Ethnic Discrimination in Hiring Decisions: Replication Material

Didier Ruedin

15 December 2015

This is the replication material for 'Ethnic Discrimination in Hiring Decisions: A Meta-Analysis of Correspondence Tests 1990–2015'. All data are in the file `meta-clean.csv`, and analyses were carried out using *R* R version 3.2.3 (2015-12-10).

R1. Variables in the Data File

This table describes all the variables in the data file `meta-clean.csv` for the purpose of easing replication (see section R2).

Variable	Description
global	case refers to the 'study' level (coded 'global') or the 'subgroup' level (coded 'subgroup'). All subgroups are part of a study.
author	author(s) of the study
year	year of the study (publication)
region	where the study was carried out; only OECD countries are included (coded 'North America', 'Europe (EU)', 'Europe (non-EU)', or 'Other')
directive	study was carried out before or after the EU directives were implemented (coded 'before', 'after', or 'not applicable')
country	country where study was carried out; Akintola (2011) is treated as two studies
testing.year	year of testing, or year when testing began where it covered multiple years
testing.year.implicit	year of testing (as above); if year not given, set to one year before publication, as 'best guess'
firstgen	information on immigrant groups tested
firstgen.explicit	generation of immigrant group tested, only explicit mention (coded 'first', 'second', 'mixed', or 'native'); in the article 'first' is referred to as *foreign-born*, wherease 'second' is referred to as *native*; the 'native' in the dataset refers to native minorities in Sweden that cannot be traced to immigration (Carlsson 2010)
firstgen.implicit	generation of immigrant group tested (as above), but considers both explicit and implicit mentions; local education was coded as 'second' generation
minority	specific group tested
gender	gender tested (coded 'female', 'male', 'both', or 'not indicated')
source	source of adverts, such as newspapers or online adverts
firm	firm characteristics
other	additional information

Variable	Description
qualification	qualification tested
education	level of education (skills-level) tested
skills	skills-level tested (coded as 'high', 'low', or 'mixed')
isced	ISCED tested
job	additional information on job tested
isco88	ISCO-88 of job tested
number.of.jobs	number of jobs applications were sent to
n.none.invited	of which neither of the applicant was invited
n.both.invited	of which both applicants were invited
n.one.or.more.success	of which minority, majority, or both candidated were invited
n.only.majority.invited	of which only majority candidate was invited
success.majority	success rate for majority candidates: $\frac{n.both.invited+n.only.majority.invited}{n.one.or.more.success}$
n.only.minority.invited	of which only minority candidate was invited
success.minority	success rate for minority candidates: $\frac{n.both.invited+n.only.minority.invited}{n.one.or.more.success}$
relative.call.back.rate	relative call back rate: $\frac{success.majority}{success.minority}$
net.discrimination.rate	net discrimination rate: $success.majority - success.minority = \frac{n.only.majority.invited-n.only.minorit.invited}{n.one.or.more.success}$
n.valid.applications	number of valid applications
n.cases	number of cases
negative.treatment.minority	number of applications by minority without response (failure)
positive.treatment.minority	number of application by minority with response (success)
negative.treatment.majority	number of applications by majority without response (failure)
positive.treatment.majority	number of applications by majority with response (success)

R2. Replication Material and Additional Analysis

Set-Up

Loading data and packages

```
  knitr::opts_chunk$set(cache=TRUE, fig.width=6.5, fig.height=6.5)
  library(metafor)  # meta analysis
```

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

```
## Loading 'metafor' package (version 1.9-8). For an overview
## and introduction to the package please type: help(metafor).
```

```
  library(knitr)    # tables in knitr
  options(digits=3) # fewer digits shown
```

```r
setwd("/home/didier/ownCloud/Attitudes/NCCR Discrimination")
# setwd("C:/Users/ruedind/switchdrive/Attitudes/Discrimination NCCR")
disc <- read.csv("meta-clean.csv", header=TRUE, sep=",", fileEncoding="UTF8")
```

How Many Studies?

```r
# how many studies?
table(disc$global)
```

```
##
##   global subgroup
##       43      738
```

```r
# by year:
table(disc$global, disc$year)
```

```
##
##            1991 1993 1994 1995 1997 1998 2003 2004 2006 2007 2009 2010
##   global      2    1    1    2    1    1    2    1    1    4    1    3
##   subgroup   35    0   19   24    0    4   57    9   18   78   36   60
##
##            2011 2012 2013 2014 2015
##   global      5    8    2    5    3
##   subgroup   36  202   14   71   75
```

```r
# by large region
table(disc$global, disc$region)
```

```
##
##            Europe EU Europe non-EU North America other
##   global          26             2            12     3
##   subgroup       374            67           229    68
```

```r
# creating year groups to summarize data
disc$yearGroup <- NA
disc$yearGroup[disc$year >= 1990 & disc$year <= 1995] <- 1990
disc$yearGroup[disc$year >= 1996 & disc$year <= 2000] <- 1996
disc$yearGroup[disc$year >= 2001 & disc$year <= 2005] <- 2001
disc$yearGroup[disc$year >= 2006 & disc$year <= 2010] <- 2006
disc$yearGroup[disc$year >= 2011 & disc$year <= 2015] <- 2011
# year group of testing year to summarize data:
disc$yearGroupTesting <- NA
disc$yearGroupTesting[disc$testing.year.implicit >= 1990 & disc$testing.year.implicit <= 1995] <- 1990
disc$yearGroupTesting[disc$testing.year.implicit >= 1996 & disc$testing.year.implicit <= 2000] <- 1996
disc$yearGroupTesting[disc$testing.year.implicit >= 2001 & disc$testing.year.implicit <= 2005] <- 2001
disc$yearGroupTesting[disc$testing.year.implicit >= 2006 & disc$testing.year.implicit <= 2010] <- 2006
disc$yearGroupTesting[disc$testing.year.implicit >= 2011 & disc$testing.year.implicit <= 2015] <- 2011
table(disc$yearGroup, disc$yearGroupTesting) ## makes absolutely no difference
```

```
##
##        1990 1996 2001 2006 2011
##   1990   73    0    0    0    0
##   1996    0    6    0    0    0
##   2001    0    7   62    0    0
##   2006    0    0   32  169    0
##   2011    0    0    0  103  318
```

```
  # mean call-back rate, in table 1:
  aggregate(disc$relative.call.back.rate, by=list(Time=disc$yearGroup, Global=disc$global), mean, na.rm=
```

```
##    Time   Global    x
## 1  1990   global 1.29
## 2  1996   global 1.50
## 3  2001   global 1.47
## 4  2006   global 1.98
## 5  2011   global 1.47
## 6  1990 subgroup 1.28
## 7  1996 subgroup 2.04
## 8  2001 subgroup 1.54
## 9  2006 subgroup 2.28
## 10 2011 subgroup 1.60
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global), mean, na.rm=TRUE) # overall
```

```
##     Global    x
## 1   global 1.55
## 2 subgroup 1.75
```

```
  table(disc$yearGroup, disc$global) # N
```

```
##
##        global subgroup
##   1990      6       78
##   1996      2        4
##   2001      3       66
##   2006      9      192
##   2011     23      398
```

```
  table(disc$yearGroup[disc$global == "global"], disc$region[disc$global == "global"]) # N
```

```
##
##        Europe EU Europe non-EU North America other
##   1990         4             0             1     1
##   1996         2             0             0     0
##   2001         0             1             2     0
##   2006         9             0             0     0
##   2011        11             1             9     2
```

There are 43 studies in the meta analysis, covering a total of 738 groups ('subgroups').

General Summary

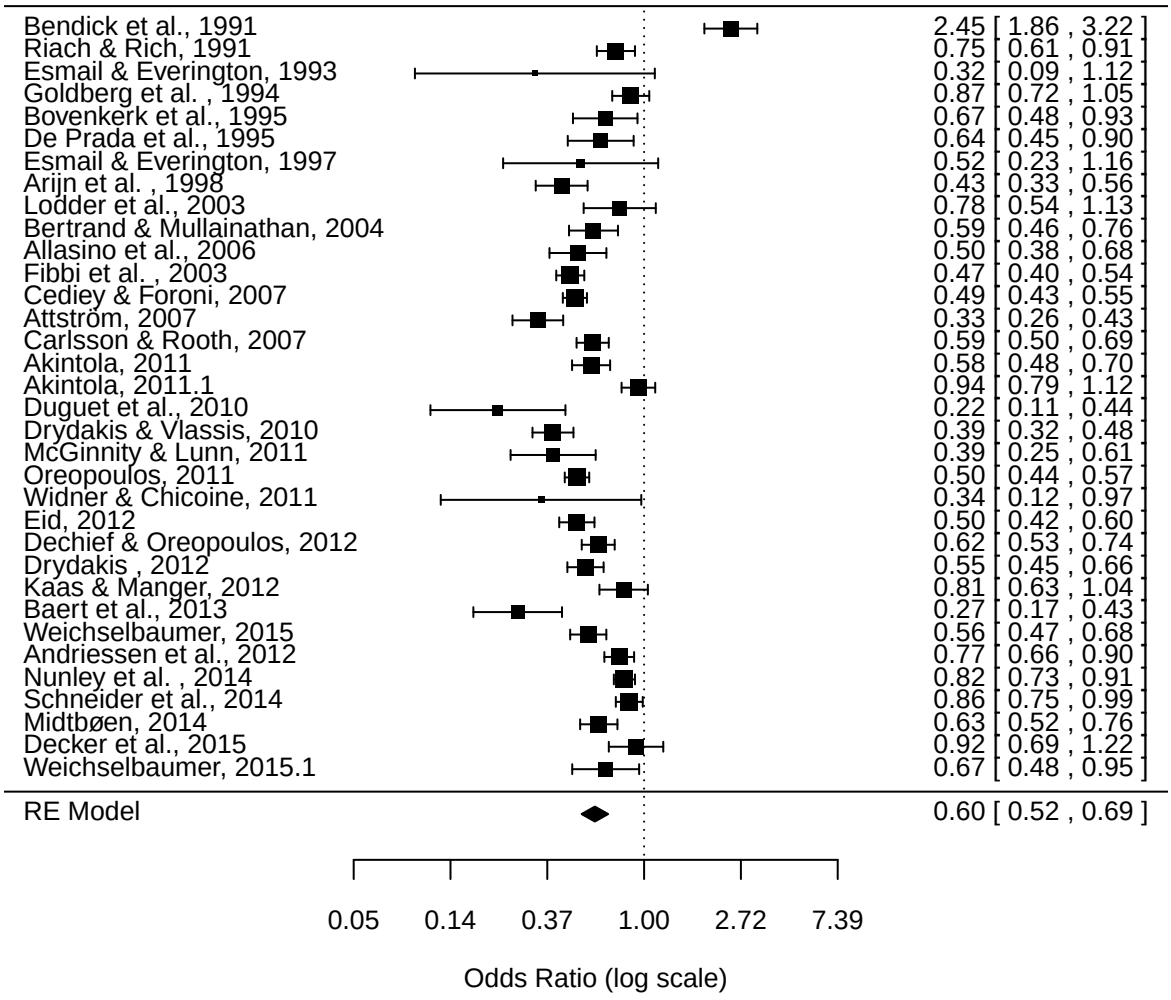
As a base line, *all* correspondence tests pooled.

Study Level

```
# only global studies
resGlobal <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$global=="global" & !is.na(disc$negative.treatment.majority)
  & !is.na(disc$negative.treatment.minority) ), measure="OR",
  slab=paste(author, year, sep=", "), method="REML")
summary(resGlobal)                    # SUPPLEMENT S6, TEXT
```

```
##
## Random-Effects Model (k = 34; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -18.5391   37.0782   41.0782   44.0712   41.4782
##
## tau^2 (estimated amount of total heterogeneity): 0.1406 (SE = 0.0403)
## tau (square root of estimated tau^2 value):      0.3750
## I^2 (total heterogeneity / total variability):   92.73%
## H^2 (total variability / sampling variability):  13.75
##
## Test for Heterogeneity:
## Q(df = 33) = 327.6821, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##  -0.5086   0.0701  -7.2543   <.0001  -0.6460  -0.3712      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# tiff(file="Fig1.tif", width = 6.4, height = 6.4, units = "in", pointsize = 9, compression = "lzw",
   forest(resGlobal, atransf=exp, cex=0.8)
```




```
# dev.off()
```

Subgroups

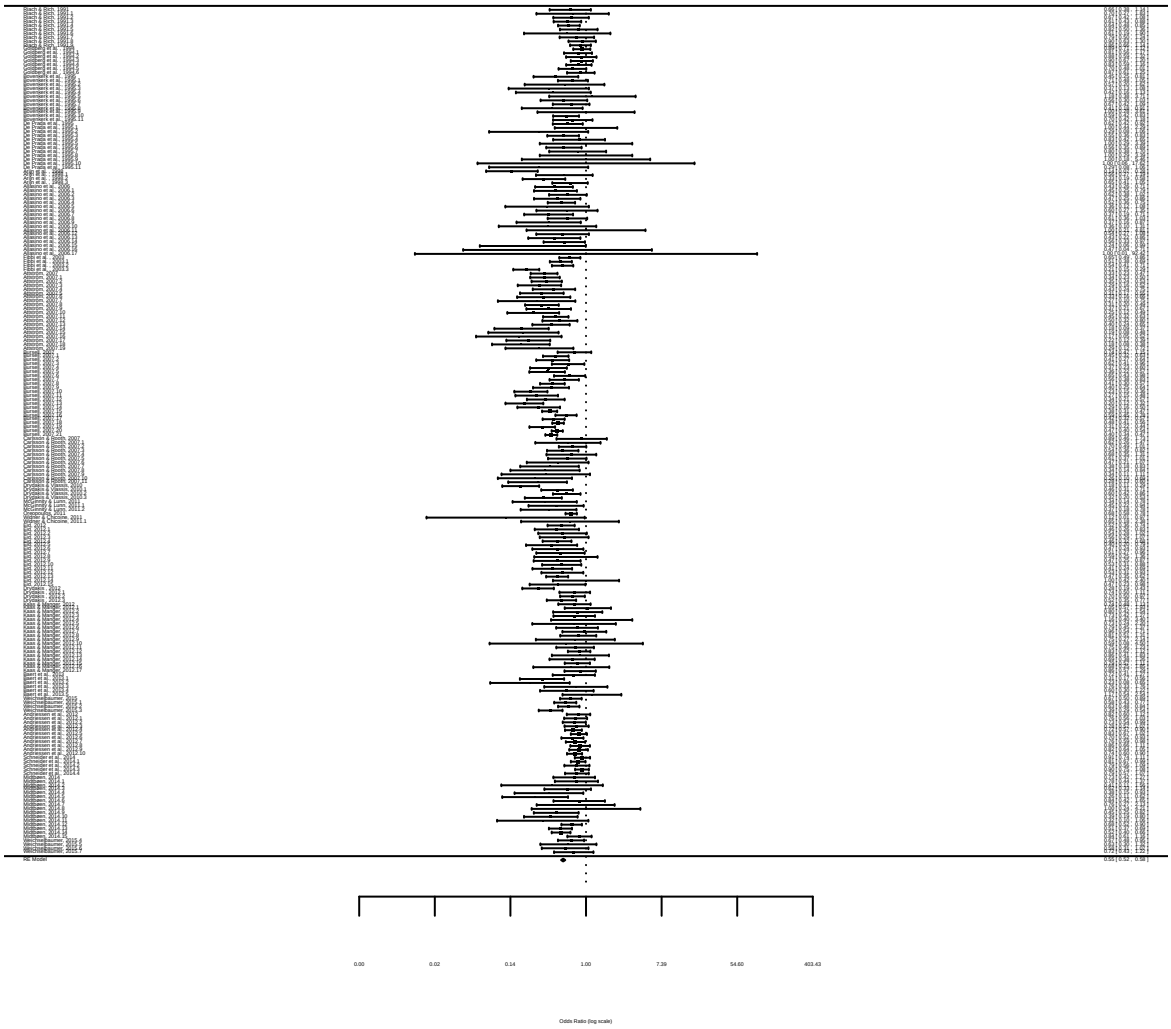
The figure will not be lisible.

```
### fit random-effects model (use slab argument to define study labels)
resSub <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$global=="subgroup" & !is.na(disc$negative.treatment.majority)
  & !is.na(disc$negative.treatment.minority) ),  measure="OR",
  slab=paste(author, year, sep=", "), method="REML")
summary(resSub)                          # SUPPLEMENT S6, TEXT
```

```
##
## Random-Effects Model (k = 215; tau^2 estimator: REML)
##
##    logLik   deviance        AIC        BIC       AICc
## -125.4099   250.8198   254.8198   261.5518   254.8767
```

```
##
## tau^2 (estimated amount of total heterogeneity): 0.0934 (SE = 0.0154)
## tau (square root of estimated tau^2 value):      0.3056
## I^2 (total heterogeneity / total variability):   68.23%
## H^2 (total variability / sampling variability):  3.15
##
## Test for Heterogeneity:
## Q(df = 214) = 647.3077, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##  -0.6041   0.0283 -21.3758   <.0001  -0.6595  -0.5487      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
forest(resSub, atransf=exp)
```



Summary Table ————-

```
  tab <- data.frame(exp(coef(resGlobal)), exp(coef(resSub))) # combining
  colnames(tab) <- c("Global", "Subgroups")                  # adding labels
  kable(tab)                                                 # SUPPLEMENT S6
```

	Global	Subgroups
intrcpt	0.601	0.547

Call-Back Rates

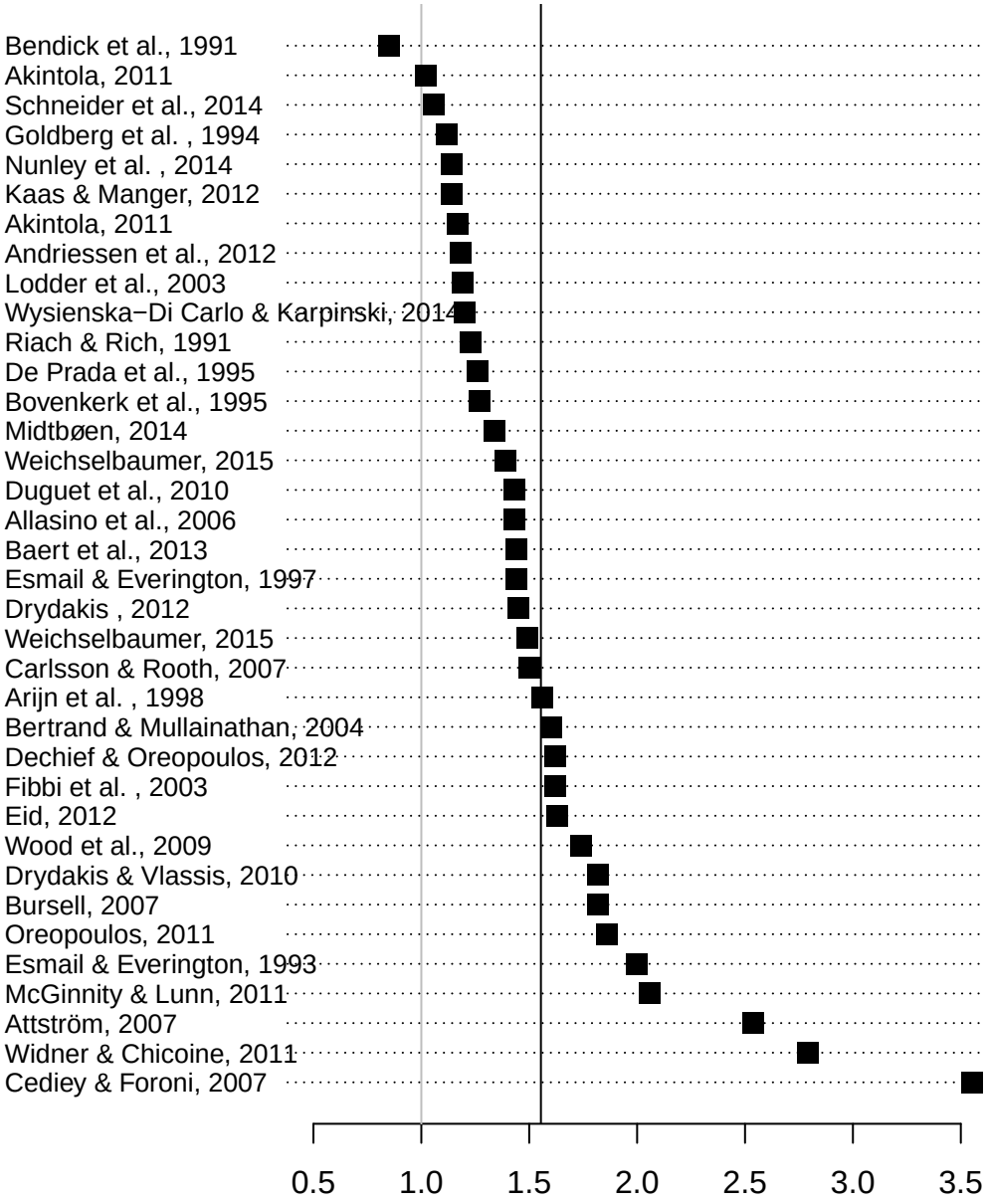
```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global), mean, na.rm=TRUE)   # SUPPLEMENT
```

```
##     Global    x
## 1   global 1.55
## 2 subgroup 1.75
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global), median, na.rm=TRUE) # SUPPLEMENT
```

```
##     Global    x
## 1   global 1.44
## 2 subgroup 1.49
```

```
  disc$study <- paste(disc$author, disc$year, sep=", ")     # study = Author, Year
  x <- subset(disc, disc$global=="global" & !is.na(disc$relative.call.back.rate) )[c("study","relative.
  xo <- x[order(x[,2], decreasing=TRUE),]
# tiff(file="Fig2.tif", width = 6.4, height = 4.8, units = "in", pointsize = 9, compression = "lzw", bg
old.par <- par(no.readonly=TRUE) # save defaults
  par(mar=c(2.1,12.1,1.1,2.1), xpd=FALSE)
  len <- length(xo[,2])
  plot(1, xlim=c(0.5, 3.5), ylim=c(1, len), type="n", bty="n", axes=FALSE, ylab="", xlab="")
  axis(1, cex=.8)
  abline(v=1, lty=1, col="grey")
  mval <- mean(disc$relative.call.back.rate[disc$global=="global"], na.rm=TRUE)
  abline(v=mval, lty=1, col="black")
  for(i in 1:len) {
    abline(h=i, lty=3)
    points(xo[i,2],i, pch=15, cex=1.5)
  }
  par(xpd=TRUE) # draw in margins
  for(i in 1:len) {
    text(-0.93,i, xo[i,1], cex=0.8, adj=c(0, 0.5)) # align left; optimized for TIFF
  }
```




```
par(old.par) # restore default par
# dev.off()
```

First and Second Generation

Call-Back Ratios

```
  # explicit:
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, First=disc$firstgen.explicit), mea
```

```
##     Global  First    x
```

```
## 1   global        1.50
## 2 subgroup        1.71
## 3   global  first 1.34
## 4 subgroup  first 1.79
## 5   global  mixed 1.48
## 6 subgroup  mixed 1.42
## 7 subgroup native  NaN
## 8   global second 1.79
## 9 subgroup second 1.82
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, First=disc$firstgen.explicit), mea
```

```
##     Global  First    x
## 1   global        1.44
## 2 subgroup        1.45
## 3   global  first 1.34
## 4 subgroup  first 1.54
## 5   global  mixed 1.39
## 6 subgroup  mixed 1.42
## 7 subgroup native   NA
## 8   global second 1.47
## 9 subgroup second 1.62
```

```
  # implicit:
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, First=disc$firstgen.implicit), mea
```

```
##     Global  First    x
## 1   global        1.53
## 2 subgroup        1.73
## 3   global  first 1.30
## 4 subgroup  first 1.93
## 5   global  mixed 1.48
## 6 subgroup  mixed 1.42
## 7 subgroup native  NaN
## 8   global second 1.61
## 9 subgroup second 1.71
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, First=disc$firstgen.implicit), mea
```

```
##     Global  First    x
## 1   global        1.56
## 2 subgroup        1.43
## 3   global  first 1.26
## 4 subgroup  first 1.64
## 5   global  mixed 1.39
## 6 subgroup  mixed 1.42
## 7 subgroup native   NA
## 8   global second 1.45
## 9 subgroup second 1.46
```

```
  # how many studies?
    table(disc$global, disc$firstgen.explicit) # TABLE 2 (184)
```

```
##
##                first mixed native second
##   global    29     2     4      0      8
##   subgroup 490    44     8     12    184
```

```
    table(disc$global, disc$firstgen.implicit) # TABLE 2 (97, 448)
```

```
##
##                first mixed native second
##   global     9     3     5      0     26
##   subgroup 173    97     8     12    448
```

```
  # significant differences?
  t.test(relative.call.back.rate ~ firstgen.implicit, data=subset(disc, firstgen.implicit %in% c("first
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by firstgen.implicit
## t = 2, df = 200, p-value = 0.06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.00984  0.45714
## sample estimates:
##  mean in group first mean in group second
##                 1.93                 1.71
```

```
  t.test(relative.call.back.rate ~ firstgen.explicit, data=subset(disc, firstgen.explicit %in% c("first
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by firstgen.explicit
## t = -0.1, df = 60, p-value = 0.9
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.351  0.309
## sample estimates:
##  mean in group first mean in group second
##                 1.79                 1.82
```

Odds-Ratios

```
  # only subgroups, explicit:
  resGen1 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
    ci=positive.treatment.majority, di=negative.treatment.majority,
```

```
  data=subset(disc, disc$firstgen.explicit=="first" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority)),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
# only subgroups, implicit:
resGen3 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$firstgen.implicit=="first" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority)),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
```

```
# only subgroups, explicit:
resGen2 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$firstgen.explicit=="second" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority) ),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
# only subgroups, implicit:
resGen4 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$firstgen.implicit=="second" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority) ),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
```

```
# extracting coefficients and combining:
tab <- data.frame(exp(coef(resGen1)), exp(coef(resGen2)), exp(coef(resGen3)),
                  exp(coef(resGen4)))
# adding labels:
colnames(tab) <- c("1st generation, explicit", "2nd generation, explicit",
                   "1st generation, implicit", "2nd generation, implicit")
kable(t(tab))
```

	intrcpt
1st generation, explicit	0.587
2nd generation, explicit	0.551
1st generation, implicit	0.587
2nd generation, implicit	0.521

Specific Groups: Turks, Arabs, Indians/Pakistani, Chinese

Call-Back Ratios

```
disc$minor <- NA
# groups of interest:
disc$minor[disc$minority %in% c("Turkish", "Turkish ", "Turks")] <- "Turk"
disc$minor[disc$minority %in% c("arab", "arab ", "Arab", "native swedes of middle eastern origin", "fo
disc$minor[disc$minority %in% c("India", "Indian", "Pakistan", "Pakistani",
  "Pakistani/Bangladeshi")] <- "Indian/Pakistani"
disc$minor[disc$minority %in% c("Chinese", "Cinese",
```

```
    "Chinese with English first name")] <- "Chinese"
  # table(disc$minor)
  table(disc$minor, disc$global) # N, TABLE 3
```

```
##
##                    global subgroup
##   Arab                  3       68
##   Chinese               0       39
##   Indian/Pakistani      1       42
##   Turk                  5       73
```

```
  # TABLE 3:
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Group=disc$minor), mean, na.rm=TRU
```

```
##     Global            Group    x
## 1   global             Arab 2.18
## 2 subgroup             Arab 2.35
## 3 subgroup          Chinese 1.64
## 4   global Indian/Pakistani 1.34
## 5 subgroup Indian/Pakistani 1.77
## 6   global             Turk 1.23
## 7 subgroup             Turk 1.33
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Group=disc$minor), median, na.rm=
```

```
##     Global            Group    x
## 1   global             Arab 2.18
## 2 subgroup             Arab 2.13
## 3 subgroup          Chinese 1.57
## 4   global Indian/Pakistani 1.34
## 5 subgroup Indian/Pakistani 1.69
## 6   global             Turk 1.14
## 7 subgroup             Turk 1.23
```

```
  # significant differences?
  summary(aov(relative.call.back.rate ~ minor, data=subset(disc, global == "subgroup")))
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## minor         3   35.2   11.72    22.6  1e-12 ***
## Residuals   208  107.7    0.52
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 526 observations deleted due to missingness
```

Odds-Ratios

```
  # only subgroups
  resM1 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
    ci=positive.treatment.majority, di=negative.treatment.majority,
```

```
  data=subset(disc, disc$minor=="Arab" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority)),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
resM2 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$minor=="Chinese" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority)),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
resM3 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$minor=="Indian/Pakistani" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority)),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
resM4 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$minor=="Turk" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority)),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
# extracting and combining coefficients:
tab <- data.frame(exp(coef(resM1)), exp(coef(resM2)), exp(coef(resM3)),
                  exp(coef(resM4)))
# adding labels:
colnames(tab) <- c("Arab", "Chinese", "Indian/Pakistani", "Turk")
kable(t(tab)) # SUPPLEMENT S8
```

	intrcpt
Arab	0.383
Chinese	0.633
Indian/Pakistani	0.593
Turk	0.779

Specific Groups, Foreign-Born and Native Minorities

```
# explicit generations
  aggregate(disc$relative.call.back.rate, by=list(Generation=disc$firstgen.explicit, Global=disc$global
```

```
##    Generation   Global            Group    x
## 1                global             Arab 1.82
## 2      second   global             Arab 2.54
## 3             subgroup             Arab 2.07
## 4      second subgroup             Arab 3.00
## 5             subgroup          Chinese 1.70
## 6      second subgroup          Chinese 1.41
## 7      second   global Indian/Pakistani 1.34
## 8             subgroup Indian/Pakistani 1.90
## 9      second subgroup Indian/Pakistani 1.60
## 10               global             Turk 1.28
## 11      mixed   global             Turk 1.39
## 12     second   global             Turk 1.10
## 13            subgroup             Turk 1.44
```

```
## 14      mixed subgroup             Turk 1.42
## 15     second subgroup             Turk 1.22
```

```
  aggregate(disc$relative.call.back.rate, by=list(Generation=disc$firstgen.explicit, Global=disc$global
```

```
##    Generation   Global            Group    x
## 1               global             Arab 1.82
## 2      second   global             Arab 2.54
## 3             subgroup             Arab 1.87
## 4      second subgroup             Arab 2.71
## 5             subgroup          Chinese 1.62
## 6      second subgroup          Chinese 1.37
## 7      second   global Indian/Pakistani 1.34
## 8             subgroup Indian/Pakistani 1.85
## 9      second subgroup Indian/Pakistani 1.52
## 10              global             Turk 1.28
## 11      mixed   global             Turk 1.39
## 12     second   global             Turk 1.10
## 13            subgroup             Turk 1.35
## 14      mixed subgroup             Turk 1.42
## 15     second subgroup             Turk 1.18
```

```
# implicit generations
  aggregate(disc$relative.call.back.rate, by=list(Generation=disc$firstgen.implicit, Global=disc$global
```

```
##    Generation   Global            Group    x
## 1               global             Arab  NaN
## 2      second   global             Arab 2.18
## 3             subgroup             Arab 2.12
## 4      second subgroup             Arab 2.48
## 5       first subgroup          Chinese 1.88
## 6      second subgroup          Chinese 1.49
## 7      second   global Indian/Pakistani 1.34
## 8       first subgroup Indian/Pakistani 2.05
## 9      second subgroup Indian/Pakistani 1.56
## 10      mixed   global             Turk 1.39
## 11     second   global             Turk 1.19
## 12      mixed subgroup             Turk 1.42
## 13     second subgroup             Turk 1.32
```

```
  aggregate(disc$relative.call.back.rate, by=list(Generation=disc$firstgen.implicit, Global=disc$global
```

```
##    Generation   Global            Group    x
## 1               global             Arab   NA
## 2      second   global             Arab 2.18
## 3             subgroup             Arab 1.83
## 4      second subgroup             Arab 2.27
## 5       first subgroup          Chinese 1.68
## 6      second subgroup          Chinese 1.47
## 7      second   global Indian/Pakistani 1.34
## 8       first subgroup Indian/Pakistani 2.09
## 9      second subgroup Indian/Pakistani 1.46
```

```
## 10      mixed   global             Turk 1.39
## 11     second   global             Turk 1.13
## 12      mixed subgroup             Turk 1.42
## 13     second subgroup             Turk 1.20
```

Europe versus US/Canada

Call-Back Ratios

Creating additional variables

```
  # continent:
  disc$continent <- NA
  disc$continent[disc$country %in% c("US", "Canada")] <- "US/Canada"
  disc$continent[disc$country %in% c("Austria", "Belgium", "France", "Germany",
          "Greece", "Ireland", "Italy", "Netherlands", "Norway", "Poland", "Spain",
          "Sweden", "Switzerland", "UK")] <- "Europe"
  # European Union:
  disc$EU <- "non EU"
  disc$EU[disc$country %in% c("Austria", "Belgium", "France", "Germany",
          "Greece", "Ireland", "Italy", "Netherlands", "Poland", "Spain", "Sweden",
          "UK")] <- "European Union"
  # German-speaking:
  disc$German <- "other"
  disc$German[disc$country %in% c("Austria", "Germany", "Switzerland")] <- "German"

  table(disc$continent[!is.na(disc$relative.call.back.rate)], disc$global[!is.na(disc$relative.call.bac
```

```
##
##              global subgroup
##   Europe         25      340
##   US/Canada       9      126
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Cont=disc$continent), mean, na.rm=
```

```
##     Global      Cont    x
## 1   global    Europe 1.58
## 2 subgroup    Europe 1.84
## 3   global US/Canada 1.52
## 4 subgroup US/Canada 1.69
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Cont=disc$continent), median, na.
```

```
##     Global      Cont    x
## 1   global    Europe 1.44
## 2 subgroup    Europe 1.52
## 3   global US/Canada 1.60
## 4 subgroup US/Canada 1.49
```

Odds-Ratios

```
resM1 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$continent=="US/Canada"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority) ),
  measure="OR", slab=paste(author, year, sep=", "), method="REML")
resM2 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$continent=="Europe" & !is.na(disc$negative.treatment.majority)
  & !is.na(disc$negative.treatment.minority) ), measure="OR",
  slab=paste(author, year, sep=", "), method="REML")
tab <- data.frame(exp(coef(resM1)), exp(coef(resM2))) # combine
colnames(tab) <- c("US/Canada", "Europe")             # labels
kable(t(tab))
```

	intrcpt
US/Canada	0.610
Europe	0.542

```
# only subgroups
resM1s <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$continent=="US/Canada" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority) ),
  measure="OR", slab=paste(author, year, sep=", "), method="EB") # REML does not converge
resM2s <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$continent=="Europe" & disc$global=="subgroup"
  & !is.na(disc$negative.treatment.majority)
  & !is.na(disc$negative.treatment.minority) ), measure="OR",
  slab=paste(author, year, sep=", "), method="EB")
tab <- data.frame(exp(coef(resM1s)), exp(coef(resM2s))) # combine
colnames(tab) <- c("US/Canada", "Europe")               # labels
kable(t(tab)) # SUPPLEMENT S9
```

	intrcpt
US/Canada	0.563
Europe	0.542

German-Speaking Countries

```
# TABLE 5
aggregate(disc$relative.call.back.rate, by=list(German=disc$German, Global=disc$global), mean, na.rm="
```

```
##   German   Global    x
```

```
## 1 German   global 1.30
## 2  other   global 1.60
## 3 German subgroup 1.43
## 4  other subgroup 1.83
```

```
  aggregate(disc$relative.call.back.rate, by=list(German=disc$German, Global=disc$global), median, na.r
```

```
##   German   Global    x
## 1 German   global 1.26
## 2  other   global 1.44
## 3 German subgroup 1.31
## 4  other subgroup 1.55
```

```
  table(disc$German[!is.na(disc$relative.call.back.rate)], disc$global[!is.na(disc$relative.call.back.ra
```

```
##
##          global subgroup
##   German      6      119
##   other      30      421
```

```
  t.test(relative.call.back.rate ~ German, data=subset(disc, global == "subgroup"))
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by German
## t = -5, df = 500, p-value = 6e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.543 -0.257
## sample estimates:
## mean in group German  mean in group other
##                 1.43                 1.83
```

```
    # only subgroups
  resG1 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
    ci=positive.treatment.majority, di=negative.treatment.majority,
    data=subset(disc, disc$German=="German" & disc$global=="subgroup"
    & !is.na(disc$negative.treatment.majority) & !is.na(disc$negative.treatment.minority) ),
    measure="OR", slab=paste(author, year, sep=", "), method="REML")
  resG2 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
    ci=positive.treatment.majority, di=negative.treatment.majority,
    data=subset(disc, disc$German=="other" & disc$global=="subgroup"
    & !is.na(disc$negative.treatment.majority)
    & !is.na(disc$negative.treatment.minority) ), measure="OR",
    slab=paste(author, year, sep=", "), method="REML")
  tab <- data.frame(exp(coef(resG1)), exp(coef(resG2))) # combine
  colnames(tab) <- c("German", "other")                 # labels
  kable(t(tab)) # SUPPLEMENT S10
```

	intrcpt
German	0.716
other	0.505

Pre/post 2000 Directives

Refer to section S3 for the coding of implementation dates.

Ratios

```
  table(disc$directive, disc$global)  # N, TABLE 4
```

```
##
##                  global subgroup
##   after              20      327
##   before              6       47
##   not applicable     17      364
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global,
            Directive=disc$directive), mean, na.rm=TRUE)   # TABLE 4
```

```
##     Global      Directive    x
## 1   global          after 1.63
## 2 subgroup          after 1.96
## 3   global         before 1.44
## 4 subgroup         before 1.36
## 5   global not applicable 1.49
## 6 subgroup not applicable 1.59
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global,
            Directive=disc$directive), median, na.rm=TRUE) # TABLE 4
```

```
##     Global      Directive    x
## 1   global          after 1.44
## 2 subgroup          after 1.65
## 3   global         before 1.35
## 4 subgroup         before 1.21
## 5   global not applicable 1.47
## 6 subgroup not applicable 1.44
```

```
  t.test(relative.call.back.rate ~ directive, data=subset(disc, global == "subgroup" & directive %in% c
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by directive
```

```
## t = 5, df = 100, p-value = 8e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.345 0.853
## sample estimates:
##  mean in group after mean in group before
##                 1.96                 1.36
```

Economic Situation

There are two variable to match economic variables: the time of testing as indicated in the study, and one that *additionally* assumes that the time of testing was one year before publication for the studies that do not indicate the time of testing.

Importing and merging economic data:

```
# get data from database: World Development Indicators
gdp <- read.csv("worldbankGDP.csv", header=TRUE, sep=",")
unemp <- read.csv("worldbankUnemployment.csv", header=TRUE, sep=",")
# Unemployment, total (% of total labor force) (modeled ILO estimate)
# setting country codes for matching
disc$CountryCode[disc$country == "Australia"] <- "AUS"
disc$CountryCode[disc$country == "Austria"] <- "AUT"
disc$CountryCode[disc$country == "Belgium"] <- "BEL"
disc$CountryCode[disc$country == "Canada"] <- "CAN"
disc$CountryCode[disc$country == "France"] <- "FRA"
disc$CountryCode[disc$country == "Germany"] <- "DEU"
disc$CountryCode[disc$country == "Greece"] <- "GRC"
disc$CountryCode[disc$country == "Ireland"] <- "IRL"
disc$CountryCode[disc$country == "Italy"] <- "ITA"
disc$CountryCode[disc$country == "Mexico"] <- "MEX"
disc$CountryCode[disc$country == "Netherlands"] <- "NLD"
disc$CountryCode[disc$country == "Norway"] <- "NOR"
disc$CountryCode[disc$country == "Poland"] <- "POL"
disc$CountryCode[disc$country == "Spain"] <- "ESP"
disc$CountryCode[disc$country == "Switzerland"] <- "CHE"
disc$CountryCode[disc$country == "Sweden"] <- "SWE"
disc$CountryCode[disc$country == "UK"] <- "GBR"
disc$CountryCode[disc$country == "US"] <- "USA"
# combining country and year in "disc"
disc$CountryYear <- paste(disc$CountryCode, disc$testing.year, sep="")
disc$CountryYear[is.na(disc$testing.year)] <- NA # year unknown
# ditto for implicit years
disc$CountryYear.implicit <- paste(disc$CountryCode, disc$testing.year.implicit, sep="")
# add "gdp" and "unemp"
disc$gdp <- NA        # empty variables
disc$unemp <- NA
disc$gdp.imp <- NA
disc$unemp.imp <- NA
# for all explicit year:
for(cy in na.omit(unique(subset(disc$CountryYear, disc$testing.year > 1990)))){
  disc$gdp[disc$CountryYear==cy] <- as.numeric(subset(gdp,
          gdp$CountryCode==substr(cy,1,3) & gdp$SeriesName=="GDPgrowth")
```

```
            [paste("YR",substr(cy,4,7), sep="")])
    disc$unemp[disc$CountryYear==cy] <- as.numeric(subset(unemp,
            unemp$CountryCode==substr(cy,1,3))
            [paste("YR",substr(cy,4,7), sep="")])
  }
  # for all (implicit) years
  for(cy in unique(subset(disc$CountryYear.implicit, disc$testing.year > 1990))){
    disc$gdp.imp[disc$CountryYear.implicit==cy] <- as.numeric(subset(gdp,
            gdp$CountryCode==substr(cy,1,3) & gdp$SeriesName=="GDPgrowth")
            [paste("YR",substr(cy,4,7), sep="")])
    disc$unemp.imp[disc$CountryYear.implicit==cy] <- as.numeric(subset(unemp,
            unemp$CountryCode==substr(cy,1,3))
            [paste("YR",substr(cy,4,7), sep="")])
  }
  # create new variables for binary contrasts: greater than mean of values in study:
  disc$gdpbin <- ifelse(disc$gdp >= mean(disc$gdp, na.rm=TRUE), "high", "low")
  disc$unempbin <- ifelse(disc$unemp >= mean(disc$unemp, na.rm=TRUE), "high", "low")
  # ditto for implicit years:
  disc$gdpbin.imp <- ifelse(disc$gdp.imp >= mean(disc$gdp.imp, na.rm=TRUE), "high", "low")
  disc$unempbin.imp <- ifelse(disc$unemp.imp >= mean(disc$unemp.imp, na.rm=TRUE), "high", "low")
```

Unemployment

```
  table(disc$unempbin, disc$global) # N
```

```
##
##        global subgroup
##   high      7      133
##   low      14      182
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Unemployment=disc$unempbin), mean
```

```
##     Global Unemployment    x
## 1   global         high 1.50
## 2 subgroup         high 2.03
## 3   global          low 1.44
## 4 subgroup          low 1.68
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Unemployment=disc$unempbin), media
```

```
##     Global Unemployment    x
## 1   global         high 1.43
## 2 subgroup         high 1.44
## 3   global          low 1.44
## 4 subgroup          low 1.50
```

```
  t.test(relative.call.back.rate ~ unempbin, data=subset(disc, global == "subgroup"))
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by unempbin
## t = 1, df = 70, p-value = 0.2
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.225  0.913
## sample estimates:
## mean in group high  mean in group low
##               2.03               1.68
```

```
    # implicit:
  table(disc$unempbin.imp, disc$global) # N
```

```
##
##        global subgroup
##   high      9      180
##   low      15      202
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Unemployment=disc$unempbin.imp), m
```

```
##     Global Unemployment    x
## 1   global         high 1.84
## 2 subgroup         high 1.94
## 3   global          low 1.52
## 4 subgroup          low 1.85
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Unemployment=disc$unempbin.imp), m
```

```
##     Global Unemployment    x
## 1   global         high 1.52
## 2 subgroup         high 1.45
## 3   global          low 1.44
## 4 subgroup          low 1.60
```

```
  t.test(relative.call.back.rate ~ unempbin.imp, data=subset(disc, global == "subgroup"))
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by unempbin.imp
## t = 0.4, df = 90, p-value = 0.7
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.393  0.583
## sample estimates:
## mean in group high  mean in group low
##               1.94               1.85
```

Without Spain (outlier):

```r
table(disc$unempbin[disc$unemp <20], disc$global[disc$unemp <20]) # N
```

```
##
##        global subgroup
##   high      6      121
##   low      14      182
```

```r
aggregate(disc$relative.call.back.rate[disc$unemp <20], by=list(Global=disc$global[disc$unemp <20],
          Unemployment=disc$unempbin[disc$unemp <20]), mean, na.rm=TRUE) # SUPPLEMENT S11
```

```
##     Global Unemployment    x
## 1   global         high 1.56
## 2 subgroup         high 2.19
## 3   global          low 1.44
## 4 subgroup          low 1.68
```

```r
aggregate(disc$relative.call.back.rate[disc$unemp <20], by=list(Global=disc$global[disc$unemp <20],
          Unemployment=disc$unempbin[disc$unemp <20]), median, na.rm=TRUE) # SUPPLEMENT S11
```

```
##     Global Unemployment    x
## 1   global         high 1.52
## 2 subgroup         high 1.52
## 3   global          low 1.44
## 4 subgroup          low 1.50
```

```r
t.test(relative.call.back.rate ~ unempbin, data=subset(disc, global == "subgroup" & unemp < 20))
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by unempbin
## t = 1, df = 60, p-value = 0.1
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.177  1.179
## sample estimates:
## mean in group high  mean in group low
##               2.19               1.68
```

Annual GDP Growth

```r
table(disc$gdpbin, disc$global)
```

```
##
##        global subgroup
##   high     10      139
##   low      11      176
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, GDP=disc$gdpbin), mean, na.rm=TRU
```

```
##     Global  GDP    x
## 1   global high 1.40
## 2 subgroup high 1.90
## 3   global  low 1.51
## 4 subgroup  low 1.67
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, GDP=disc$gdpbin), median, na.rm=T
```

```
##     Global  GDP    x
## 1   global high 1.43
## 2 subgroup high 1.43
## 3   global  low 1.45
## 4 subgroup  low 1.54
```

```
  t.test(relative.call.back.rate ~ gdpbin, data=subset(disc, global == "subgroup"))
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by gdpbin
## t = 1, df = 100, p-value = 0.2
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.128  0.595
## sample estimates:
## mean in group high  mean in group low
##               1.90               1.67
```

```
  # implicit:
  table(disc$gdpbin.imp, disc$global)
```

```
##
##        global subgroup
##   high     12      183
##   low      12      199
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, GDP=disc$gdpbin.imp), mean, na.rm
```

```
##     Global  GDP    x
## 1   global high 1.70
## 2 subgroup high 2.06
## 3   global  low 1.51
## 4 subgroup  low 1.64
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, GDP=disc$gdpbin.imp), median, na.
```

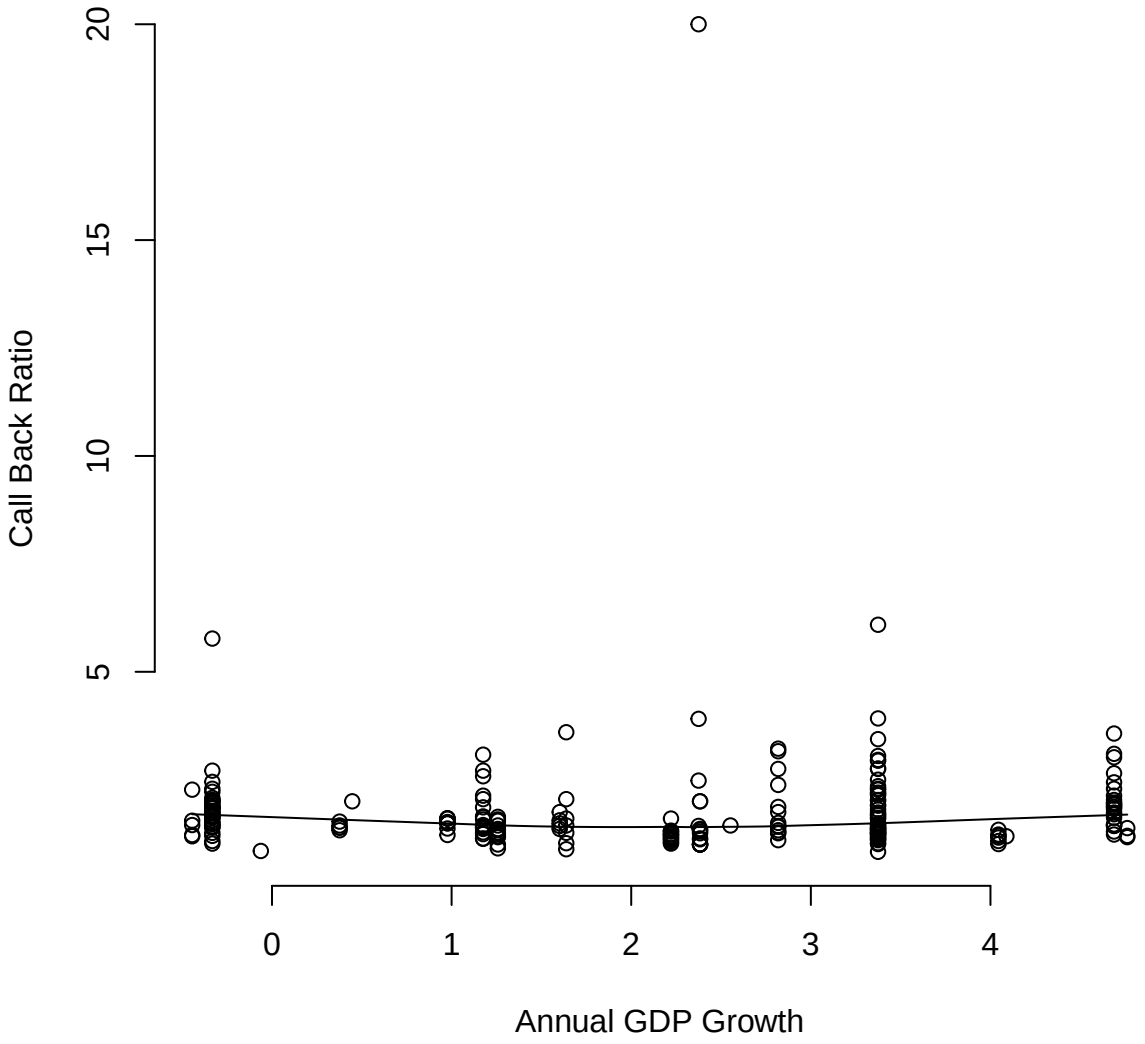
```
##     Global  GDP    x
## 1   global high 1.44
## 2 subgroup high 1.68
## 3   global  low 1.45
## 4 subgroup  low 1.51
```

```
t.test(relative.call.back.rate ~ gdpbin.imp, data=subset(disc, global == "subgroup"))
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by gdpbin.imp
## t = 3, df = 200, p-value = 0.01
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.0992 0.7398
## sample estimates:
## mean in group high  mean in group low
##               2.06               1.64
```

Scatter Plots

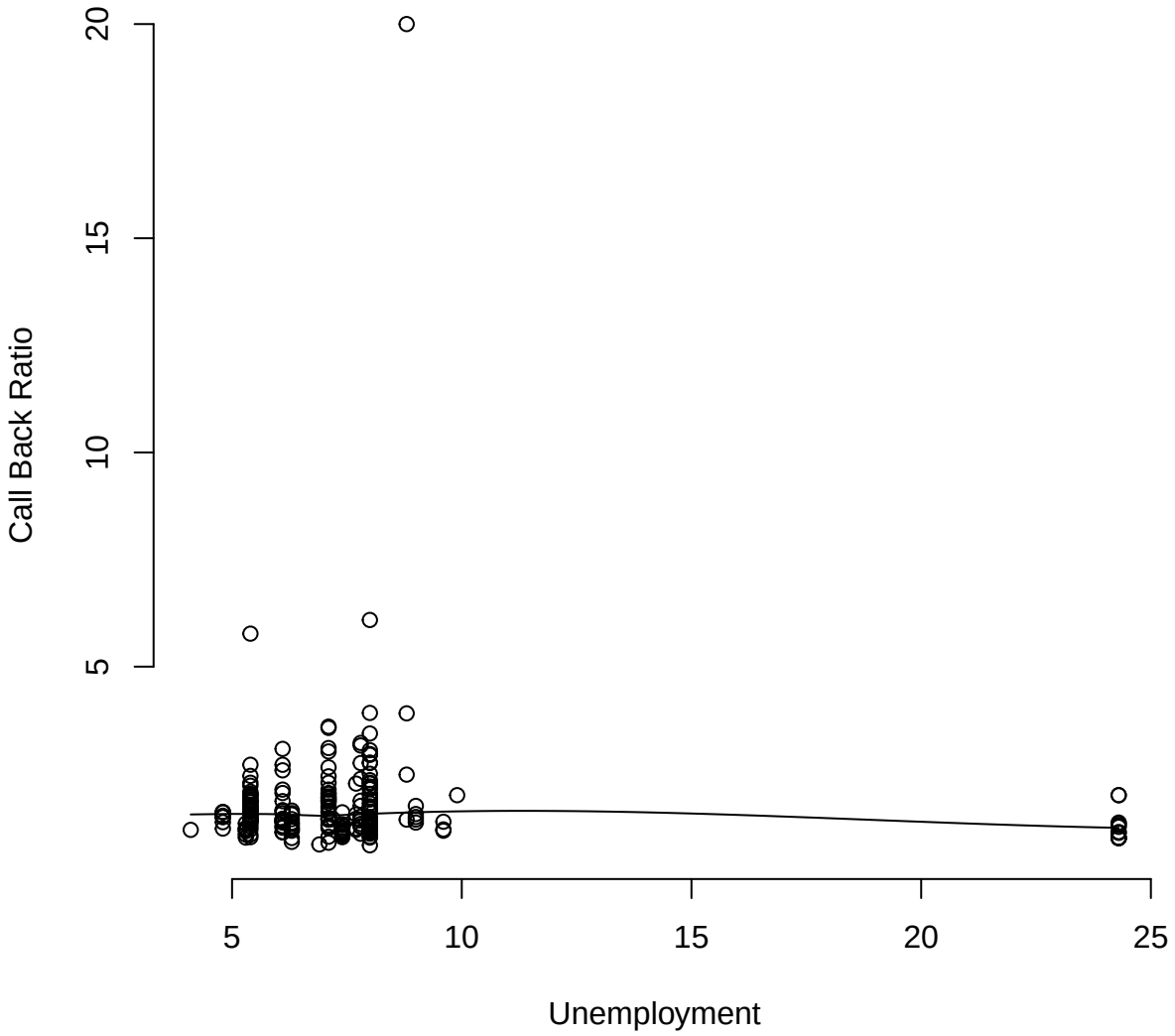
```
library(car)
scatter.smooth(disc$relative.call.back.rate ~ disc$gdp, bty="n", ylab="Call Back Ratio", xlab="Annual
```




```
  cor.test(disc$relative.call.back.rate, disc$gdp)    # SUPPLEMENT S11 (GDP)
```

```
##
##  Pearson's product-moment correlation
##
## data:  disc$relative.call.back.rate and disc$gdp
## t = 0.6, df = 200, p-value = 0.5
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.0886  0.1696
## sample estimates:
##    cor
## 0.0412
```

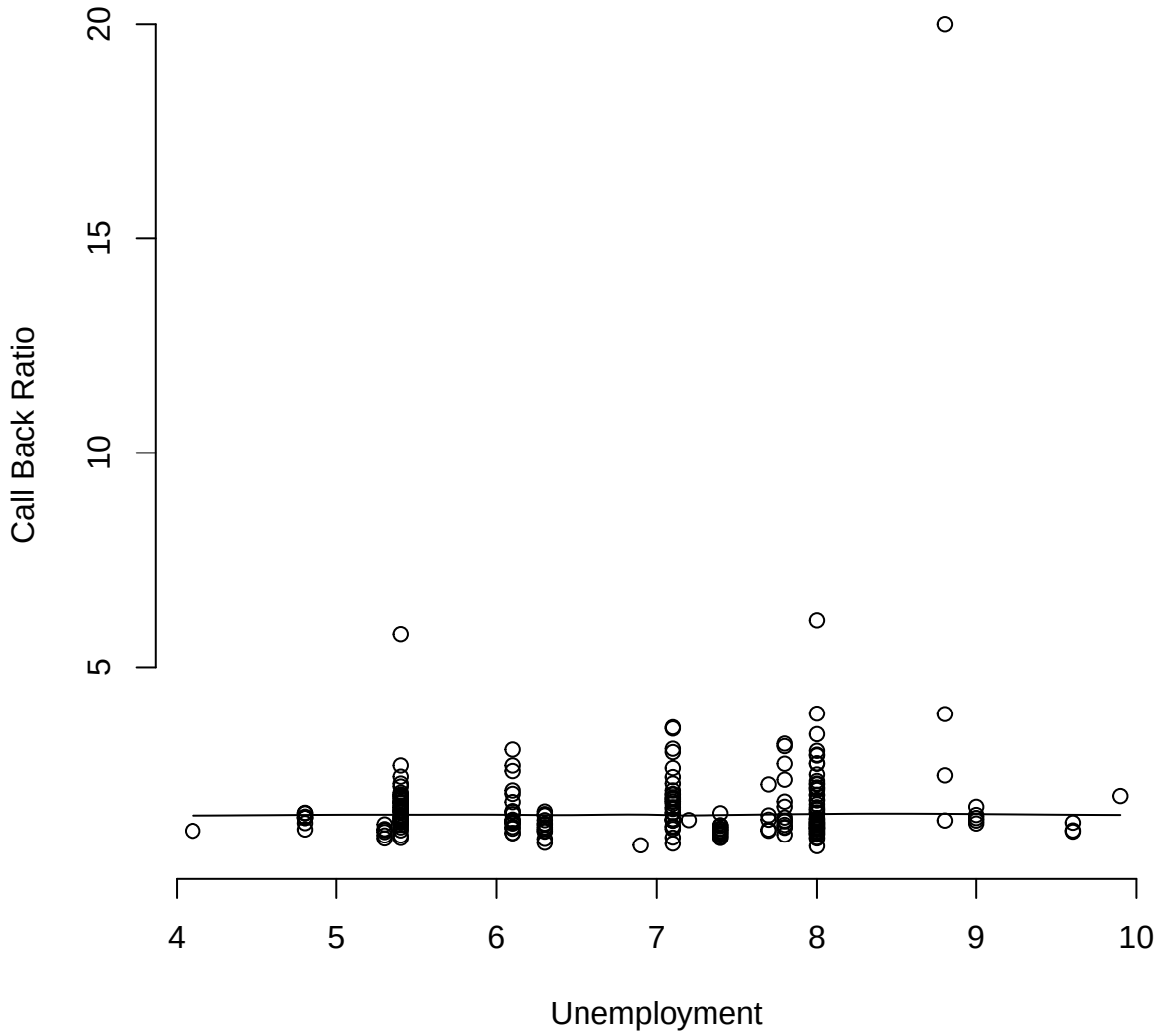
```
  scatter.smooth(disc$relative.call.back.rate ~ disc$unemp, bty="n", ylab="Call Back Ratio", xlab="Unemp
```




```
  cor.test(disc$relative.call.back.rate, disc$unemp) # SUPPLEMENT S11 (unemployment)
```

```
##
##  Pearson's product-moment correlation
##
## data:  disc$relative.call.back.rate and disc$unemp
## t = -0.7, df = 200, p-value = 0.5
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.1741  0.0841
## sample estimates:
##     cor
## -0.0458
```

```
  # removing outlier (Spain)
  scatter.smooth(disc$relative.call.back.rate[disc$unemp <20] ~ disc$unemp[disc$unemp <20], bty="n", yla
```




```
cor.test(disc$relative.call.back.rate[disc$unemp <20], disc$unemp[disc$unemp <20])
```

```
##
##  Pearson's product-moment correlation
##
## data:  disc$relative.call.back.rate[disc$unemp < 20] and disc$unemp[disc$unemp < 20]
## t = 2, df = 200, p-value = 0.04
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.00409 0.26553
## sample estimates:
##   cor
## 0.137
```

Gender: Male versus Female

```r
  table(disc$gender, disc$global)    # N
```

```
##
##                 global subgroup
##   both              17      232
##   female             4      110
##   male              15      256
##   not indicated      7      140
```

```r
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Gender=disc$gender), mean, na.rm="
```

```
##     Global        Gender    x
## 1   global          both 1.69
## 2 subgroup          both 1.70
## 3   global        female 1.34
## 4 subgroup        female 1.74
## 5   global          male 1.53
## 6 subgroup          male 1.63
## 7   global not indicated 1.42
## 8 subgroup not indicated 2.15
```

```r
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Gender=disc$gender), median, na.rm
```

```
##     Global        Gender    x
## 1   global          both 1.60
## 2 subgroup          both 1.49
## 3   global        female 1.39
## 4 subgroup        female 1.44
## 5   global          male 1.44
## 6 subgroup          male 1.37
## 7   global not indicated 1.31
## 8 subgroup not indicated 1.71
```

```r
  t.test(relative.call.back.rate ~ gender, data=subset(disc, global == "subgroup" & gender %in% c("male
```

```
##
##  Welch Two Sample t-test
##
## data:  relative.call.back.rate by gender
## t = 0.8, df = 90, p-value = 0.4
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.165  0.400
## sample estimates:
## mean in group female   mean in group male
##                 1.74                 1.63
```

Odds-Ratios

```r
# only subgroups
resM1 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$gender=="male" & disc$global=="subgroup"
              & !is.na(disc$negative.treatment.majority)
              & !is.na(disc$negative.treatment.minority) ), measure="OR",
  slab=paste(author, year, sep=", "), method="REML")
resM2 <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
  ci=positive.treatment.majority, di=negative.treatment.majority,
  data=subset(disc, disc$gender=="female" & disc$global=="subgroup"
              & !is.na(disc$negative.treatment.majority)
              & !is.na(disc$negative.treatment.minority) ), measure="OR",
  slab=paste(author, year, sep=", "), method="REML")
tab <- data.frame(exp(coef(resM1)), exp(coef(resM2)))
colnames(tab) <- c("male", "female")
kable(t(tab))
```

	intrcpt
male	0.587
female	0.543

Source of Jobs Adverts: Online versus Newspaper

```r
disc$asource <- NA
disc$asource[disc$source %in% c("newspaper",
  "newspaper ads, industry lists", "newspapers", "Suburban newspaper",
  "Washington Post")] <- "newspaper"
disc$asource[disc$source %in% c("newspaper and online",
  "online + newspaper", "online + newspapers",
  "online + some from newspapers")] <- "mixed"
disc$asource[disc$source %in% c("online", "online ")] <- "online"
aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Source=disc$asource), mean, na.rm=
```

```
##     Global    Source    x
## 1   global     mixed 1.59
## 2 subgroup     mixed 1.95
## 3   global newspaper 0.85
## 4 subgroup newspaper 1.98
## 5   global    online 1.44
## 6 subgroup    online 1.75
```

```r
aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Source=disc$asource), median, na.
```

```
##     Global    Source    x
## 1   global     mixed 1.49
## 2 subgroup     mixed 1.73
## 3   global newspaper 0.85
```

```
## 4 subgroup newspaper 1.98
## 5   global    online 1.42
## 6 subgroup    online 1.50
```

```
  summary(aov(relative.call.back.rate ~ asource, data=subset(disc, global == "subgroup")))
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## asource       2    2.7   1.349    1.69   0.19
## Residuals   321  256.0   0.798
## 414 observations deleted due to missingness
```

Firm Characteristics

```
  # firm size:
  disc$afirm <- NA
  disc$afirm[disc$firm %in% c("small", "small (<100)", "small (<50)",
    "small (<6)", "small company (0-15)")] <- "small"
  disc$afirm[disc$firm %in% c("medium", "medium (6-50)",
    "medium size company (16-50)", "medium-sized")] <- "medium"
  disc$afirm[disc$firm %in% c("big (>50)", "large", "large ",
    "large  (>500)", "large >500", "large company (more than 51)",
    "middle (100-499)")] <- "large"
  # private versus public employers
  disc$afirm[disc$firm %in% c("private", "Private", "private sector",
    "private sector ")] <- "private"
  disc$afirm[disc$firm %in% c("public", "Public",
    "pubic sector")] <- "public"

  table(disc$afirm, disc$global) # N
```

```
##
##           global subgroup
##   large        0        8
##   medium       0        4
##   private      0        5
##   public       0        1
##   small        0        9
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Characteristics=disc$afirm), mean
```

```
##     Global Characteristics    x
## 1 subgroup           large 1.47
## 2 subgroup          medium 1.18
## 3 subgroup         private 1.65
## 4 subgroup          public 1.19
## 5 subgroup           small 1.48
```

```
  aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Characteristics=disc$afirm), media
```

```
##     Global Characteristics    x
## 1 subgroup           large 1.39
## 2 subgroup          medium 1.10
## 3 subgroup         private 1.57
## 4 subgroup          public 1.19
## 5 subgroup           small 1.29
```

```
  summary(aov(relative.call.back.rate ~ afirm, data=subset(disc, global == "subgroup" & afirm %in% c("p
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## afirm        1  0.175  0.1748    3.44   0.14
## Residuals    4  0.203  0.0508
```

```
  summary(aov(relative.call.back.rate ~ afirm, data=subset(disc, global == "subgroup" & afirm %in% c("s
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## afirm        2   0.30   0.148    0.73    0.5
## Residuals   18   3.64   0.202
```

Publication Bias

Using a meta-analysis, it is also possible to enumerate the estimated publication bias. The follwing figures are so-called funnel plots, indicating as black dots the estimates for each of the studies, and as white dots the studies that are needed to obtain a balanced picture. The presence of white dots in the figure indicates that conventional tests of publication bias indicate clearly that publication bias is likely, with some studies finding less discrimination or no discrimination 'missing'. While publication bias is usually linked to the non-publication of studies failing to produce 'statistically significant' results, we are puzzled about this particular finding. Given that studies report discrimination against minority groups rather consistently, we suspect that a study finding no difference between the minority and majority population, or even one that indicates positive discrimination in favour of the minority groups would actually be *easier* to publish.

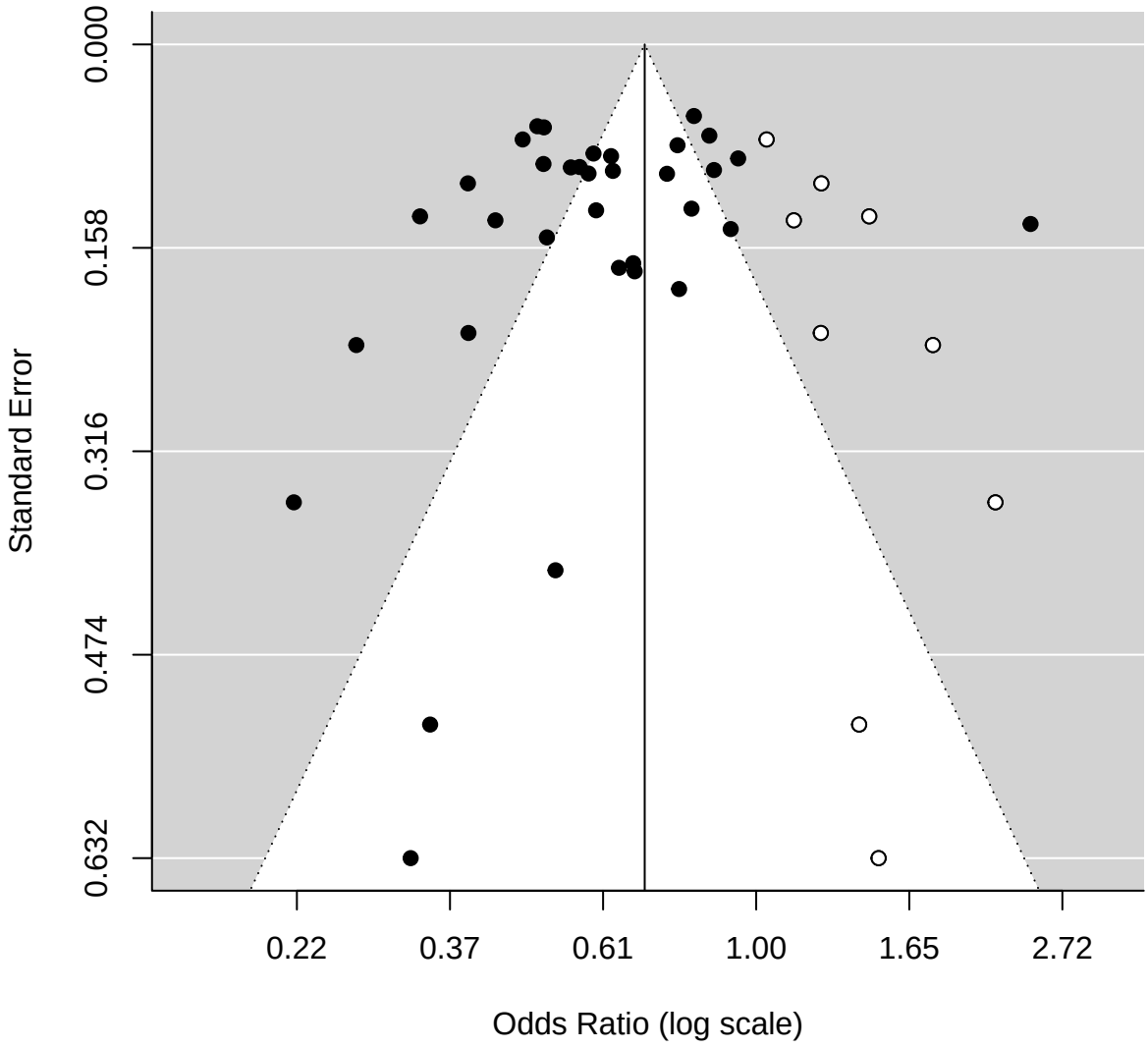
Study Level

```
  taf <- trimfill(resGlobal)   # TRIM & FILL
  summary(taf)
```

```
##
## Estimated number of missing studies on the right side: 9 (SE = 3.8189)
##
## Random-Effects Model (k = 43; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -31.8536   63.7072   67.7072   71.1826   68.0149
##
## tau^2 (estimated amount of total heterogeneity): 0.2055 (SE = 0.0517)
```

```
## tau (square root of estimated tau^2 value):      0.4533
## I^2 (total heterogeneity / total variability):   94.37%
## H^2 (total variability / sampling variability):  17.77
##
## Test for Heterogeneity:
## Q(df = 42) = 493.6423, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##  -0.3641   0.0750  -4.8554   <.0001  -0.5111  -0.2171      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

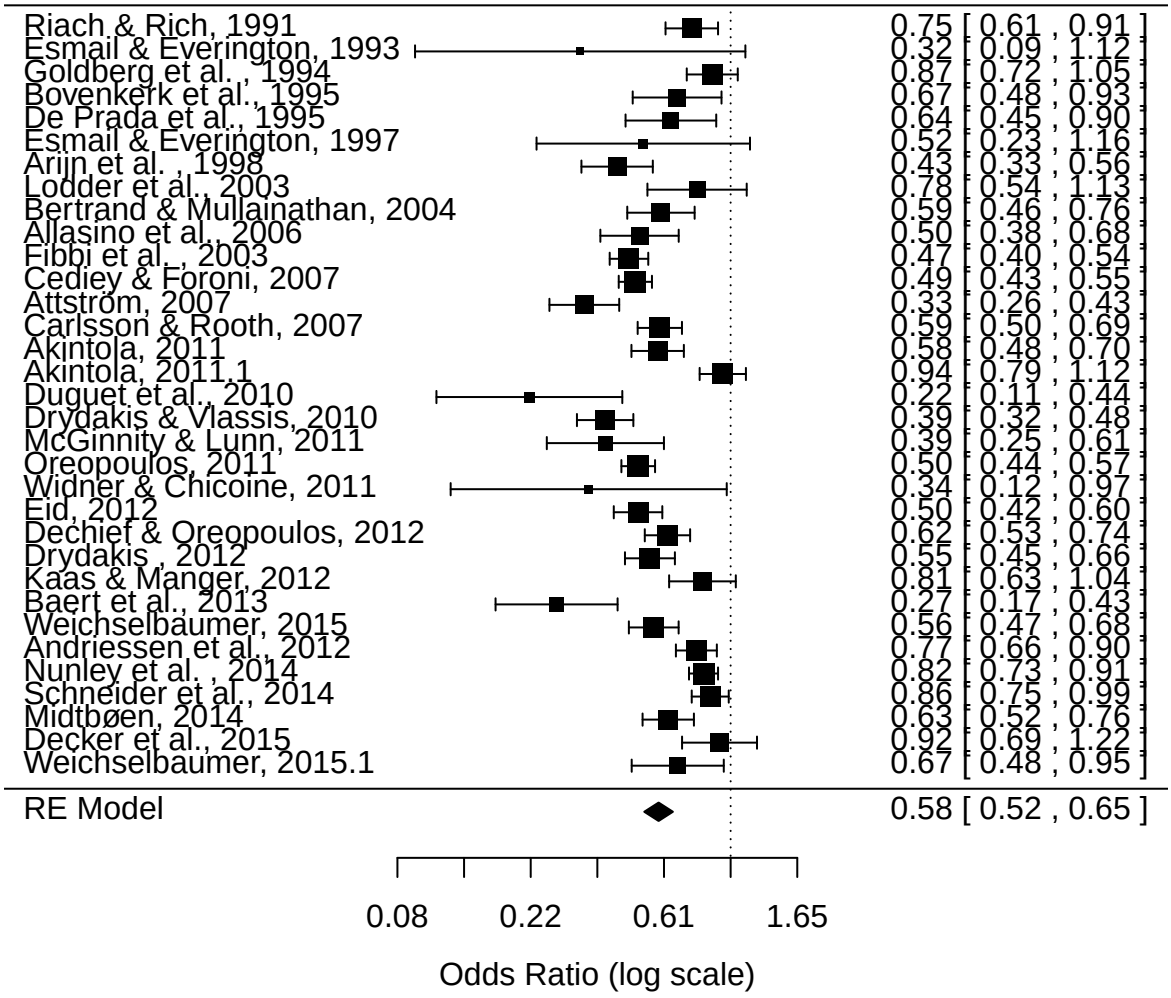
```
funnel(taf, atransf=exp) ### draw funnel plot with missing studies filled in
```



```r
  # only global studies; without the outlier:
  resGlobal.mod <- rma(ai=positive.treatment.minority, bi=negative.treatment.minority,
    ci=positive.treatment.majority, di=negative.treatment.majority,
    data=subset(disc, disc$global=="global" & !is.na(disc$negative.treatment.majority)
    & !is.na(disc$negative.treatment.minority)  & disc$author != "Bendick et al."), measure="OR",
    slab=paste(author, year, sep=", "), method="REML")
  summary(resGlobal.mod)
```

```
##
## Random-Effects Model (k = 33; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -10.6068   21.2137   25.2137   28.1452   25.6275
##
## tau^2 (estimated amount of total heterogeneity): 0.0764 (SE = 0.0240)
## tau (square root of estimated tau^2 value):      0.2764
## I^2 (total heterogeneity / total variability):   87.52%
## H^2 (total variability / sampling variability):  8.01
##
## Test for Heterogeneity:
## Q(df = 32) = 231.2032, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##  -0.5404   0.0549  -9.8521   <.0001  -0.6479  -0.4329      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
  forest(resGlobal.mod, atransf=exp)
```

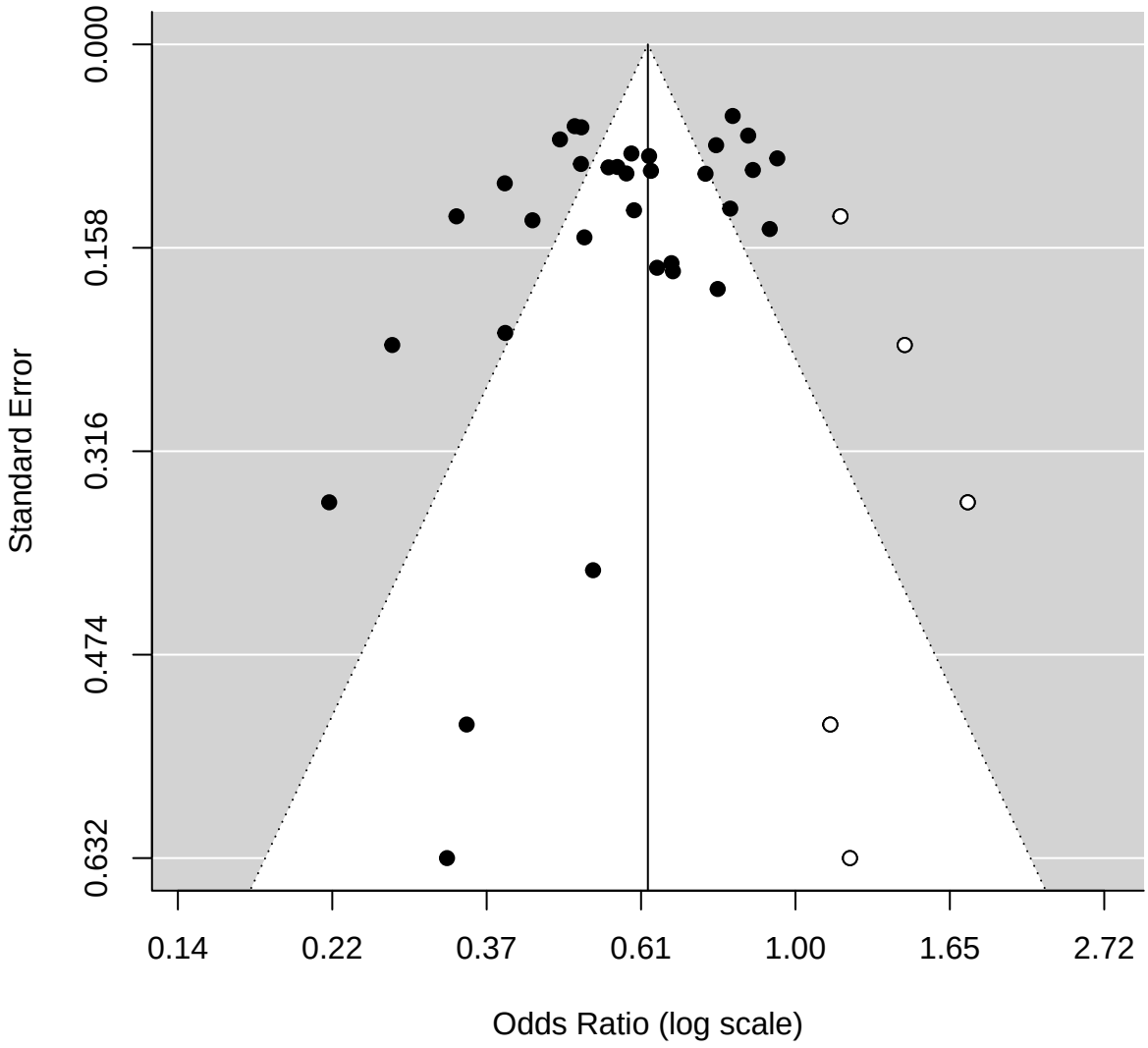



```
taf.mod <- trimfill(resGlobal.mod)    # TRIM & FILL
summary(taf.mod) # SUPPLEMENT S12
```

```
##
## Estimated number of missing studies on the right side: 5 (SE = 3.8084)
##
## Random-Effects Model (k = 38; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -19.6494   39.2988   43.2988   46.5206   43.6517
##
## tau^2 (estimated amount of total heterogeneity): 0.1076 (SE = 0.0312)
## tau (square root of estimated tau^2 value):      0.3280
## I^2 (total heterogeneity / total variability):   89.81%
## H^2 (total variability / sampling variability):  9.81
##
## Test for Heterogeneity:
## Q(df = 37) = 276.2092, p-val < .0001
##
## Model Results:
```

```
##
## estimate       se     zval     pval    ci.lb    ci.ub
##  -0.4779   0.0606  -7.8800   <.0001  -0.5967  -0.3590      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
funnel(taf.mod, atransf=exp)
```



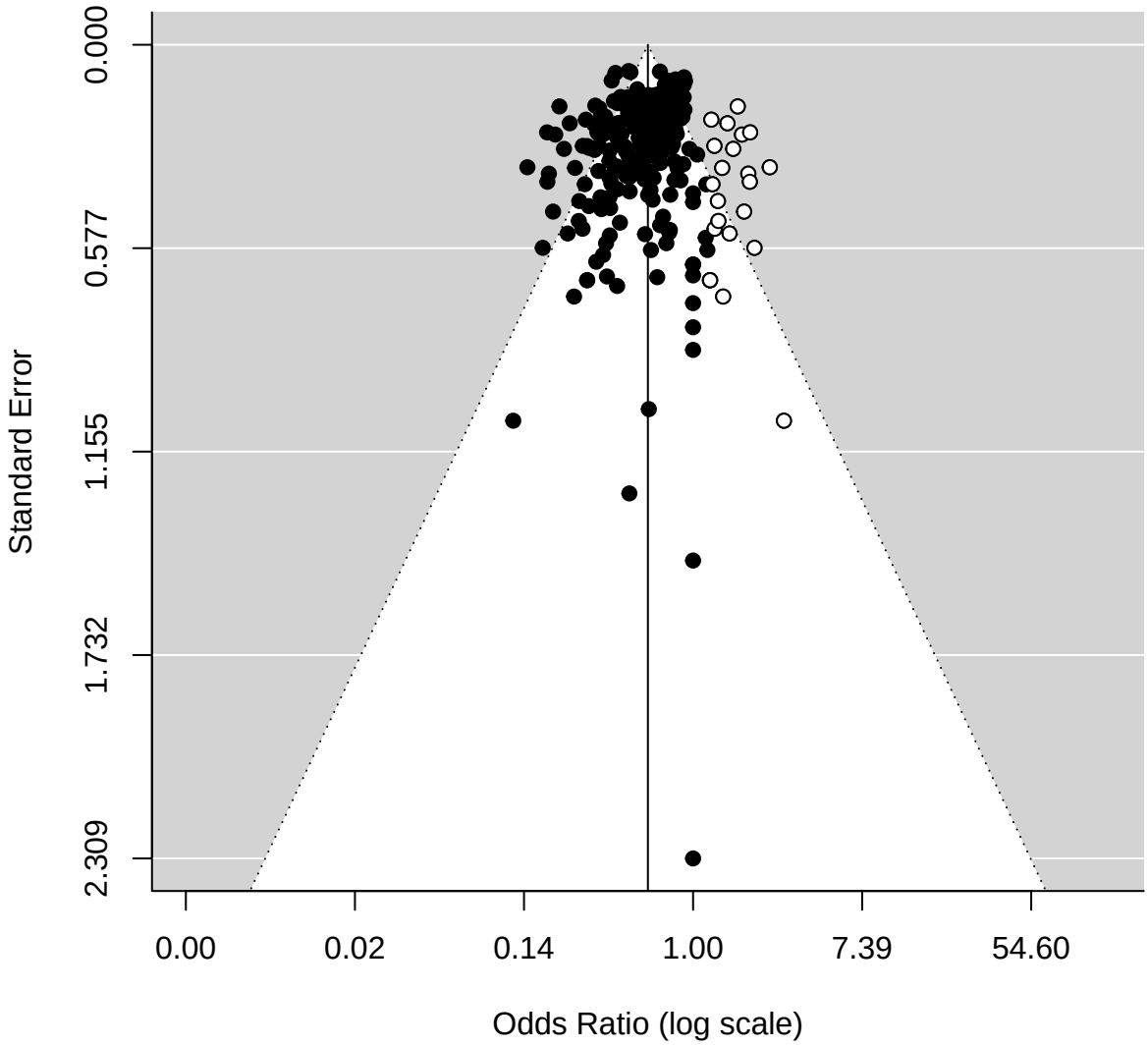

```
# savePlot("funnel-study.png") # SUPPLEMENT S12
```

Subgroups

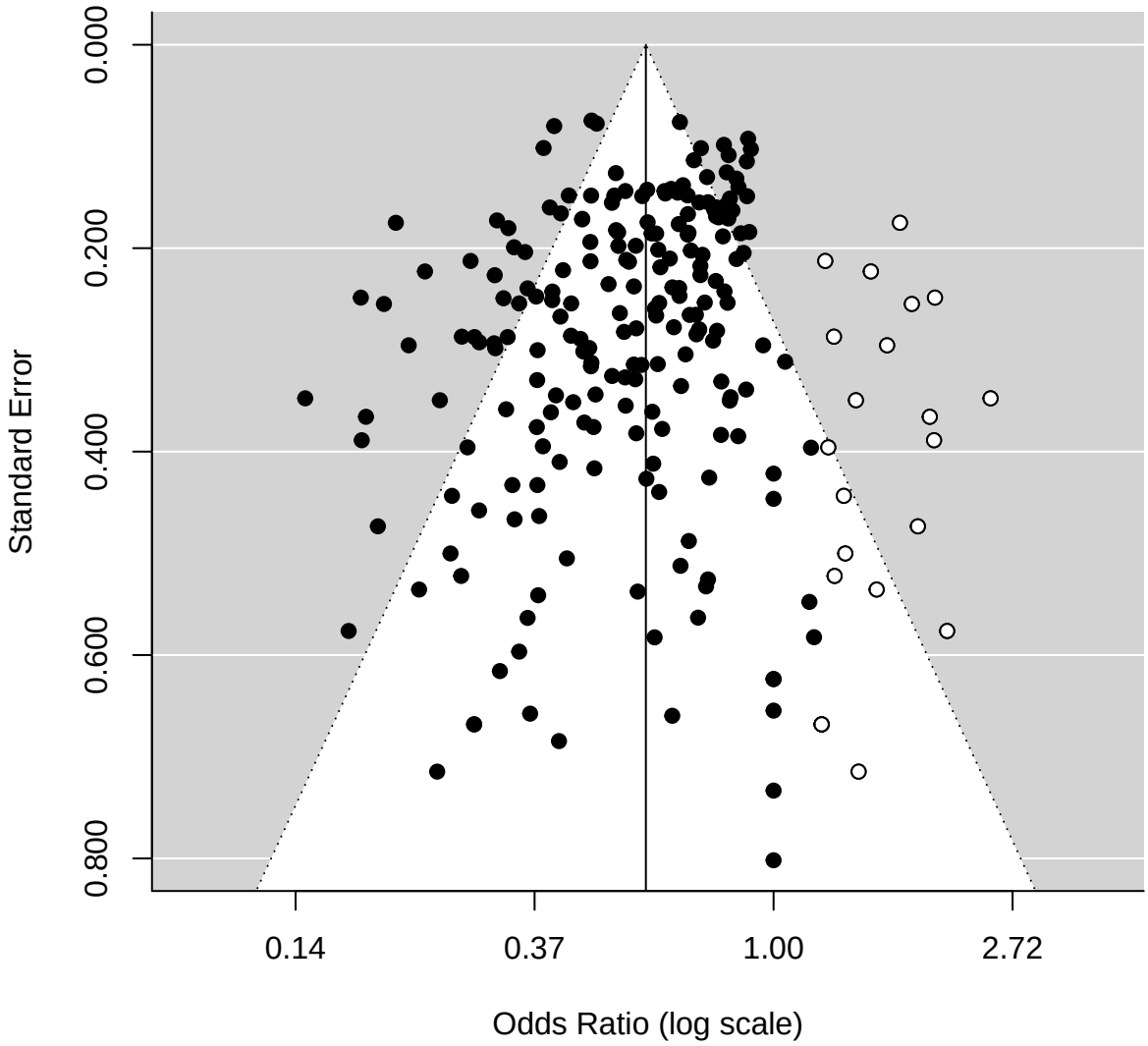
```
taf <- trimfill(resSub) # trim & fill
summary(taf) # SUPPLEMENT S12
```

```
##
## Estimated number of missing studies on the right side: 22 (SE = 9.3887)
##
## Random-Effects Model (k = 237; tau^2 estimator: REML)
##
##    logLik   deviance        AIC        BIC       AICc
## -178.0571   356.1141   360.1141   367.0418   360.1656
##
## tau^2 (estimated amount of total heterogeneity): 0.1440 (SE = 0.0204)
## tau (square root of estimated tau^2 value):      0.3795
## I^2 (total heterogeneity / total variability):   75.76%
## H^2 (total variability / sampling variability):  4.13
##
## Test for Heterogeneity:
## Q(df = 236) = 848.8932, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##  -0.5343   0.0315 -16.9800   <.0001  -0.5960  -0.4726      ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
funnel(taf, atransf=exp)
```




```
funnel(taf, atransf=exp, ylim=c(0, .8)) # set ylim to see something; cuts 5 studies (largest SE)
```




```
# savePlot("funnel-subgroup.png") # SUPPLEMENT S12
```

Multivariate Regression Analysis

Results are reported including the Knapp and Hartung adjustment.

Global

```
# metafor doesn't like characters (sometimes), and as.numeric() doesn't seem to work here, either:
# EU/non-EU:
disc$EU.num <- NA
disc$EU.num[disc$EU == "European Union"] <- 1
disc$EU.num[disc$EU == "non EU"] <- 0
# US/Canada versus rest
disc$USC.num <- NA
```

```r
  disc$USC.num[!is.na(disc$country)] <- 0              # country defined
  disc$USC.num[disc$continent == "US/Canada"] <- 1     # US and Canada
  # 2nd generation:
  disc$g2.explicit.num <- NA
  disc$g2.explicit.num[!disc$firstgen.explicit == ""] <- 0        # excluding implicit
  disc$g2.explicit.num[disc$firstgen.explicit == "second"] <- 1   # explicit G2 only
      # table(disc$firstgen.explicit, disc$g2.explicit.num)
  disc$g2.implicit.num <- NA
  disc$g2.implicit.num[!disc$firstgen.implicit == ""] <- 0        # excluding implicit
  disc$g2.implicit.num[disc$firstgen.implicit == "second"] <- 1   # explicit G2 only
      # table(disc$firstgen.implicit, disc$g2.implicit.num)
  # time dummy: after 2006
  disc$yg <- 0
  disc$yg[disc$yearGroup %in% c(2006, 2011)] <- 1
      # table(disc$yg, disc$yearGroup)
  disc$skills[disc$skills==""] <- NA # properly treating NA
  disc$agender <- disc$gender
  disc$agender[disc$gender=="not indicated"] <- NA
  # regression analysis:
   ord <- escalc(ai=positive.treatment.minority, bi=negative.treatment.minority, ci=positive.treatment.m
     RG1 <- rma(data=ord, yi ~ skills, vi=vi, knha=TRUE) # Knapp & Hartung adjustment
```

```
## Warning in rma(data = ord, yi ~ skills, vi = vi, knha = TRUE): Studies with
## NAs omitted from model fitting.
```

```
## Warning in rma(data = ord, yi ~ skills, vi = vi, knha = TRUE): Redundant
## predictors dropped from the model.
```

```r
  summary(RG1)
```

```
##
## Mixed-Effects Model (k = 30; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -17.3556   34.7112   42.7112   47.8946   44.5294
##
## tau^2 (estimated amount of residual heterogeneity):     0.1696 (SE = 0.0534)
## tau (square root of estimated tau^2 value):             0.4119
## I^2 (residual heterogeneity / unaccounted variability): 93.35%
## H^2 (unaccounted variability / sampling variability):   15.04
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 27) = 273.1182, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3):
## F(df1 = 2, df2 = 27) = 0.2502, p-val = 0.7804
##
## Model Results:
##
##             estimate      se     tval    pval    ci.lb    ci.ub
## intrcpt      -0.6080  0.1571  -3.8710  0.0006  -0.9303  -0.2857  ***
```

```
## skillshigh    0.1578  0.2241   0.7042  0.4873  -0.3020   0.6176
## skillslow     0.0665  0.1990   0.3342  0.7408  -0.3417   0.4747
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
    RG2 <- rma(data=ord, yi ~ skills + continent, vi=vi, knha=TRUE) # Knapp & Hartung adjustment
```

```
## Warning in rma(data = ord, yi ~ skills + continent, vi = vi, knha = TRUE):
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(data = ord, yi ~ skills + continent, vi = vi, knha = TRUE):
## Redundant predictors dropped from the model.
```

```
  summary(RG2)
```

```
##
## Mixed-Effects Model (k = 29; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -14.6261   29.2522   39.2522   45.3466   42.4101
##
## tau^2 (estimated amount of residual heterogeneity):     0.1482 (SE = 0.0493)
## tau (square root of estimated tau^2 value):             0.3849
## I^2 (residual heterogeneity / unaccounted variability): 92.50%
## H^2 (unaccounted variability / sampling variability):   13.34
## R^2 (amount of heterogeneity accounted for):            10.23%
##
## Test for Residual Heterogeneity:
## QE(df = 25) = 249.1466, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## F(df1 = 3, df2 = 25) = 1.9345, p-val = 0.1498
##
## Model Results:
##
##                     estimate      se     tval    pval    ci.lb    ci.ub
## intrcpt              -0.7065  0.1542  -4.5807  0.0001  -1.0241  -0.3888
## skillshigh            0.0753  0.2136   0.3523  0.7275  -0.3647   0.5152
## skillslow             0.0511  0.1897   0.2695  0.7898  -0.3396   0.4418
## continentUS/Canada    0.3835  0.1675   2.2894  0.0308   0.0385   0.7284
##
## intrcpt             ***
## skillshigh
## skillslow
## continentUS/Canada    *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Subgroups

```
  # preparing data:
  ord <- escalc(ai=positive.treatment.minority, bi=negative.treatment.minority, ci=positive.treatment.ma
  # multivariate regressions:
  RS1 <- rma(data=ord, yi ~ skills, vi=vi, knha=TRUE) # Knapp & Hartung adjustment
```

```
## Warning in rma(data = ord, yi ~ skills, vi = vi, knha = TRUE): Studies with
## NAs omitted from model fitting.
```

```
## Warning in rma(data = ord, yi ~ skills, vi = vi, knha = TRUE): Redundant
## predictors dropped from the model.
```

```
  summary(RS1) # TEXT (on skills), SUPPLEMENT S13 (M1)
```

```
##
## Mixed-Effects Model (k = 181; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -89.3547  178.7095  186.7095  199.4366  186.9407
##
## tau^2 (estimated amount of residual heterogeneity):     0.0603 (SE = 0.0134)
## tau (square root of estimated tau^2 value):             0.2456
## I^2 (residual heterogeneity / unaccounted variability): 54.68%
## H^2 (unaccounted variability / sampling variability):   2.21
## R^2 (amount of heterogeneity accounted for):            38.92%
##
## Test for Residual Heterogeneity:
## QE(df = 178) = 363.0930, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3):
## F(df1 = 2, df2 = 178) = 27.8614, p-val < .0001
##
## Model Results:
##
##             estimate      se     tval    pval    ci.lb    ci.ub
## intrcpt      -0.6830  0.0691  -9.8840  <.0001  -0.8193  -0.5466  ***
## skillshigh    0.2802  0.0821   3.4136  0.0008   0.1182   0.4421  ***
## skillslow    -0.1615  0.0795  -2.0302  0.0438  -0.3184  -0.0045    *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
  RS2 <- rma(data=ord, yi ~ skills + g2.implicit.num, vi=vi, knha=TRUE)
```

```
## Warning in rma(data = ord, yi ~ skills + g2.implicit.num, vi = vi, knha =
## TRUE): Studies with NAs omitted from model fitting.
```

```
## Warning in rma(data = ord, yi ~ skills + g2.implicit.num, vi = vi, knha =
## TRUE): Redundant predictors dropped from the model.
```

```r
summary(RS2) # TEXT (on G2), SUPPLEMENT S13 (M2)
```

```
##
## Mixed-Effects Model (k = 158; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -70.8049  141.6099  151.6099  166.7946  152.0153
##
## tau^2 (estimated amount of residual heterogeneity):     0.0535 (SE = 0.0135)
## tau (square root of estimated tau^2 value):             0.2313
## I^2 (residual heterogeneity / unaccounted variability): 51.09%
## H^2 (unaccounted variability / sampling variability):   2.04
## R^2 (amount of heterogeneity accounted for):            49.28%
##
## Test for Residual Heterogeneity:
## QE(df = 154) = 289.0030, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## F(df1 = 3, df2 = 154) = 26.8207, p-val < .0001
##
## Model Results:
##
##                  estimate      se     tval    pval    ci.lb    ci.ub
## intrcpt           -0.3302  0.1119  -2.9506  0.0037  -0.5514  -0.1091   **
## skillshigh         0.3094  0.0793   3.9032  0.0001   0.1528   0.4659  ***
## skillslow         -0.2795  0.0802  -3.4835  0.0006  -0.4380  -0.1210  ***
## g2.implicit.num   -0.3542  0.0913  -3.8796  0.0002  -0.5346  -0.1739  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
RS3 <- rma(data=ord, yi ~ skills + g2.implicit.num + continent, vi=vi, knha=TRUE)
```

```
## Warning in rma(data = ord, yi ~ skills + g2.implicit.num + continent, vi =
## vi, : Studies with NAs omitted from model fitting.

## Warning in rma(data = ord, yi ~ skills + g2.implicit.num + continent, vi =
## vi, : Redundant predictors dropped from the model.
```

```r
summary(RS3)
```

```
##
## Mixed-Effects Model (k = 140; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -62.1356  124.2712  134.2712  148.8345  134.7328
##
## tau^2 (estimated amount of residual heterogeneity):     0.0592 (SE = 0.0148)
## tau (square root of estimated tau^2 value):             0.2434
## I^2 (residual heterogeneity / unaccounted variability): 55.38%
## H^2 (unaccounted variability / sampling variability):   2.24
```

```
## R^2 (amount of heterogeneity accounted for):            49.87%
##
## Test for Residual Heterogeneity:
## QE(df = 136) = 279.2592, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## F(df1 = 3, df2 = 136) = 25.8873, p-val < .0001
##
## Model Results:
##
##                  estimate      se     tval    pval    ci.lb    ci.ub
## intrcpt           -0.1451  0.1562  -0.9291  0.3545  -0.4540   0.1638
## skillshigh         0.3063  0.0827   3.7042  0.0003   0.1428   0.4698  ***
## skillslow         -0.2840  0.0837  -3.3932  0.0009  -0.4495  -0.1185  ***
## g2.implicit.num   -0.5381  0.1408  -3.8229  0.0002  -0.8164  -0.2597  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RS3b <- rma(data=ord, yi ~ skills + g2.implicit.num + continent + relevel(agender, ref="male"), vi=vi
```

```
## Warning in rma(data = ord, yi ~ skills + g2.implicit.num + continent +
## relevel(agender, : Studies with NAs omitted from model fitting.
```

```
## Warning in rma(data = ord, yi ~ skills + g2.implicit.num + continent +
## relevel(agender, : Redundant predictors dropped from the model.
```

```
summary(RS3b) # SUPPLEMENT S13 (M3)
```

```
##
## Mixed-Effects Model (k = 137; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -56.5168  113.0337  127.0337  147.1600  127.9442
##
## tau^2 (estimated amount of residual heterogeneity):     0.0487 (SE = 0.0133)
## tau (square root of estimated tau^2 value):             0.2207
## I^2 (residual heterogeneity / unaccounted variability): 50.74%
## H^2 (unaccounted variability / sampling variability):   2.03
## R^2 (amount of heterogeneity accounted for):            59.39%
##
## Test for Residual Heterogeneity:
## QE(df = 131) = 243.6458, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4,5,6):
## F(df1 = 5, df2 = 131) = 18.5078, p-val < .0001
##
## Model Results:
##
##                                    estimate      se     tval    pval
## intrcpt                             -0.0623  0.1550  -0.4022  0.6882
## skillshigh                           0.1745  0.0953   1.8318  0.0693
```

```
## skillslow                               -0.3699  0.0854  -4.3325  <.0001
## g2.implicit.num                         -0.4537  0.1449  -3.1318  0.0021
## relevel(agender, ref = "male")both      -0.2367  0.0776  -3.0496  0.0028
## relevel(agender, ref = "male")female     0.0309  0.1022   0.3020  0.7632
##                                           ci.lb    ci.ub
## intrcpt                                 -0.3690   0.2443
## skillshigh                              -0.0140   0.3630    .
## skillslow                               -0.5388  -0.2010  ***
## g2.implicit.num                         -0.7403  -0.1671   **
## relevel(agender, ref = "male")both      -0.3903  -0.0832   **
## relevel(agender, ref = "male")female    -0.1714   0.2331
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RS4 <- rma(data=ord, yi ~ continent, vi=vi, knha=TRUE)
```

```
## Warning in rma(data = ord, yi ~ continent, vi = vi, knha = TRUE): Studies
## with NAs omitted from model fitting.
```

```
summary(RS4)
```

```
##
## Mixed-Effects Model (k = 187; tau^2 estimator: REML)
##
##    logLik   deviance        AIC        BIC       AICc
## -112.3169   224.6338   230.6338   240.2948   230.7664
##
## tau^2 (estimated amount of residual heterogeneity):     0.1037 (SE = 0.0175)
## tau (square root of estimated tau^2 value):             0.3220
## I^2 (residual heterogeneity / unaccounted variability): 71.39%
## H^2 (unaccounted variability / sampling variability):   3.50
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 185) = 618.0908, p-val < .0001
##
## Test of Moderators (coefficient(s) 2):
## F(df1 = 1, df2 = 185) = 0.2933, p-val = 0.5887
##
## Model Results:
##
##                     estimate      se      tval    pval    ci.lb    ci.ub
## intrcpt              -0.6100  0.0324  -18.8394  <.0001  -0.6739  -0.5462
## continentUS/Canada   -0.0580  0.1070   -0.5416  0.5887  -0.2691   0.1532
##
## intrcpt              ***
## continentUS/Canada
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
  RS5 <- rma(data=ord, yi ~ skills + German, vi=vi, knha=TRUE)
```

```
## Warning in rma(data = ord, yi ~ skills + German, vi = vi, knha = TRUE):
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(data = ord, yi ~ skills + German, vi = vi, knha = TRUE):
## Redundant predictors dropped from the model.
```

```
  summary(RS5) # SUPPLEMENT S13 (M4)
```

```
##
## Mixed-Effects Model (k = 181; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -81.8727  163.7454  173.7454  189.6262  174.0963
##
## tau^2 (estimated amount of residual heterogeneity):     0.0514 (SE = 0.0122)
## tau (square root of estimated tau^2 value):             0.2268
## I^2 (residual heterogeneity / unaccounted variability): 50.48%
## H^2 (unaccounted variability / sampling variability):   2.02
## R^2 (amount of heterogeneity accounted for):            47.93%
##
## Test for Residual Heterogeneity:
## QE(df = 177) = 331.2343, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## F(df1 = 3, df2 = 177) = 26.1682, p-val < .0001
##
## Model Results:
##
##              estimate      se     tval    pval    ci.lb    ci.ub
## intrcpt       -0.3882  0.0970  -4.0017  <.0001  -0.5796  -0.1968  ***
## skillshigh     0.1474  0.0844   1.7467  0.0824  -0.0191   0.3140    .
## skillslow     -0.1806  0.0753  -2.3968  0.0176  -0.3293  -0.0319    *
## Germanother   -0.2968  0.0720  -4.1230  <.0001  -0.4388  -0.1547  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
  RS6 <- rma(data=ord, yi ~ skills + minor, vi=vi, knha=TRUE)
```

```
## Warning in rma(data = ord, yi ~ skills + minor, vi = vi, knha = TRUE):
## Studies with NAs omitted from model fitting.
```

```
## Warning in rma(data = ord, yi ~ skills + minor, vi = vi, knha = TRUE):
## Redundant predictors dropped from the model.
```

```
  summary(RS6) # SUPPLEMENT S13 (M5)
```

```
##
## Mixed-Effects Model (k = 92; tau^2 estimator: REML)
##
##   logLik  deviance       AIC       BIC      AICc
## -11.6016   23.2033   35.2033   49.9987   36.2533
##
## tau^2 (estimated amount of residual heterogeneity):     0.0060 (SE = 0.0061)
## tau (square root of estimated tau^2 value):             0.0777
## I^2 (residual heterogeneity / unaccounted variability): 12.73%
## H^2 (unaccounted variability / sampling variability):   1.15
## R^2 (amount of heterogeneity accounted for):            95.44%
##
## Test for Residual Heterogeneity:
## QE(df = 87) = 97.1726, p-val = 0.2139
##
## Test of Moderators (coefficient(s) 2,3,4,5):
## F(df1 = 4, df2 = 87) = 53.1197, p-val < .0001
##
## Model Results:
##
##                        estimate      se      tval    pval    ci.lb
## intrcpt                 -0.8311  0.0487  -17.0820  <.0001  -0.9278
## skillshigh               0.1402  0.0922    1.5210  0.1319  -0.0430
## skillslow               -0.2375  0.0678   -3.5044  0.0007  -0.3721
## minorIndian/Pakistani    0.3139  0.0817    3.8422  0.0002   0.1515
## minorTurk                0.5004  0.0834    5.9979  <.0001   0.3346
##                         ci.ub
## intrcpt                -0.7344  ***
## skillshigh              0.3234
## skillslow              -0.1028  ***
## minorIndian/Pakistani   0.4762  ***
## minorTurk               0.6663  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
  # multilevel:
  RM1 <- rma.mv(yi=yi, V=vi, data=ord, random = ~ 1 | study, mods= ~ skills)
```

```
## Warning in rma.mv(yi = yi, V = vi, data = ord, random = ~1 | study, mods =
## ~skills): Rows with NAs omitted from model fitting.
```

```
## Warning in rma.mv(yi = yi, V = vi, data = ord, random = ~1 | study, mods =
## ~skills): Redundant predictors dropped from the model.
```

```
  summary(RM1) # removed from SUPPLEMENT S13
```

```
##
## Multivariate Meta-Analysis Model (k = 181; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -76.5544  153.1087  161.1087  173.8359  161.3399
```

```
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed  factor
## sigma^2    0.0449  0.2119     20     no   study
##
## Test for Residual Heterogeneity:
## QE(df = 178) = 363.0930, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3):
## QM(df = 2) = 13.7456, p-val = 0.0010
##
## Model Results:
##
##             estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt      -0.5483  0.0784  -6.9902  <.0001  -0.7020  -0.3946  ***
## skillshigh    0.0940  0.0739   1.2712  0.2036  -0.0509   0.2389
## skillslow    -0.1392  0.0697  -1.9982  0.0457  -0.2757  -0.0027    *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
RM2 <- rma.mv(yi=yi, V=vi, data=ord, random = ~ 1 | study, mods= ~ skills + continent)
```

```
## Warning in rma.mv(yi = yi, V = vi, data = ord, random = ~1 | study, mods =
## ~skills + : Rows with NAs omitted from model fitting.
```

```
## Warning in rma.mv(yi = yi, V = vi, data = ord, random = ~1 | study, mods =
## ~skills + : Redundant predictors dropped from the model.
```

```
summary(RM2)
```

```
##
## Multivariate Meta-Analysis Model (k = 163; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -70.1449  140.2898  150.2898  165.6343  150.6819
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed  factor
## sigma^2    0.0510  0.2258     19     no   study
##
## Test for Residual Heterogeneity:
## QE(df = 159) = 351.1051, p-val < .0001
##
## Test of Moderators (coefficient(s) 2,3,4):
## QM(df = 3) = 12.6585, p-val = 0.0054
##
## Model Results:
##
##                     estimate      se     zval    pval    ci.lb    ci.ub
```

```
## intrcpt              -0.5375  0.0839  -6.4033  <.0001  -0.7020  -0.3730
## skillshigh            0.0918  0.0750   1.2228  0.2214  -0.0553   0.2388
## skillslow            -0.1383  0.0704  -1.9654  0.0494  -0.2762  -0.0004
## continentUS/Canada   -0.0809  0.1755  -0.4607  0.6450  -0.4248   0.2631
##
## intrcpt             ***
## skillshigh
## skillslow            *
## continentUS/Canada
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Overview

```
    #table 1: overview over time
    table(disc$yearGroup, disc$global)
```

```
##
##        global subgroup
##   1990      6       78
##   1996      2        4
##   2001      3       66
##   2006      9      192
##   2011     23      398
```

```
    table(disc$yearGroup[disc$global=="global"], disc$region[disc$global=="global"])
```

```
##
##        Europe EU Europe non-EU North America other
##   1990         4             0             1     1
##   1996         2             0             0     0
##   2001         0             1             2     0
##   2006         9             0             0     0
##   2011        11             1             9     2
```

```
    aggregate(disc$relative.call.back.rate, by=list(Global=disc$global, Time=disc$yearGroup), mean, na.
```

```
##      Global Time    x
## 1    global 1990 1.29
## 2  subgroup 1990 1.28
## 3    global 1996 1.50
## 4  subgroup 1996 2.04
## 5    global 2001 1.47
## 6  subgroup 2001 1.54
## 7    global 2006 1.98
## 8  subgroup 2006 2.28
## 9    global 2011 1.47
## 10 subgroup 2011 1.60
```