigma^2.2  0.0109  0.1046   107    no      ID
##
## Test for Residual Heterogeneity:
## QE(df = 102) = 3906.0102, p-val < .0001
##
## Test of Moderators (coefficients 1:5):
## QM(df = 5) = 21.9667, p-val = 0.0005
```

```
##
## Model Results:
##
##
##               estimate      se      zval
## Broad.research.field.reorderedTEMCP      -0.0027  0.0256  -0.1052
## Broad.research.field.reorderedBiological\nsciences -0.0034  0.0271  -0.1236
## Broad.research.field.reorderedMultidisciplinary -0.0500  0.0380  -1.3167
## Broad.research.field.reorderedMedicine      -0.0656  0.0185  -3.5450
## Broad.research.field.reorderedSocial\nsciences -0.0839  0.0304  -2.7640
##
##               pval      ci.lb      ci.ub
## Broad.research.field.reorderedTEMCP      0.9162  -0.0529  0.0475
## Broad.research.field.reorderedBiological\nsciences 0.9017  -0.0565  0.0498
## Broad.research.field.reorderedMultidisciplinary 0.1879  -0.1245  0.0244
## Broad.research.field.reorderedMedicine      0.0004  -0.1018  -0.0293
## Broad.research.field.reorderedSocial\nsciences 0.0057  -0.1434  -0.0244
##
## Broad.research.field.reorderedTEMCP
## Broad.research.field.reorderedBiological\nsciences
## Broad.research.field.reorderedMultidisciplinary
## Broad.research.field.reorderedMedicine      ***
## 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 = 5.13, label = paste("-0.109 [-0.173, -0.045]")) +
  annotate("text", size = 6, y = 0.5, x = 4.13, label = paste("-0.096 [-0.155, -0.038]")) +
  annotate("text", size = 6, y = 0.5, x = 3.13, label = paste("-0.083 [-0.113, -0.031]")) +
  annotate("text", size = 6, y = 0.5, x = 2.13, label = paste("-0.017 [-0.069, 0.034]")) +
  annotate("text", size = 6, y = 0.5, x = 1.13, label = paste("0.007 [-0.051, 0.064]")) +
  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)'
```

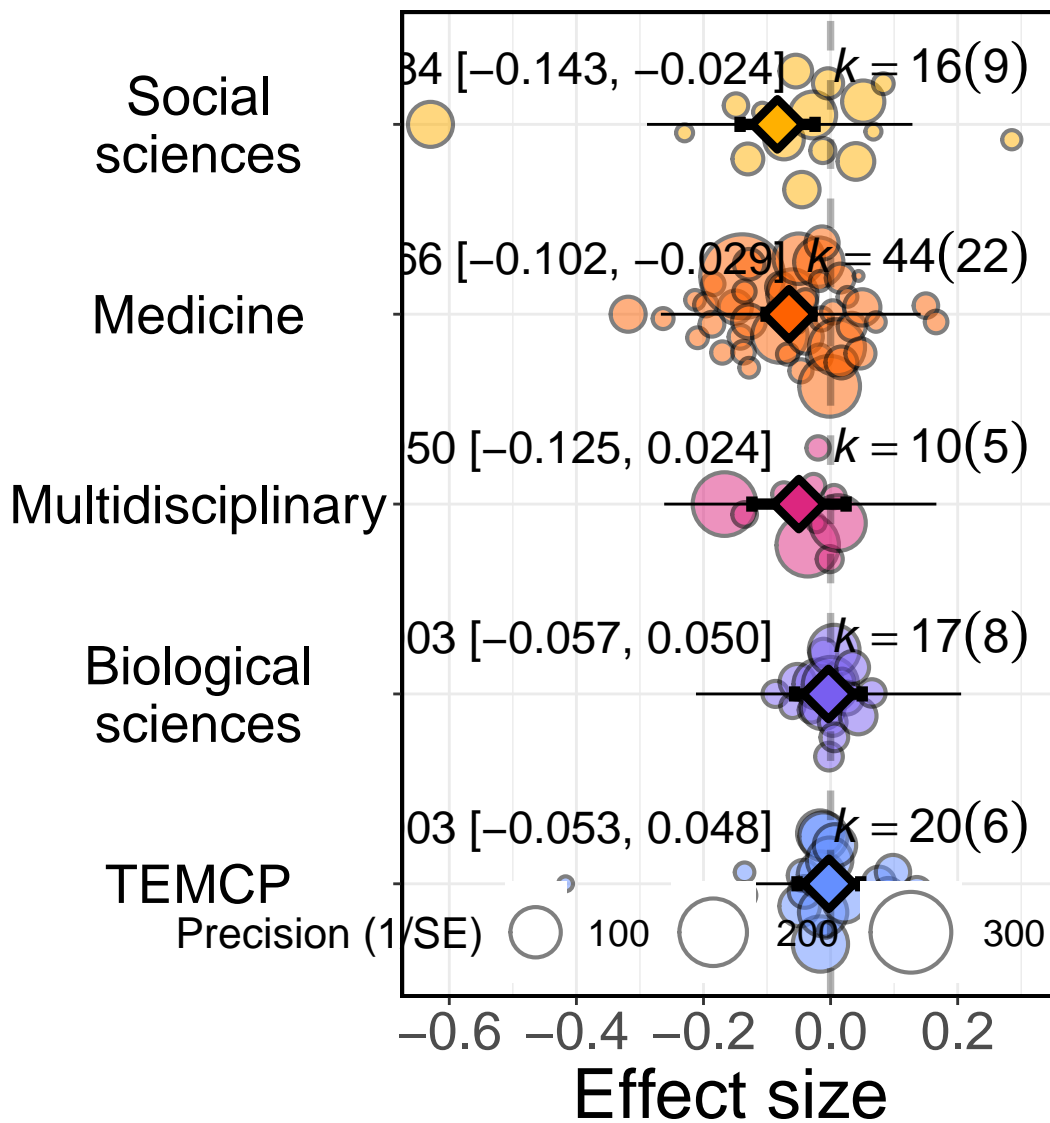


```
ggsave("figure3.png", width = 10, height = 10, dpi = 300)
```

Figure 3 Measured data

```
my.orchard(research_field1m, mod = "Broad.research.field.reordered",
  group = "Article.ID", data = measured_data, whisker = 0.08,
  xlab = "Effect size", alpha = 0.5, transfm = "tanh", cb = FALSE) +
  annotate("text", size = 6, y = -0.45, x = 5.3, label = paste("-0.084 [-0.143, -0.024]")) +
  annotate("text", size = 6, y = -0.45, x = 4.3, label = paste("-0.066 [-0.102, -0.029]")) +
  annotate("text", size = 6, y = -0.45, x = 3.3, label = paste("-0.050 [-0.125, 0.024]")) +
  annotate("text", size = 6, y = -0.45, x = 2.3, label = paste("-0.003 [-0.057, 0.050]")) +
  annotate("text", size = 6, y = -0.45, x = 1.3, label = paste("-0.003 [-0.053, 0.048]")) +
  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("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$nmnenprepandemic <- round(as.numeric(all_data$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
all_data$nmenduringpandemic <- round(as.numeric(all_data$N.during.pandemic) -
  as.numeric(all_data$nwomenduringpandemic), digits = 0)

## Warning: NAs introduced by coercion
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: 8 rows with NAs omitted from model fitting.
# 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 = 99; method: REML)
##
##      logLik   Deviance      AIC      BIC      AICc
##      67.3853  -134.7706  -126.7706  -116.4718  -126.3358
##
## Variance Components:
##
##           estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000     30     no  Article.ID
## sigma^2.2  0.0113  0.1061     99     no           ID
##
## Test for Residual Heterogeneity:
## QE(df = 97) = 3576.7319, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 10.2852, p-val = 0.0013
##
## Model Results:
##
##              estimate      se      zval      pval
## intrcpt          -0.1230  0.0282  -4.3559  <.0001
## cbind(nmenprepandemic/nwomenprepandemic)  0.0346  0.0108   3.2071  0.0013
##              ci.lb      ci.ub
## intrcpt          -0.1783  -0.0676  ***
## cbind(nmenprepandemic/nwomenprepandemic)  0.0135  0.0557   **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

measured_data_numbers <- measured_data
measured_data_numbers$X..women.authors.pre.pandemic <- as.numeric(measured_data_numbers$X..women.authors.p

## Warning: NAs introduced by coercion
measured_data_numbers$X..women.authors.during.pandemic <- as.numeric(measured_data_numbers$X..women.authors

## Warning: NAs introduced by coercion
measured_data_numbers <- subset(measured_data_numbers, !is.na(measured_data_numbers$X..women.authors.pre.p
measured_data_numbers <- subset(measured_data_numbers, !is.na(measured_data_numbers$X..women.authors.durin
socialsciences <- subset(measured_data_numbers, measured_data_numbers$Broad.research.field ==

```

```

  "Social sciences")
medicine <- subset(measured_data_numbers, measured_data_numbers$Broad.research.field ==
  "Medicine")
multi <- subset(measured_data_numbers, measured_data_numbers$Broad.research.field ==
  "Multidisciplinary")
temcp <- subset(measured_data_numbers, measured_data_numbers$Broad.research.field ==
  "TEMCP")
bio <- subset(measured_data_numbers, measured_data_numbers$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))

## [1] 0.3582624
mean((socialsciences$X..women.authors.during.pandemic))

## [1] 0.3340385
mean((medicine$X..women.authors.pre.pandemic))

## [1] 0.3656482
mean((medicine$X..women.authors.during.pandemic))

## [1] 0.3375599
mean((multi$X..women.authors.pre.pandemic))

## [1] 0.3621285
mean((multi$X..women.authors.during.pandemic))

## [1] 0.3416098
mean((temcp$X..women.authors.pre.pandemic))

## [1] 0.22973
mean((temcp$X..women.authors.during.pandemic))

## [1] 0.2206314
mean((bio$X..women.authors.pre.pandemic))

## [1] 0.326504
mean((bio$X..women.authors.during.pandemic))

## [1] 0.3274816
mean((measured_data_numbers$X..women.authors.pre.pandemic))

## [1] 0.332232
mean((measured_data_numbers$X..women.authors.during.pandemic))

## [1] 0.3135032
all_data$Broad.research.field.reordered.opposite <- factor(all_data$Broad.research.field,
  c("Social sciences", "Medicine", "Multidisciplinary", "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.pandemic)

## 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")

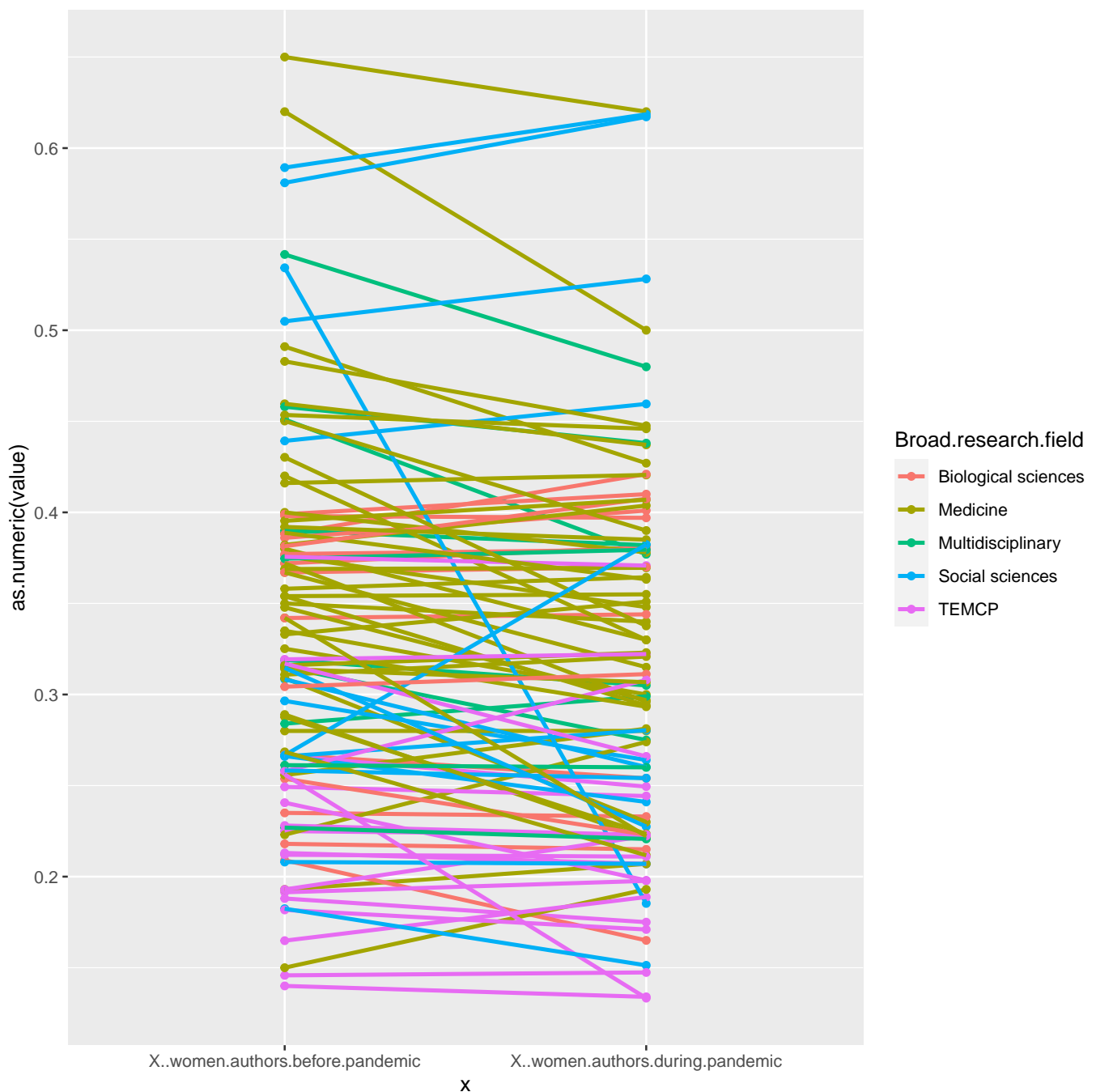
# This plot shows the above model about how previous gender
# bias influence gender bias during the pandemic
ggplot(all_data_long, aes(x = factor(Period), y = as.numeric(value),
  color = factor(Broad.research.field), group = ID)) + geom_point(position = position_jitter(width = 0))
  geom_smooth(method = "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

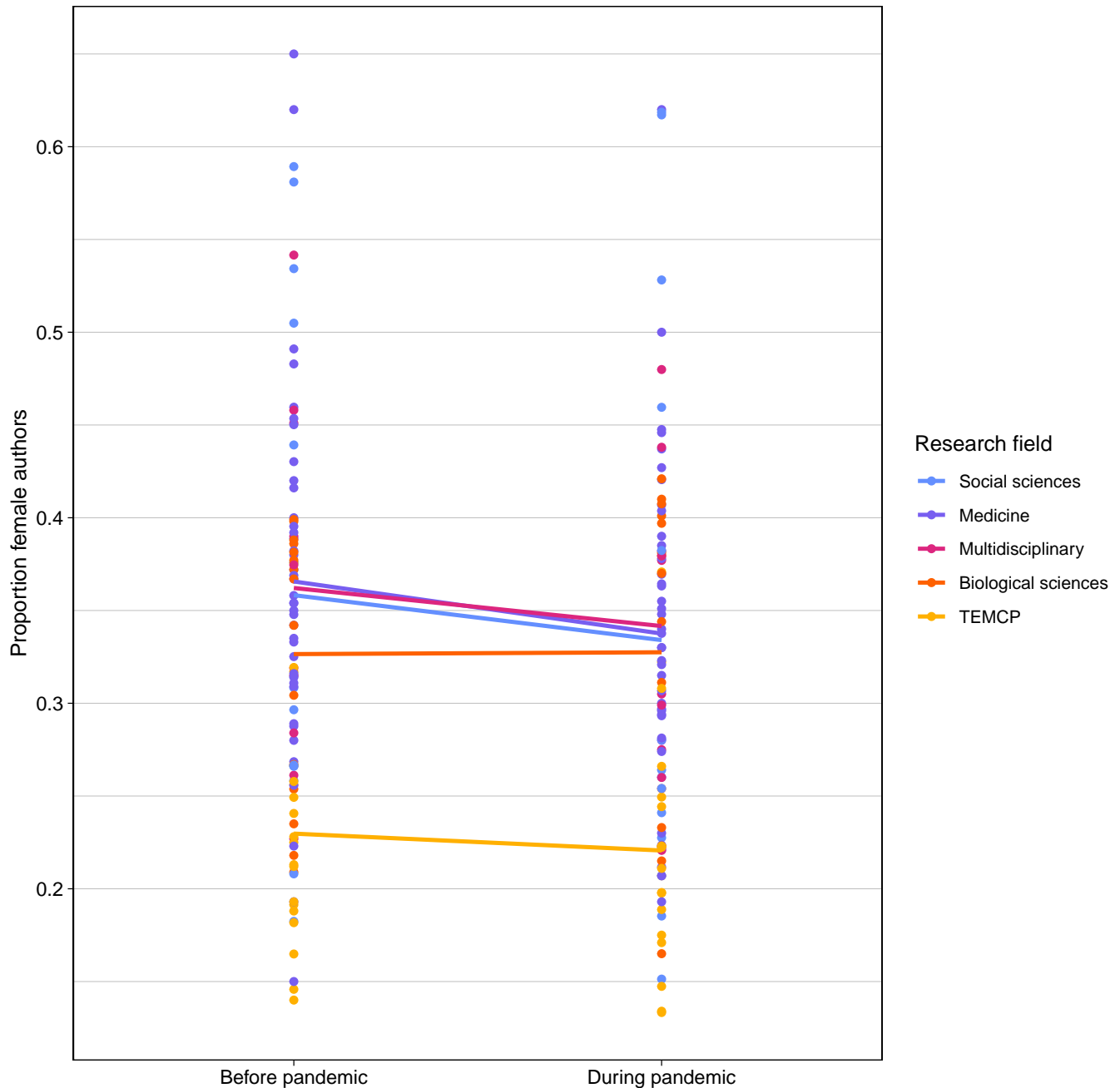
```
ggplot(all_data_long, aes(x = factor(Period), y = as.numeric(value),
  color = factor(Broad.research.field.reordered.opposite),
  group = Broad.research.field.reordered.opposite)) + geom_point(position = position_jitter(width = 0)) +
  geom_smooth(method = "lm", se = FALSE) + labs(color = "Research 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"))
```

```
## Warning: The `size` argument of `element_line()` is deprecated as of ggplot2 3.4.0.
## i Please use the `linewidth` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## `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 = 107; method: REML)
##
##      logLik    Deviance      AIC      BIC      AICc
##  68.0390 -136.0779 -122.0779 -103.7031 -120.8865
##
## Variance Components:
##
##      estim    sqrt  nlvls  fixed    factor
## sigma^2.1  0.0000  0.0000   34    no  Article.ID
## sigma^2.2  0.0121  0.1102  107    no         ID
##
## Test for Residual Heterogeneity:
## QE(df = 102) = 4161.4977, p-val < .0001
##
## Test of Moderators (coefficients 2:5):
## QM(df = 4) = 0.2657, p-val = 0.9919
##
## Model Results:
##
##                                     estimate
## intrcpt                             -0.0396
## Specific.productivity.measure.reorderedMiddle\authorship -0.0058
## Specific.productivity.measure.reorderedFirst\authorship -0.0003
## Specific.productivity.measure.reorderedCorresponding\authorship -0.0187
## Specific.productivity.measure.reorderedAny\authorship -0.0054
##                                     se
## intrcpt                             0.0277
## Specific.productivity.measure.reorderedMiddle\authorship 0.0706
## Specific.productivity.measure.reorderedFirst\authorship 0.0324
## Specific.productivity.measure.reorderedCorresponding\authorship 0.0432
## Specific.productivity.measure.reorderedAny\authorship 0.0420
```

```

##                                                                 zval
## intrcpt                                                         -1.4314
## Specific.productivity.measure.reorderedMiddle\nauthorship      -0.0820
## Specific.productivity.measure.reorderedFirst\nauthorship       -0.0084
## Specific.productivity.measure.reorderedCorresponding\nauthorship -0.4318
## Specific.productivity.measure.reorderedAny\nauthorship          -0.1294
##                                                                 pval
## intrcpt                                                         0.1523
## Specific.productivity.measure.reorderedMiddle\nauthorship      0.9347
## Specific.productivity.measure.reorderedFirst\nauthorship       0.9933
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.6659
## Specific.productivity.measure.reorderedAny\nauthorship          0.8971
##                                                                 ci.lb
## intrcpt                                                         -0.0938
## Specific.productivity.measure.reorderedMiddle\nauthorship      -0.1442
## Specific.productivity.measure.reorderedFirst\nauthorship       -0.0638
## Specific.productivity.measure.reorderedCorresponding\nauthorship -0.1034
## Specific.productivity.measure.reorderedAny\nauthorship          -0.0877
##                                                                 ci.ub
## intrcpt                                                         0.0146
## Specific.productivity.measure.reorderedMiddle\nauthorship      0.1326
## Specific.productivity.measure.reorderedFirst\nauthorship       0.0633
## Specific.productivity.measure.reorderedCorresponding\nauthorship 0.0661
## Specific.productivity.measure.reorderedAny\nauthorship          0.0768
##
## ---
## 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 = 107; method: REML)
##
##      logLik   Deviance      AIC      BIC      AICc
## 68.0390 -136.0779 -122.0779 -103.7031 -120.8865
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000   34    no  Article.ID
## sigma^2.2  0.0121  0.1102  107    no          ID
##
## Test for Residual Heterogeneity:
## QE(df = 102) = 4161.4977, p-val < .0001
##
## Test of Moderators (coefficients 1:5):
## QM(df = 5) = 13.1902, p-val = 0.0217
##
## Model Results:
##
##                                                                 estimate
## Specific.productivity.measure.reorderedLast\nauthorship      -0.0396
## Specific.productivity.measure.reorderedMiddle\nauthorship     -0.0454
## Specific.productivity.measure.reorderedFirst\nauthorship      -0.0399

```

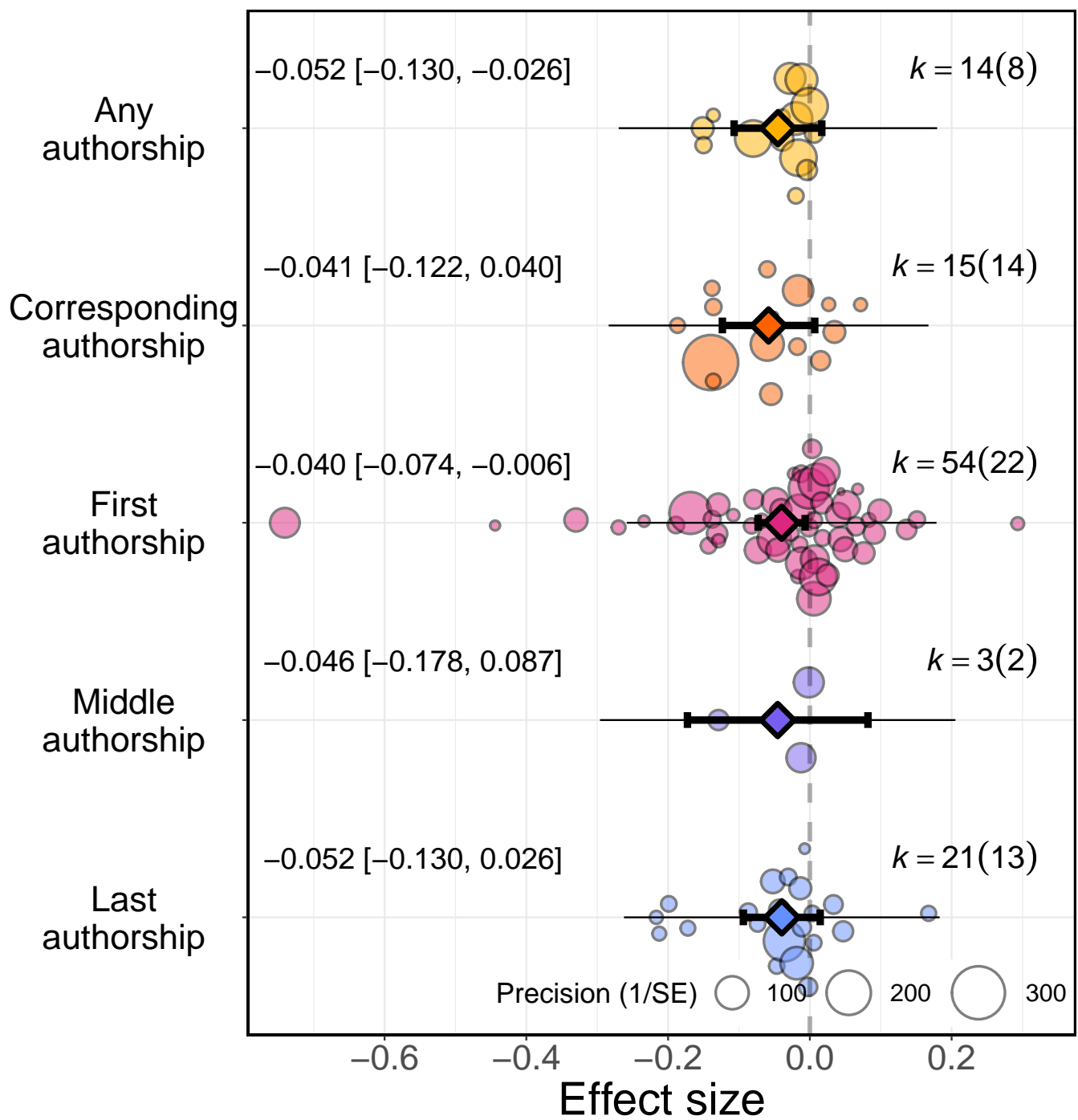


```
## Specific.productivity.measure.reorderedCorresponding\authorship -0.0583
## Specific.productivity.measure.reorderedAny\authorship -0.0450
## se
## Specific.productivity.measure.reorderedLast\authorship 0.0277
## Specific.productivity.measure.reorderedMiddle\authorship 0.0650
## Specific.productivity.measure.reorderedFirst\authorship 0.0169
## Specific.productivity.measure.reorderedCorresponding\authorship 0.0332
## Specific.productivity.measure.reorderedAny\authorship 0.0316
## zval
## Specific.productivity.measure.reorderedLast\authorship -1.4314
## Specific.productivity.measure.reorderedMiddle\authorship -0.6986
## Specific.productivity.measure.reorderedFirst\authorship -2.3547
## Specific.productivity.measure.reorderedCorresponding\authorship -1.7529
## Specific.productivity.measure.reorderedAny\authorship -1.4269
## pval
## Specific.productivity.measure.reorderedLast\authorship 0.1523
## Specific.productivity.measure.reorderedMiddle\authorship 0.4848
## Specific.productivity.measure.reorderedFirst\authorship 0.0185
## Specific.productivity.measure.reorderedCorresponding\authorship 0.0796
## Specific.productivity.measure.reorderedAny\authorship 0.1536
## ci.lb
## Specific.productivity.measure.reorderedLast\authorship -0.0938
## Specific.productivity.measure.reorderedMiddle\authorship -0.1727
## Specific.productivity.measure.reorderedFirst\authorship -0.0731
## Specific.productivity.measure.reorderedCorresponding\authorship -0.1234
## Specific.productivity.measure.reorderedAny\authorship -0.1069
## ci.ub
## Specific.productivity.measure.reorderedLast\authorship 0.0146
## Specific.productivity.measure.reorderedMiddle\authorship 0.0819
## Specific.productivity.measure.reorderedFirst\authorship -0.0067 *
## Specific.productivity.measure.reorderedCorresponding\authorship 0.0069 .
## Specific.productivity.measure.reorderedAny\authorship 0.0168
##
## ---
## 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 = -0.56, x = 1.3, label = paste("-0.052 [-0.130, 0.026]")) +
  annotate("text", size = 6, y = -0.56, x = 2.3, label = paste("-0.046 [-0.178, 0.087]")) +
  annotate("text", size = 6, y = -0.56, x = 3.3, label = paste("-0.040 [-0.074, -0.006]")) +
  annotate("text", size = 6, y = -0.56, x = 4.3, label = paste("-0.041 [-0.122, 0.040]")) +
  annotate("text", size = 6, y = -0.56, x = 5.3, label = paste("-0.052 [-0.130, -0.026]")) +
  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", width = 10, height = 10, dpi = 300)
```

Heterogeneity test

```
##                I2_Total                I2_ID.article
##                97.898495                8.835454
## I2_ID.article/ID.observation
##                89.063041

##                I2_Total                I2_ID.article
##                97.898                8.835
## I2_ID.article/ID.observation
##                89.063

##                I2_Total                I2_ID.article
##                98.7                0.0
##                I2_ID.observation    I2_Self.reported.or.measured
##                52.1                46.6
## I2_Specific.productivity.measure    I2_Broad.productivity.measure
##                0.0                0.0
```

Publication bias

Figure 6

Funnel plots.

```
png("figure6.png", width = 205, height = 205, units = "mm", res = 300)
plot(1:10, 1:10, mar = c(3.5, 0, 3, 0))
par(mfrow = c(2, 1))

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))
f1 <- 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 = "a) Survey responses")

f1
```

```
##          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
```

```
meta_measured <- rma.mv(yi = yi, V = vi, data = as.data.frame(measured_data))
f2 <- 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 = "b) Article output")
f2
```

```
##      x      y slab
## 1 -0.7407 76.028592 1
## 2 -0.4441 4.532131 2
## 3 -0.3300 40.422604 3
## 4 -0.2699 10.283082 4
## 5 -0.2339 6.007875 5
## 6 -0.2166 8.739376 6
## 7 -0.2125 10.051901 7
## 8 -0.1990 13.633547 8
## 9 -0.1889 16.160594 9
## 10 -0.1867 12.334404 10
## 11 -0.1721 12.085971 11
## 12 -0.1685 174.077656 12
## 13 -0.1511 35.333263 13
## 14 -0.1500 15.625000 14
## 15 -0.1429 14.236408 15
## 16 -0.1400 326.666667 16
## 17 -0.1380 12.752571 17
## 18 -0.1377 17.246507 18
## 19 -0.1363 12.085971 19
## 20 -0.1362 8.636107 20
## 21 -0.1361 15.971914 21
## 22 -0.1309 29.123469 22
## 23 -0.1292 36.961064 23
## 24 -0.1289 28.182994 24
## 25 -0.1288 8.772267 25
## 26 -0.1079 7.329401 26
## 27 -0.0869 18.433375 27
## 28 -0.0823 13.037688 28
## 29 -0.0800 125.000000 29
## 30 -0.0791 24.112141 30
## 31 -0.0737 13.719241 31
## 32 -0.0734 54.717566 32
## 33 -0.0678 10.613702 33
## 34 -0.0601 99.014754 34
## 35 -0.0600 15.076923 35
## 36 -0.0548 33.922677 36
## 37 -0.0522 42.524326 37
## 38 -0.0502 3.132688 38
## 39 -0.0502 100.503782 39
## 40 -0.0487 59.028134 40
## 41 -0.0467 13.492586 41
## 42 -0.0452 41.558582 42
```

## 43	-0.0429	29.880715	43
## 44	-0.0408	31.204324	44
## 45	-0.0393	13.492586	45
## 46	-0.0389	37.190400	46
## 47	-0.0360	166.666667	47
## 48	-0.0305	16.536269	48
## 49	-0.0280	83.333333	49
## 50	-0.0268	13.719241	50
## 51	-0.0222	7.257628	51
## 52	-0.0197	13.719241	52
## 53	-0.0195	95.782629	53
## 54	-0.0186	94.915800	54
## 55	-0.0176	16.188098	55
## 56	-0.0169	8.960056	56
## 57	-0.0166	80.064077	57
## 58	-0.0160	125.000000	58
## 59	-0.0149	83.333333	59
## 60	-0.0143	13.493815	60
## 61	-0.0135	35.333263	61
## 62	-0.0125	72.168784	62
## 63	-0.0122	16.694514	63
## 64	-0.0118	90.535746	64
## 65	-0.0113	89.442719	65
## 66	-0.0112	21.853030	66
## 67	-0.0073	5.026776	67
## 68	-0.0040	27.451403	68
## 69	-0.0025	21.853030	69
## 70	-0.0015	77.849894	70
## 71	-0.0015	156.173762	71
## 72	-0.0014	18.698940	72
## 73	-0.0003	125.000000	73
## 74	0.0033	21.853030	74
## 75	0.0035	13.749033	75
## 76	0.0055	13.719241	76
## 77	0.0057	101.534617	77
## 78	0.0059	21.853030	78
## 79	0.0062	14.438269	79
## 80	0.0069	65.372045	80
## 81	0.0101	129.099445	81
## 82	0.0116	127.000127	82
## 83	0.0152	24.477769	83
## 84	0.0170	31.654447	84
## 85	0.0180	14.403780	85
## 86	0.0221	66.964953	86
## 87	0.0252	34.503278	87
## 88	0.0266	8.823339	88
## 89	0.0333	22.456323	89
## 90	0.0347	34.340141	90
## 91	0.0399	45.596075	91
## 92	0.0435	45.314325	92
## 93	0.0441	3.025958	93
## 94	0.0470	27.216553	94
## 95	0.0499	44.280744	95
## 96	0.0514	64.956980	96
## 97	0.0651	21.398025	97
## 98	0.0675	5.291968	98

```
## 99 0.0716 8.515960 99
## 100 0.0762 32.915401 100
## 101 0.0831 10.137797 101
## 102 0.0909 32.897585 102
## 103 0.0985 38.433122 103
## 104 0.1366 25.751310 104
## 105 0.1511 15.723193 105
## 106 0.1676 12.559009 106
## 107 0.2933 8.172324 107
```

```
dev.off()
```

```
## pdf
## 2
```

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
```

```
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 = 130; method: REML)
##
##      logLik    Deviance      AIC      BIC      AICc
## 60.5164 -121.0328 -111.0328 -96.8118 -110.5369
##
## Variance Components:
##
##      estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000   55    no      ID.article
## sigma^2.2  0.0129  0.1135  130    no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 127) = 4865.4360, p-val < .0001
##
## Test of Moderators (coefficients 1:3):
## QM(df = 3) = 48.1049, p-val < .0001
##
## Model Results:
##
##      estimate      se      zval      pval
## Variance.as.standard.error      -0.4183  0.2768  -1.5111  0.1308
## Self.reported.or.measuredMeasured      -0.0250  0.0173  -1.4476  0.1477
## Self.reported.or.measuredSelf-reported      -0.1574  0.0444  -3.5485  0.0004
##      ci.lb      ci.ub
## Variance.as.standard.error      -0.9609  0.1243
## Self.reported.or.measuredMeasured      -0.0588  0.0088
## Self.reported.or.measuredSelf-reported      -0.2444  -0.0705  ***
##
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

##Sensitivity analyses Repeat analysis excluding the studies that use the following less direct measured of productivity: Research time (N = 4), job-loss (N = 1), burnout (N = 1), and number of projects (N = 1). It looks like excluding these does not change the results much.

```
# Create dataframe without less direct measures of
# productivity
sensitive_data <- subset(all_data, all_data$Broad.specific.productivity !=
  "Job loss")
sensitive_data <- subset(sensitive_data, Broad.specific.productivity !=
  "Burn-out")
sensitive_data <- subset(sensitive_data, Broad.specific.productivity !=
  "Work hours")
sensitive_data <- subset(sensitive_data, Broad.specific.productivity !=
  "Projects")
sensitive_data <- subset(sensitive_data, Broad.specific.productivity !=
  "Job loss")
```

Repeat model 1 (all effect sizes)

```
sm1 <- rma.mv(yi, vi, random = ~1 | ID.article/ID.observation,
  data = sensitive_data)
sm1
```

```
##
## Multivariate Meta-Analysis Model (k = 127; method: REML)
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0003  0.0174    52    no      ID.article
## sigma^2.2  0.0129  0.1135   127    no  ID.article/ID.observation
##
## Test for Heterogeneity:
## Q(df = 126) = 4899.9629, p-val < .0001
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## -0.0586  0.0125  -4.7070  <.0001  -0.0830  -0.0342  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(sm1)
```

```
##
## Multivariate Meta-Analysis Model (k = 127; method: REML)
##
##      logLik  Deviance      AIC      BIC      AICc
##  65.3903 -130.7806 -124.7806 -116.2718 -124.5839
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0003  0.0174    52    no      ID.article
## sigma^2.2  0.0129  0.1135   127    no  ID.article/ID.observation
```

```
##
## Test for Heterogeneity:
## Q(df = 126) = 4899.9629, p-val < .0001
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## -0.0586  0.0125  -4.7070  <.0001  -0.0830  -0.0342  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Repeat model 2 (separated by study type)

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

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

## Publications  Submissions      Survey
##           48           64           15

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

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

##
## Multivariate Meta-Analysis Model (k = 127; method: REML)
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0000   52    no      ID.article
## sigma^2.2  0.0130  0.1139  127    no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 124) = 4861.8375, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 4.7140, p-val = 0.0947
##
## Model Results:
##
##                                     estimate      se      zval
## intrcpt                           -0.0513  0.0166  -3.0991
## Broad.productivity.measure.reorderedPublications  0.0055  0.0244   0.2238
## Broad.productivity.measure.reorderedSurvey        -0.0920  0.0452  -2.0361
##                                     pval      ci.lb      ci.ub
## intrcpt                           0.0019 -0.0838  -0.0189  **
## Broad.productivity.measure.reorderedPublications  0.8229 -0.0424   0.0533
## Broad.productivity.measure.reorderedSurvey        0.0417 -0.1806  -0.0034  *
##
## ---
```



```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(sm.area)
```

```
##
## Multivariate Meta-Analysis Model (k = 127; method: REML)
##
##      logLik      Deviance      AIC      BIC      AICc
## 65.6626 -131.3252 -121.3252 -107.2238 -120.8167
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1 0.0000 0.0000   52    no      ID.article
## sigma^2.2 0.0130 0.1139  127    no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 124) = 4861.8375, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 4.7140, p-val = 0.0947
##
## Model Results:
##
##                                     estimate      se      zval
## intrcpt                        -0.0513 0.0166 -3.0991
## Broad.productivity.measure.reorderedPublications 0.0055 0.0244 0.2238
## Broad.productivity.measure.reorderedSurvey      -0.0920 0.0452 -2.0361
##
##                                     pval      ci.lb      ci.ub
## intrcpt                        0.0019 -0.0838 -0.0189 **
## Broad.productivity.measure.reorderedPublications 0.8229 -0.0424 0.0533
## Broad.productivity.measure.reorderedSurvey      0.0417 -0.1806 -0.0034 *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

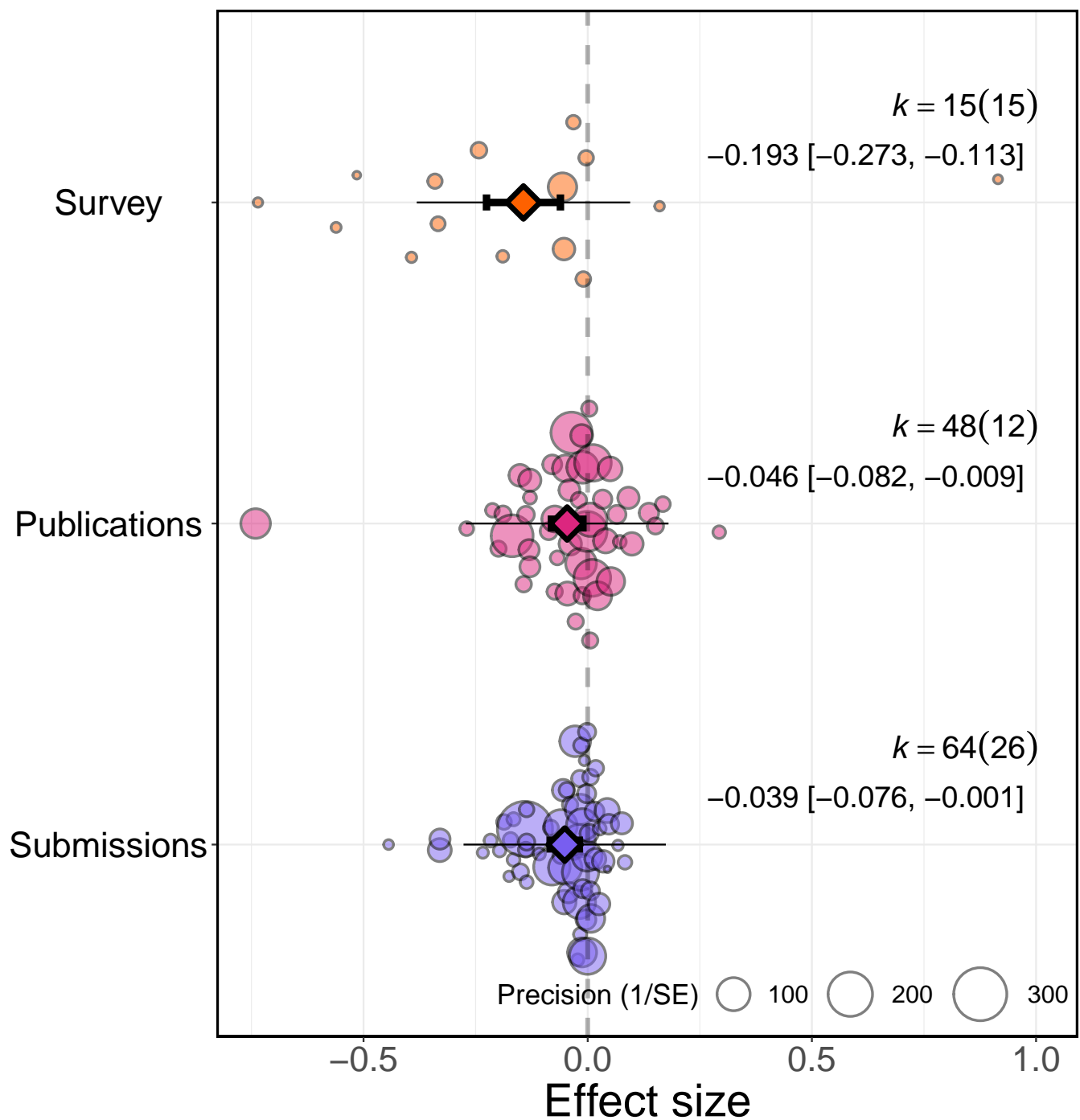
```
##
## Multivariate Meta-Analysis Model (k = 127; method: REML)
##
## Variance Components:
##
##      estim      sqrt  nlvls  fixed      factor
## sigma^2.1 0.0000 0.0000   52    no      ID.article
## sigma^2.2 0.0130 0.1139  127    no  ID.article/ID.observation
##
## Test for Residual Heterogeneity:
## QE(df = 124) = 4861.8375, p-val < .0001
##
## Test of Moderators (coefficients 1:3):
## QM(df = 3) = 27.7770, p-val < .0001
##
## Model Results:
##
##                                     estimate      se      zval
```

```
## Broad.productivity.measure.reorderedSubmissions      -0.0513  0.0166  -3.0991
## Broad.productivity.measure.reorderedPublications      -0.0459  0.0179  -2.5597
## Broad.productivity.measure.reorderedSurvey            -0.1433  0.0420  -3.4089
##
##               pval      ci.lb      ci.ub
## Broad.productivity.measure.reorderedSubmissions      0.0019  -0.0838  -0.0189  **
## Broad.productivity.measure.reorderedPublications      0.0105  -0.0810  -0.0107   *
## Broad.productivity.measure.reorderedSurvey            0.0007  -0.2257  -0.0609  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

sp1 <- my.orchard(sm.area1, mod = "Broad.productivity.measure.reordered",
  group = "Article.ID", data = sensitive_data, xlab = "Effect size",
  whisker = 0.05, transfm = "none") + annotate("text", size = 6,
  y = 0.62, x = 3.15, label = paste("-0.193 [-0.273, -0.113]")) +
  annotate("text", size = 6, y = 0.62, x = 2.15, label = paste("-0.046 [-0.082, -0.009]")) +
  annotate("text", size = 6, y = 0.62, x = 1.15, 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)'

sp1
```



```
ggsave("figureS2.png", width = 10, height = 10, dpi = 300)
```

Leave-one-out model

```
# For survey-studies only
msurvey <- rma.uni(yi, vi, random = ~1 | ID.article/ID.observation,
  slab = paste(Author), data = survey_data)
```

```
## Warning: Extra argument ('random') disregarded.
```

```
m
```

```
##
## Multivariate Meta-Analysis Model (k = 130; method: REML)
```

```
##
## Variance Components:
##
##          estim      sqrt  nlvls  fixed          factor
## sigma^2.1  0.0013  0.0366    55    no          ID.article
## sigma^2.2  0.0135  0.1163   130    no  ID.article/ID.observation
##
## Test for Heterogeneity:
## Q(df = 129) = 4939.2102, p-val < .0001
##
## Model Results:
##
## estimate      se      zval      pval      ci.lb      ci.ub
## -0.0706  0.0144  -4.9170  <.0001  -0.0988  -0.0425  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
llosurvey <- leave1out(msurvey)
llosurvey
```

```
##
##          estimate      se      zval      pval      ci.lb      ci.ub      Q
## Diaz et al.          -0.1828  0.0445  -4.1068  0.0000  -0.2700  -0.0955  119.9661
## Maguire et al.       -0.2060  0.0620  -3.3241  0.0009  -0.3275  -0.0846  141.9133
## Plaunova et al.      -0.2112  0.0644  -3.2793  0.0010  -0.3374  -0.0850  144.3802
## Yildirim & Elsen-Ziya -0.2164  0.0649  -3.3344  0.0009  -0.3437  -0.0892  147.3691
## Hoggarth et al.      -0.2112  0.0650  -3.2509  0.0012  -0.3386  -0.0839  142.8381
## Guintivano et al.    -0.2175  0.0663  -3.2802  0.0010  -0.3475  -0.0875  146.8956
## Davis et al.         -0.2182  0.0676  -3.2265  0.0013  -0.3508  -0.0857  141.6849
## Ghaffarizadeh et al. -0.2187  0.0677  -3.2319  0.0012  -0.3513  -0.0861  142.9060
## Staniscuaski et al.   -0.2184  0.0680  -3.2112  0.0013  -0.3518  -0.0851  103.4918
## Barber et al.        -0.2236  0.0684  -3.2704  0.0011  -0.3576  -0.0896  144.9627
## Gao et al.           -0.2253  0.0675  -3.3389  0.0008  -0.3575  -0.0930  148.7270
## Krukowski et al..1    -0.2258  0.0678  -3.3311  0.0009  -0.3587  -0.0930  148.6341
## Ellinas et al.        -0.2260  0.0676  -3.3452  0.0008  -0.3584  -0.0936  148.8064
## Shalaby et al.        -0.2262  0.0670  -3.3765  0.0007  -0.3575  -0.0949  148.9952
## Krukowski et al..2    -0.2274  0.0678  -3.3520  0.0008  -0.3603  -0.0944  148.8787
## Krukowski et al..3    -0.2275  0.0679  -3.3482  0.0008  -0.3606  -0.0943  148.8315
## Deryugena et al.      -0.2341  0.0675  -3.4674  0.0005  -0.3664  -0.1018  118.7065
## Myers et al.          -0.2342  0.0674  -3.4724  0.0005  -0.3663  -0.1020  144.8496
## Candido et al.        -0.2340  0.0669  -3.4989  0.0005  -0.3651  -0.1029  148.5361
## Rodriguez-Rivero et al. -0.2355  0.0666  -3.5361  0.0004  -0.3661  -0.1050  147.6368
## Breuning et al.       -0.2358  0.0665  -3.5460  0.0004  -0.3661  -0.1055  147.5120
## Stenson et al.        -0.2363  0.0644  -3.6721  0.0002  -0.3625  -0.1102  147.7250
## Camerlink et al.      -0.2446  0.0507  -4.8209  0.0000  -0.3440  -0.1451  132.8159
##
##          Qp      tau2      I2      H2
## Diaz et al.    0.0000  0.0271  86.8171  7.5856
## Maguire et al. 0.0000  0.0651  94.0702  16.8639
## Plaunova et al. 0.0000  0.0713  94.5478  18.3414
## Yildirim & Elsen-Ziya 0.0000  0.0737  94.7310  18.9788
## Hoggarth et al. 0.0000  0.0723  94.5957  18.5038
## Guintivano et al. 0.0000  0.0764  94.8848  19.5497
## Davis et al.   0.0000  0.0788  94.8970  19.5962
## Ghaffarizadeh et al. 0.0000  0.0789  94.9248  19.7037
## Staniscuaski et al. 0.0000  0.0795  93.8845  16.3519
## Barber et al.   0.0000  0.0808  94.9005  19.6096
```

```

## Gao et al.          0.0000 0.0792 95.0411 20.1657
## Krukowski et al..1  0.0000 0.0798 95.0497 20.2008
## Ellinas et al.      0.0000 0.0793 95.0436 20.1759
## Shalaby et al.      0.0000 0.0783 95.0063 20.0252
## Krukowski et al..2  0.0000 0.0798 95.0367 20.1479
## Krukowski et al..3  0.0000 0.0800 95.0298 20.1200
## Deryugena et al.    0.0000 0.0780 91.0729 11.2018
## Myers et al.        0.0000 0.0778 93.0365 14.3606
## Candido et al.      0.0000 0.0769 94.8211 19.3093
## Rodriguez-Rivero et al. 0.0000 0.0759 94.6868 18.8210
## Breuning et al.     0.0000 0.0756 94.6757 18.7819
## Stenson et al.      0.0000 0.0714 94.5591 18.3793
## Camerlink et al.    0.0000 0.0390 90.4887 10.5138

png(file = "figureS3.png", width = 3.25, height = 3.25, units = "in",
     res = 1200, pointsize = 4)
par(mar = c(5, 5, 2, 2), xaxs = "i", yaxs = "i", cex.axis = 2,
     cex.lab = 2)
forest(slab = l1osurvey$slab, l1osurvey$estimate, sei = l1osurvey$se,
       header = TRUE, xlab = "Leave One Out Estimate", refline = coef(l1osurvey))
dev.off()

## pdf
## 2

```