FDI Impacts on Income Inequality

Arthur Johnson

2024-04-15

Setup and summary information

General Set-up

```
setwd("/Users/arthurjohnson/Library/CloudStorage/OneDrive-UniversityofEdinburgh/Year
↪  4/Dissertation/Meta-Analysis/Dissertation_Data_Analysis")
FDII <- read_excel("FDI_Inequality_R_input.xlsx")
FDII$control_variables_1 <- as.factor(FDII$control_variables_1)
FDII$control_variables_2 <- as.factor(FDII$control_variables_2)
FDII$control_variables_3 <- as.factor(FDII$control_variables_3)
FDII$control_variables_4 <- as.factor(FDII$control_variables_4)
FDII$control_variables_5 <- as.factor(FDII$control_variables_5)
FDII$estimation_methods <- as.factor(FDII$estimation_methods)
FDII$coefficient <- as.numeric(FDII$coefficient)
FDII$journal_rank <- as.numeric(FDII$journal_rank)
FDII$log_journal_rank <- log(FDII$journal_rank)
FDII$se <- as.numeric(FDII$se)
FDII$t_value <- as.numeric(FDII$t_value)
FDII$t_value_calculated <- as.numeric(FDII$t_value_calculated)
FDII$p_value <- as.numeric(FDII$p_value)
```

```
# This data is from the Maddison Project (2020)
Income_Data <- read.csv("Country_Income_Data.csv")
Income_Data$avg_GDPPC_pc <-
↪  log((Income_Data$Developed_Sum)/(Income_Data$Country_table_count*(Income_Data$sample_period_end-I
FDII$mean_log_GDPPC <- Income_Data$avg_GDPPC_pc
Income_Data$log10_GDPPC_pc <-
↪  log10((Income_Data$Developed_Sum)/(Income_Data$Country_table_count*(Income_Data$sample_period_end
FDII$mean_log10_GDPPC <- Income_Data$log10_GDPPC_pc

Income_Data[, 200:8648] <- apply(Income_Data[, 200:8648], 2, function(x) {
  x <- as.numeric(x)
  x[x == 0] <- NA
  return(x)
})

Income_Data$median_GDPPC <- apply(Income_Data[, 200:8648], 1, median, na.rm = TRUE)
FDII$median_log_GDPPC <- log(Income_Data$median_GDPPC)
```

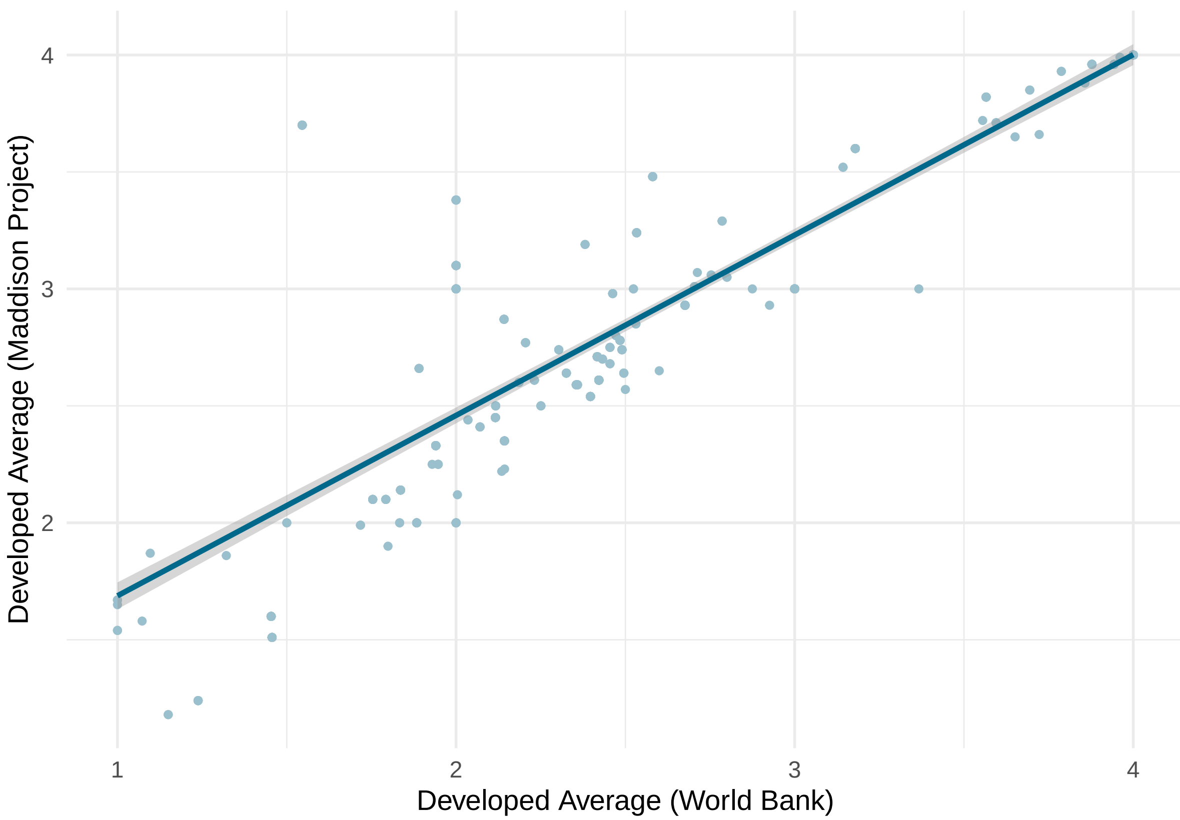
```
MP_df <- read.csv("MADDISON_PROJECT_DATA.csv")
MP_df <- MP_df[!MP_df$unique_id %in% c("16005", "16006", "45013", "45014", "45015",
↪  "45016", "45017", "45018", "45019", "45020", "45021", "45022", "45023",
↪  "45024"), ]
```

```
MP_df <- MP_df[, -(4:195)]
MP_df <- MP_df[, -(11:11434)]
```

```
FDII$MP_LI_count <- MP_df$LI_count
FDII$MP_LMI_count <- MP_df$LMI_count
FDII$MP_MHI_count <- MP_df$MHI_count
FDII$MP_HI_count <- MP_df$HI_count
FDII$MP_dev_avg <- MP_df$Developed_Average
FDII$MP_dev_sum <- MP_df$Developed_Sum
FDII$MP_incl_count <- MP_df$Developed_Count
```

```
# Plot thw two different averages
ggplot(FDII, aes(x = developed_average, y = MP_dev_avg)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  labs(x = "Developed Average (World Bank)",
       y = "Developed Average (Maddison Project)") +
  theme_minimal()
```

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



Collective country count

```
collective_country_codes <- unlist(strsplit(as.character(FDII$country_code),
↪  ";\\s*"))
collective_country_code_frequencies <-
↪  as.data.frame(table(collective_country_codes))
names(collective_country_code_frequencies) <- c('ISO_code', 'frequency')
collective_country_code_frequencies$ISO_code <-
↪  as.character(collective_country_code_frequencies$ISO_code)
```

```
collective_country_code_frequencies$country_name <-
↪  countrycode(collective_country_code_frequencies$ISO_code, origin = 'iso3c',
↪  'iso3c')
collective_country_code_frequencies$ISO_code <-
↪  as.numeric(collective_country_code_frequencies$ISO_code)
collective_country_code_frequencies$country_name <-
↪  countrycode(collective_country_code_frequencies$ISO_code, 'iso3n',
↪  'country.name')
ordered_collective_frequencies <-
↪  collective_country_code_frequencies[order(collective_country_code_frequencies$frequency,
↪  decreasing = TRUE), ]
collective_count_output <- ordered_collective_frequencies[, c("country_name",
↪  "frequency")]
sum(collective_count_output$frequency)
```

```
## [1] 18238
```

Individual country count

```
single_country_studies <- FDII[FDII$level != 3, ]
single_country_codes <-
↪  unlist(strsplit(as.character(single_country_studies$country_code), ";\\s*"))
single_country_code_frequencies <- as.data.frame(table(single_country_codes))
names(single_country_code_frequencies) <- c('ISO_code', 'frequency')
single_country_code_frequencies$ISO_code <-
↪  as.character(single_country_code_frequencies$ISO_code)
single_country_code_frequencies$country_name <-
↪  countrycode(single_country_code_frequencies$ISO_code, origin = 'iso3c', 'iso3c')
single_country_code_frequencies$ISO_code <-
↪  as.numeric(single_country_code_frequencies$ISO_code)
single_country_code_frequencies$country_name <-
↪  countrycode(single_country_code_frequencies$ISO_code, 'iso3n', 'country.name')
ordered_single_frequencies <-
↪  single_country_code_frequencies[order(single_country_code_frequencies$frequency,
↪  decreasing = TRUE), ]
single_count_output <- ordered_single_frequencies[, c('country_name', 'frequency')]
sum(single_count_output$frequency)
```

```
## [1] 252
```

Multi-country count

```
multi_country_studies <- FDII[FDII$level == 3, ]
multi_country_codes <-
↪  unlist(strsplit(as.character(multi_country_studies$country_code), ";\\s*"))
multi_country_code_frequencies <- as.data.frame(table(multi_country_codes))
names(multi_country_code_frequencies) <- c('ISO_code', 'frequency')
multi_country_code_frequencies$ISO_code <-
↪  as.character(multi_country_code_frequencies$ISO_code)
multi_country_code_frequencies$country_name <-
↪  countrycode(multi_country_code_frequencies$ISO_code, origin = 'iso3c', 'iso3c')
multi_country_code_frequencies$ISO_code <-
↪  as.numeric(multi_country_code_frequencies$ISO_code)
multi_country_code_frequencies$country_name <-
↪  countrycode(multi_country_code_frequencies$ISO_code, 'iso3n', 'country.name')
```

```r
ordered_multi_frequencies <-
↪  multi_country_code_frequencies[order(multi_country_code_frequencies$frequency,
↪  decreasing = TRUE), ]
multi_count_output <- ordered_multi_frequencies[, c('country_name', 'frequency')]
sum(multi_count_output$frequency)
```

```
## [1] 17986
```

Find counts and other relevant data to for the summary table creation

```r
# Count of regression tests used:
table(FDII$estimation_methods)
```

```
##
##                          2SLS                          3SLS
##                            21                             1
##                          ARDL               Between effects
##                             3                             2
##                           CCE                          DOLS
##                             4                            24
##                           ECM                        FM-OLS
##                             2                             6
##                           GLS                           GMM
##                             7                           104
##                            IV Johansen's cointegration test
##                            28                            16
##                          LIML                          LSDV
##                             5                             1
##                           OLS                         Panel
##                           139                           215
##                         Parks                        Probit
##                             2                             4
##                 Random effect                           SUR
##                            24                             1
##                          SURE
##                             7
```

```r
# Count of published papers:
table(FDII$if_published)
```

```
##
##   0   1
##  50 566
```

```r
# Count of FDI measure used:
# Number 1 and 9 are the Gini coefficients
table(FDII$FDI)
```

```
##
##   1   2   3   4   5   6   7   8   9  11  12  13  15  18
## 263   6   2  26  44  44  14  16 180   4   4   4   2   7
```

```r
# Count of Inequality measure used:
table(FDII$Inequality)
```

```
##
##   1   3   4   5   8   9  10  13  14  15  16  17
```

```
## 413  20  37  91   3  11   1   6  19  13   1   1
```

```r
# Count of single and multi-country studies:
table(FDII$level)
```

```
##
##   1   2   3
## 148 104 364
```

```r
# Count of regression tests used given they were published:
table(FDII$estimation_methods, FDII$if_published==1)
```

```
##
##                                FALSE TRUE
##   2SLS                             2   19
##   3SLS                             0    1
##   ARDL                             0    3
##   Between effects                  0    2
##   CCE                              0    4
##   DOLS                            10   14
##   ECM                              0    2
##   FM-OLS                           0    6
##   GLS                              0    7
##   GMM                              8   96
##   IV                               0   28
##   Johansen's cointegration test   10    6
##   LIML                             0    5
##   LSDV                             0    1
##   OLS                             10  129
##   Panel                            9  206
##   Parks                            0    2
##   Probit                           0    4
##   Random effect                    1   23
##   SUR                              0    1
##   SURE                             0    7
```

```r
# Count of Inequality measure used given they were published:
table(FDII$Inequality, FDII$if_published==1)
```

```
##
##      FALSE TRUE
##   1     42  371
##   3      0   20
##   4      8   29
##   5      0   91
##   8      0    3
##   9      0   11
##   10     0    1
##   13     0    6
##   14     0   19
##   15     0   13
##   16     0    1
##   17     0    1
```

```r
# Count of FDI measure used given they were published
table(FDII$FDI, FDII$if_published==1)
```

```
##
```

```
##      FALSE TRUE
##   1     38  225
##   2      0    6
##   3      0    2
##   4      0   26
##   5      0   44
##   6      0   44
##   7      0   14
##   8      0   16
##   9      4  176
##   11     4    0
##   12     4    0
##   13     0    4
##   15     0    2
##   18     0    7
```

```r
#Replace all NAs with 'BLANK' in the control variables
FDII$control_variables_1 <- as.character(FDII$control_variables_1)
FDII$control_variables_1[is.na(FDII$control_variables_1)] <- "NO CONTROL"
FDII$control_variables_2 <- as.character(FDII$control_variables_2)
FDII$control_variables_2[is.na(FDII$control_variables_2)] <- "NO CONTROL"
FDII$control_variables_3 <- as.character(FDII$control_variables_3)
FDII$control_variables_3[is.na(FDII$control_variables_3)] <- "NO CONTROL"
FDII$control_variables_4 <- as.character(FDII$control_variables_4)
FDII$control_variables_4[is.na(FDII$control_variables_4)] <- "NO CONTROL"
FDII$control_variables_5 <- as.character(FDII$control_variables_5)
FDII$control_variables_5[is.na(FDII$control_variables_5)] <- "NO CONTROL"
```

```r
# Info on control variable counts
selected_columns <- FDII[, c("control_variables_1", "control_variables_2",
↪  "control_variables_3", "control_variables_4", "control_variables_5")]
unlist <- tolower(unlist(selected_columns, use.names = FALSE))
count_table <- table(unlist)
count_df <- as.data.frame(count_table)
names(count_df) <- c("Variable", "Included")
count_df$notincl <- 616 - count_df$Included
count_df[order(-count_df$Included), ]
```

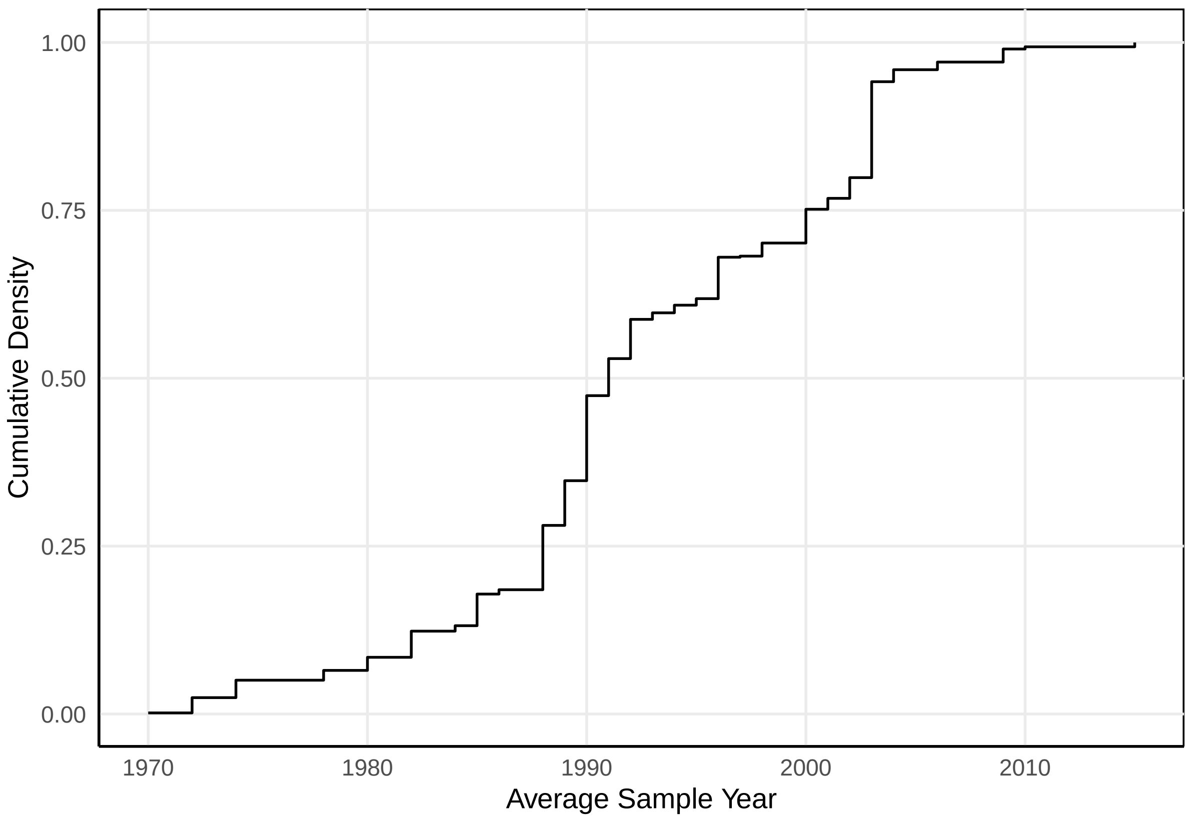
```
##                               Variable Included notincl
## 47                          no control      813    -197
## 16 education_secondary_school_enrollment      292     324
## 28                               gdppc      263     353
## 71                               trade      176     440
## 52                                 pop      167     449
## 39                           inflation      162     454
## 25                                 gdp      118     498
## 72                        unemployment      115     501
## 32                             gov_exp       98     518
## 29                             gdppc^2       88     528
## 48                            openness       74     542
## 3                          agriculture       70     546
## 27                               gdpgr       65     551
## 57                      private_credit       45     571
## 35                       human_capital       40     576
## 7                              capital       39     577
## 54                               popgr       34     582
## 36                              import       32     584
```

```
## 22                            exports   28    588
## 75                        value-added   26    590
## 66                             tariff   23    593
## 19                      exchange_rate   20    596
## 67                               tech   20    596
## 38                industry_employment   16    600
## 58                                r&d   16    600
## 73                              union   15    601
## 37                           industry   13    603
## 30                               gini   12    604
## 11                                cpi   11    605
## 15 education_middle_school_enrollment   11    605
## 61                      sector_dualism   11    605
## 8                           communist   10    606
## 13                          democracy   10    606
## 53                        pop_over_65   10    606
## 55                    portfolio_inflow    9    607
## 9                             company    8    608
## 26                              gdp^2    8    608
## 33                         gov_quality    8    608
## 34          herfindahl-hirschman_index    8    608
## 44     manufacturing_real_gross_output    8    608
## 45                        market_share    8    608
## 24               financial_development    7    609
## 6                        black_market    6    610
## 74                        urbanisation    6    610
## 68                         tech_import    4    612
## 5                                asia    3    613
## 10                       country_size    3    613
## 17                              elect    3    613
## 18                           exchange    3    613
## 23                        fdi*exports    3    613
## 41                                lac    3    613
## 63                      shadow_economy    3    613
## 64                     socialist_state    3    613
## 69                            tourism    3    613
## 14                    economic_freedom    2    614
## 21                   export_incentives    2    614
## 40                  labor_productivity    2    614
## 42                              landpc    2    614
## 43                                  m2    2    614
## 46                           migration    2    614
## 62                             service    2    614
## 65                           soe_share    2    614
## 70                           tourism^2    2    614
## 1                                   _    1    615
## 2                 agricultural_exports    1    615
## 4                                  aid    1    615
## 12                                debt    1    615
## 20                       export_growth    1    615
## 31                                 gni    1    615
## 49                                 pcm    1    615
## 50                               phone    1    615
## 51               political_integration    1    615
## 56                             poverty    1    615
## 59               relative_productivity    1    615
```

```
## 60                                      remit         1      615
```

Average year visualisation

```r
FDII$avg_sample_year <- round((FDII$sample_period_start + (FDII$sample_period_end -
↪  FDII$sample_period_start)/2),0)

ggplot(FDII, aes(x = avg_sample_year), width = 8, height = 6) +
stat_ecdf(geom = "step", pad = FALSE) +
labs(x = "Average Sample Year",
  y = "Cumulative Density") +
  theme_minimal() +
  theme(plot.background = element_rect(fill = "transparent", color = NA),
  panel.background = element_rect(fill = "white"),
  panel.grid.minor = element_blank(),
  panel.border = element_blank(),
  axis.line = element_line(color = "black"))
```

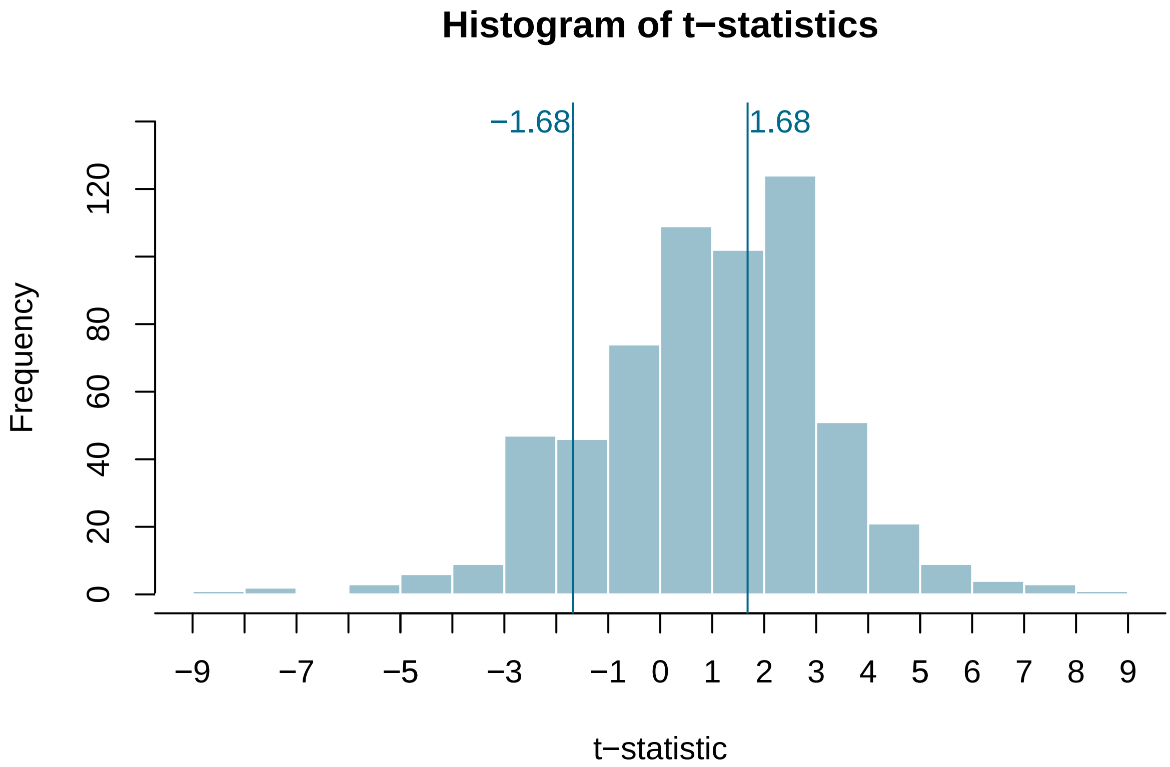



```r
ggsave("average_sample_year_plot.png", bg = "transparent", width = 8, height = 6)
```

t-statistic density by strength

```r
hist(FDII$t_value_calculated,
  breaks = 30,
  xlim = c(-9, 9),
  ylim = c(0, 140),
  xlab = "t-statistic",
  col = "lightblue3",
  border = "white",
```

```
  main = "Histogram of t-statistics")
axis(1, at = -10:10)
abline(v = -1.68, col = "deepskyblue4", lwd = 1, lty = 1)
text(x = -2.5, y = 140, '-1.68', col = "deepskyblue4")
abline(v = 1.68, col = "deepskyblue4", lwd = 1, lty = 1)
text(x = 2.3, y = 140, '1.68', col = "deepskyblue4")
```



Further setup for MRA and initial tests

Set up Partial Correlation Coefficient

```
FDII$PCC <-
↪   (FDII$t_value_calculated)/sqrt((FDII$t_value_calculated)^2+FDII$degree_freedom)
FDII$PCC_se <- sqrt((1-(FDII$PCC)^2)/FDII$degree_freedom)
mean(FDII$PCC)
```

```
## [1] 0.08913153
```

```
mean(FDII$PCC_se)
```

```
## [1] 0.1014291
```

This is from Stanley and Doucouliagos (2012) - ENSURE citation!!

Calculate averages within and out of OECD countries

```
# Number and proportion of studies that were conducted in only OECD countries (both
↪   individual and multi-country studies)
print(paste("The number of studies that focused solely on OECD countries is",
↪   length(FDII$if_OECD[FDII$if_OECD==1]), "-- This includes individual-level
↪   studies"))
```

```
## [1] "The number of studies that focused solely on OECD countries is 188 -- This includes individua
```

```r
print(paste("The number of studies that focused solely on non-OECD countries is",
↪  length(FDII$if_OECD[FDII$if_OECD==0]), "-- This includes individual-level
↪  studies"))
```

```
## [1] "The number of studies that focused solely on non-OECD countries is 428 -- This includes indiv
```

```r
print(paste("The proportion of studies that focused solely on OECD countries is",
↪  round(100*length(FDII$if_OECD[FDII$if_OECD==1])/(length(FDII$if_OECD[FDII$if_OECD==1])
↪  + length(FDII$if_OECD[FDII$if_OECD==0])), 2),"%"))
```

```
## [1] "The proportion of studies that focused solely on OECD countries is 30.52 %"
```

```r
OECD_multi <- length(FDII$if_OECD[FDII$if_OECD == 1 & FDII$level == 3])
print(paste("The number of multi-country studies that focused solely on OECD
↪  countries is", OECD_multi))
```

```
## [1] "The number of multi-country studies that focused solely on OECD countries is 22"
```

```r
nonOECD_multi <- length(FDII$if_OECD[FDII$if_OECD == 0 & FDII$level == 3])
print(paste("The number of multi-country studies that focused solely on non-OECD
↪  countries is", nonOECD_multi))
```

```
## [1] "The number of multi-country studies that focused solely on non-OECD countries is 342"
```

```r
print(paste("The proportion of multi-country studies that focused solely on OECD
↪  countries is", round(100*OECD_multi/(OECD_multi + nonOECD_multi), 2),"%"))
```

```
## [1] "The proportion of multi-country studies that focused solely on OECD countries is 6.04 %"
```

```r
OECD_single <- length(FDII$if_OECD[FDII$if_OECD == 1 & FDII$level != 3])
print(paste("The number of single-country studies that focused solely on OECD
↪  countries is", OECD_single))
```

```
## [1] "The number of single-country studies that focused solely on OECD countries is 166"
```

```r
nonOECD_single <- length(FDII$if_OECD[FDII$if_OECD == 0 & FDII$level != 3])
print(paste("The number of single-country studies that focused solely on non-OECD
↪  countries is", nonOECD_single))
```

```
## [1] "The number of single-country studies that focused solely on non-OECD countries is 86"
```

```r
print(paste("The proportion of single-country studies that focused solely on OECD
↪  countries is", round(100*OECD_single/(OECD_single + nonOECD_single), 2),"%"))
```

```
## [1] "The proportion of single-country studies that focused solely on OECD countries is 65.87 %"
```

```r
print(paste("The mean of the PCC of non-OECD focussed studies (single- and
↪  multi-country studies)", round(mean(FDII$PCC[FDII$if_OECD==0]),3)))
```

```
## [1] "The mean of the PCC of non-OECD focussed studies (single- and multi-country studies) 0.138"
```

```r
print(paste("The standard error of the PCC of non-OECD focussed studies (single- and
↪  multi-country studies)", round(mean(FDII$PCC_se[FDII$if_OECD==0]),3)))
```

```
## [1] "The standard error of the PCC of non-OECD focussed studies (single- and multi-country studies
```

```r
print(paste("The mean of the PCC of OECD-focussed studies (single- and multi-country
↪  studies)", round(mean(FDII$PCC[FDII$if_OECD==1]),3)))
```

```
## [1] "The mean of the PCC of OECD-focussed studies (single- and multi-country studies) -0.022"
```

```
print(paste("The standard error of the PCC of OECD-focussed studies (single- and
↪  multi-country studies)", round(mean(FDII$PCC_se[FDII$if_OECD==1]),3)))
```

```
## [1] "The standard error of the PCC of OECD-focussed studies (single- and multi-country studies) 0.
```

```
print(paste("The mean of the PCC of OECD-focussed studies (single-country studies)",
↪  round(mean(FDII$PCC[FDII$if_OECD == 1 & FDII$level != 3]),3)))
```

```
## [1] "The mean of the PCC of OECD-focussed studies (single-country studies) -0.013"
```

```
print(paste("The standard error of the PCC of OECD-focussed studies (single-country
↪  studies)", round(mean(FDII$PCC_se[FDII$if_OECD == 1 & FDII$level != 3]),3)))
```

```
## [1] "The standard error of the PCC of OECD-focussed studies (single-country studies) 0.068"
```

```
print(paste("The mean of the PCC of non-OECD focussed studies (single-country
↪  studies)", round(mean(FDII$PCC[FDII$if_OECD == 0 & FDII$level != 3]),3)))
```

```
## [1] "The mean of the PCC of non-OECD focussed studies (single-country studies) 0.066"
```

```
print(paste("The standard error of the PCC of non-OECD focussed studies
↪  (single-country studies)", round(mean(FDII$PCC_se[FDII$if_OECD == 0 & FDII$level
↪  != 3]),3)))
```

```
## [1] "The standard error of the PCC of non-OECD focussed studies (single-country studies) 0.181"
```

```
print(paste("The mean of the PCC of OECD-focussed studies (multi-country studies)",
↪  round(mean(FDII$PCC[FDII$if_OECD == 1 & FDII$level == 3]),3)))
```

```
## [1] "The mean of the PCC of OECD-focussed studies (multi-country studies) -0.089"
```

```
print(paste("The standard error of the PCC of OECD-focussed studies (multi-country
↪  studies)", round(mean(FDII$PCC_se[FDII$if_OECD == 1 & FDII$level == 3]),3)))
```

```
## [1] "The standard error of the PCC of OECD-focussed studies (multi-country studies) 0.076"
```

```
print(paste("The mean of the PCC of non-OECD focussed studies (multi-country
↪  studies)", round(mean(FDII$PCC[FDII$if_OECD == 0 & FDII$level == 3]),3)))
```

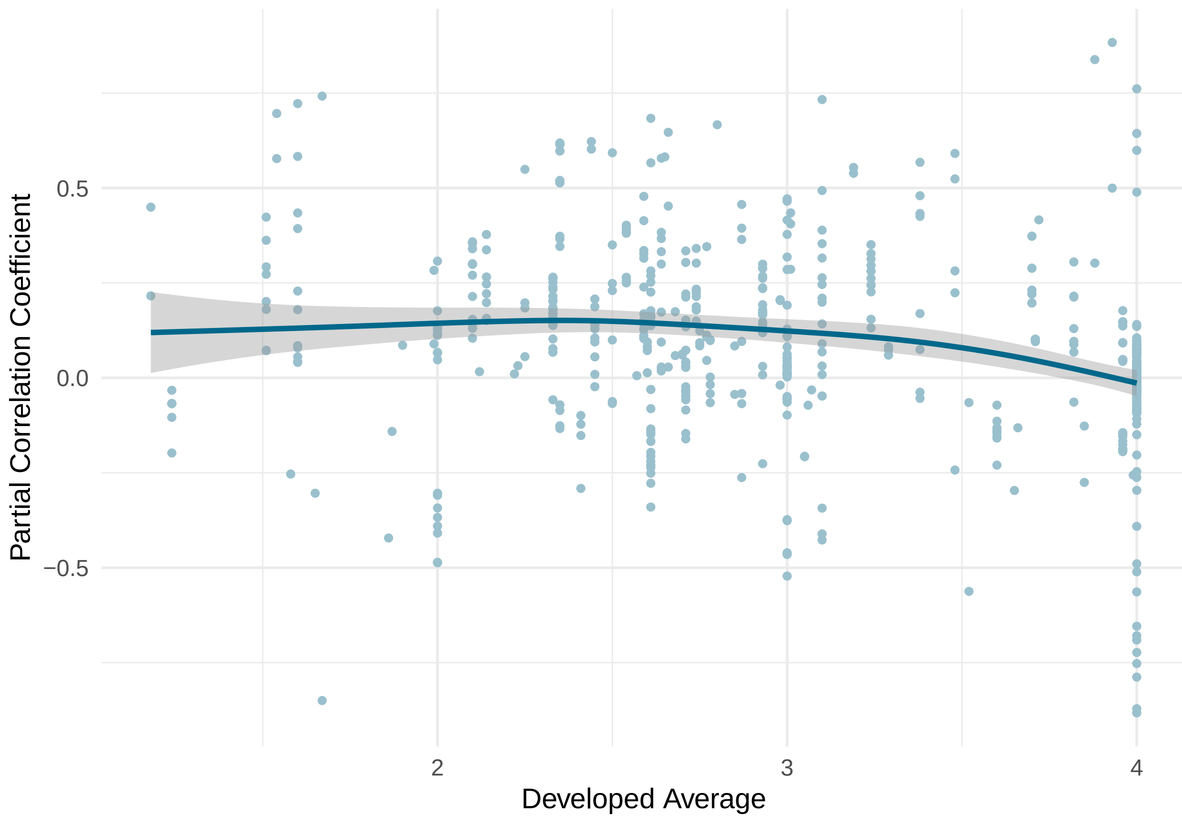
```
## [1] "The mean of the PCC of non-OECD focussed studies (multi-country studies) 0.156"
```

```
print(paste("The standard error of the PCC of non-OECD focussed studies
↪  (multi-country studies)", round(mean(FDII$PCC_se[FDII$if_OECD == 0 & FDII$level
↪  == 3]),3)))
```

```
## [1] "The standard error of the PCC of non-OECD focussed studies (multi-country studies) 0.1"
```

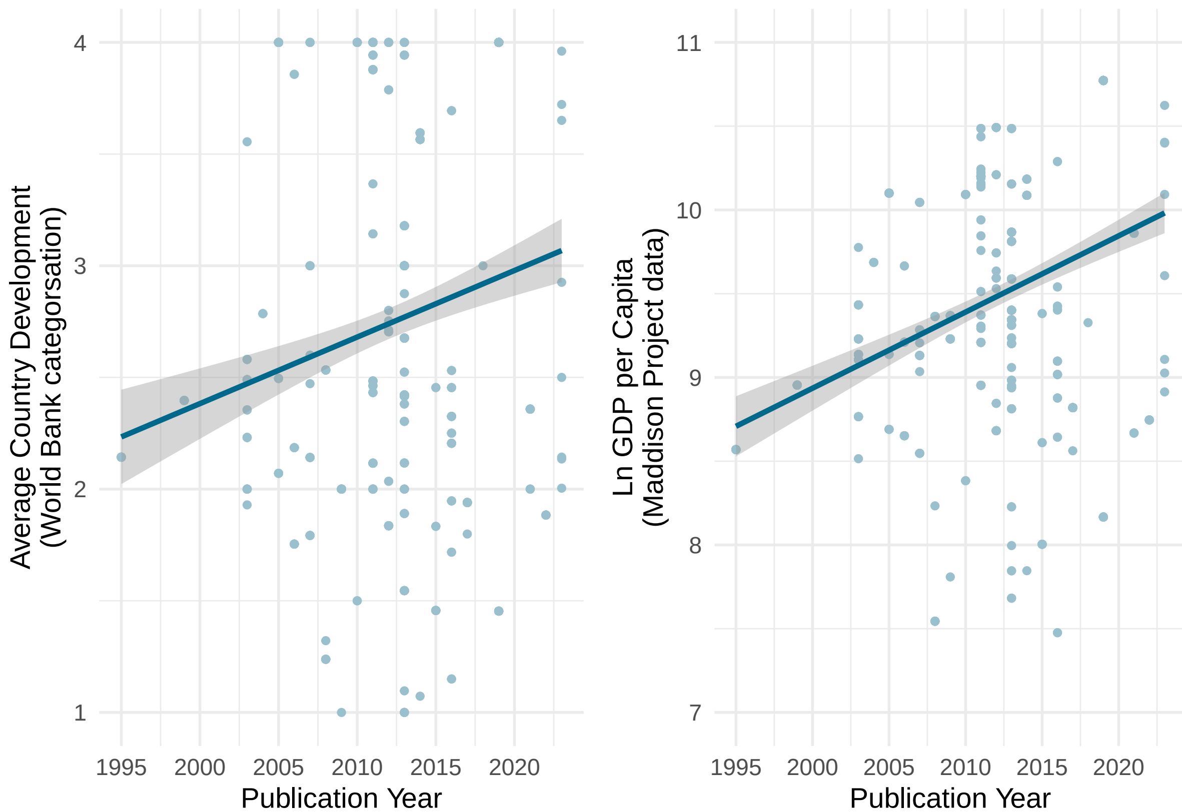
Initial plots

```
ggplot(data = FDII, aes(x = MP_dev_avg, y = PCC)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "gam", formula = y ~ s(x, bs = "cs"), color = "deepskyblue4")
  ↪  +
  labs(x = "Developed Average",
       y = "Partial Correlation Coefficient") +
  theme_minimal()
```



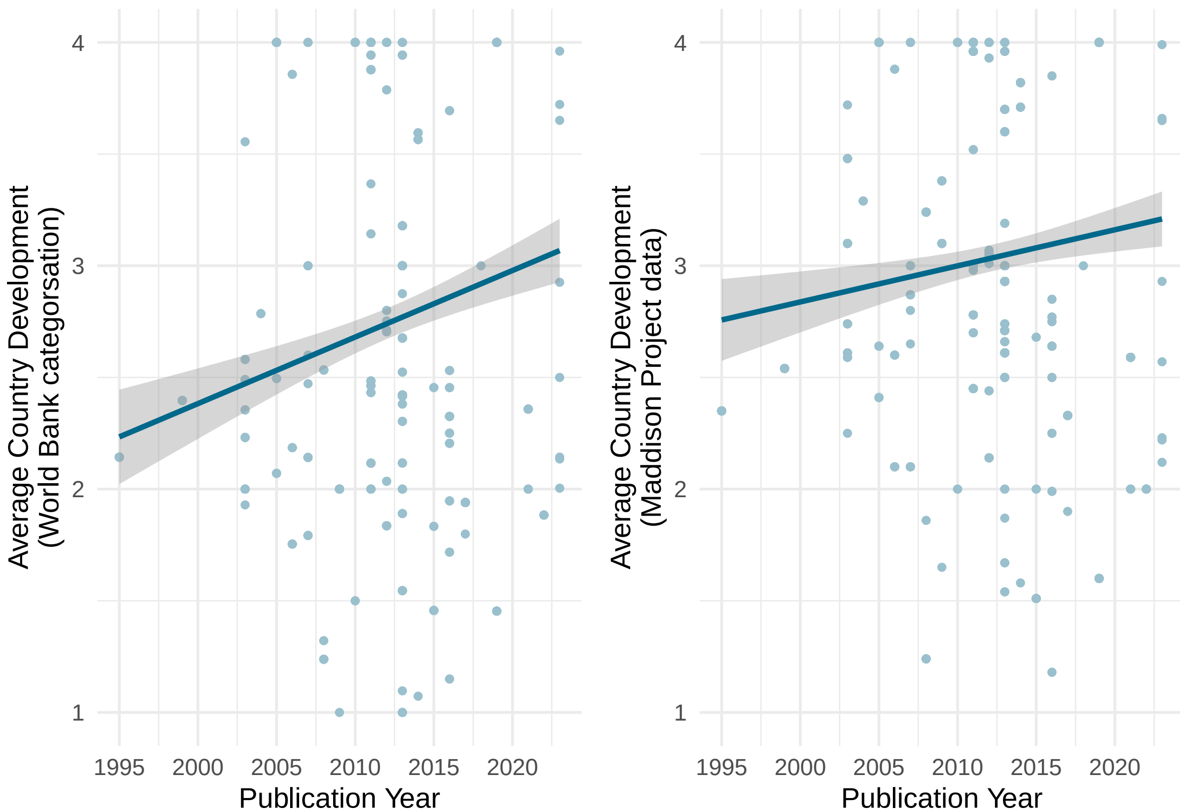

```r
WB_pub <- ggplot(data = FDII, aes(x = publication_time, y = developed_average)) +
  geom_point(color = "lightblue3", size = 1) +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  labs(x = "Publication Year",
       y = "Average Country Development\n(World Bank categorsation)") +
  theme_minimal()

MP_pub <- ggplot(data = FDII, aes(x = publication_time, y =
↪  Income_Data$avg_GDPPC_pc)) +
  geom_point(color = "lightblue3", size = 1) +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  ylim(7, 11) +
  labs(x = "Publication Year",
       y = "Ln GDP per Capita\n(Maddison Project data)")+
  theme_minimal()

grid.arrange(WB_pub, MP_pub, ncol = 2)
```

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



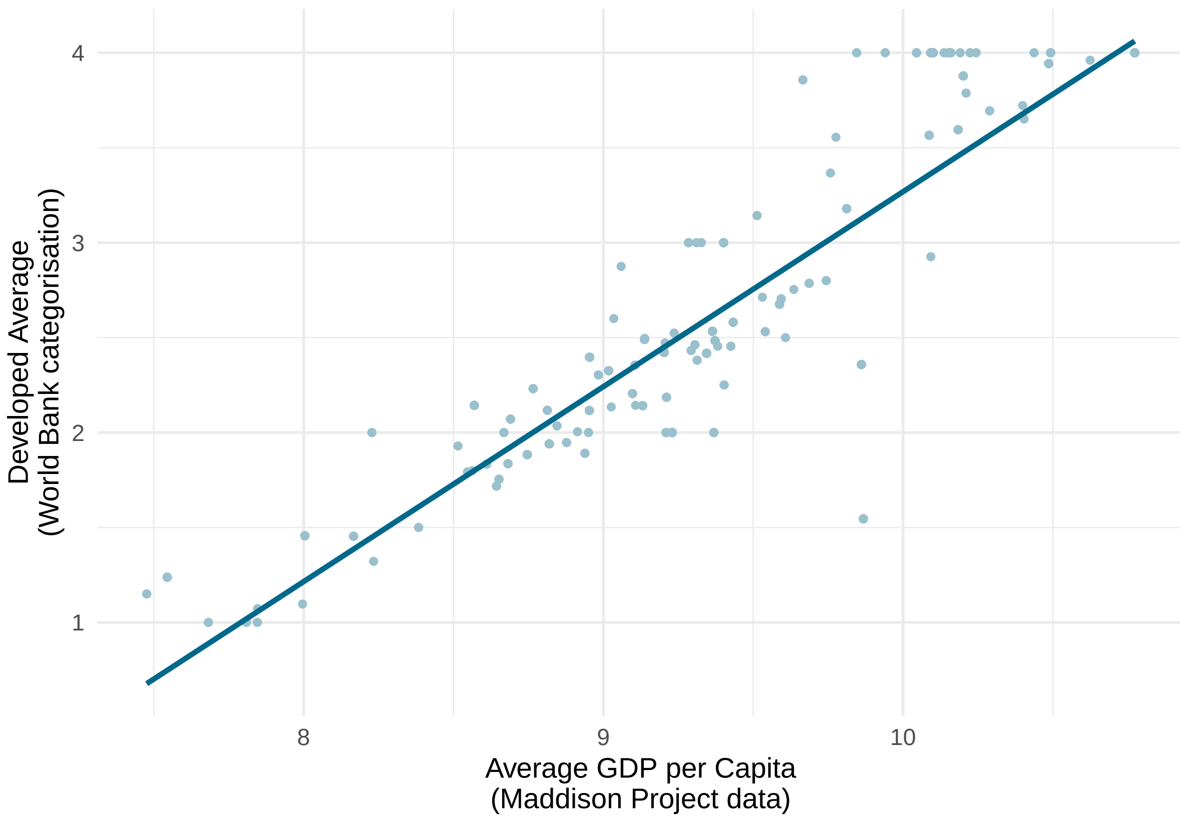

```
WB_pub <- ggplot(data = FDII, aes(x = publication_time, y = developed_average)) +
  geom_point(color = "lightblue3", size = 1) +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  labs(x = "Publication Year",
       y = "Average Country Development\n(World Bank categorsation)") +
  theme_minimal()

MP_pub <- ggplot(data = FDII, aes(x = publication_time, y = MP_dev_avg)) +
  geom_point(color = "lightblue3", size = 1) +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  ylim(1, 4) +
  labs(x = "Publication Year",
       y = "Average Country Development\n(Maddison Project data)")+
  theme_minimal()

grid.arrange(WB_pub, MP_pub, ncol = 2)
```

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



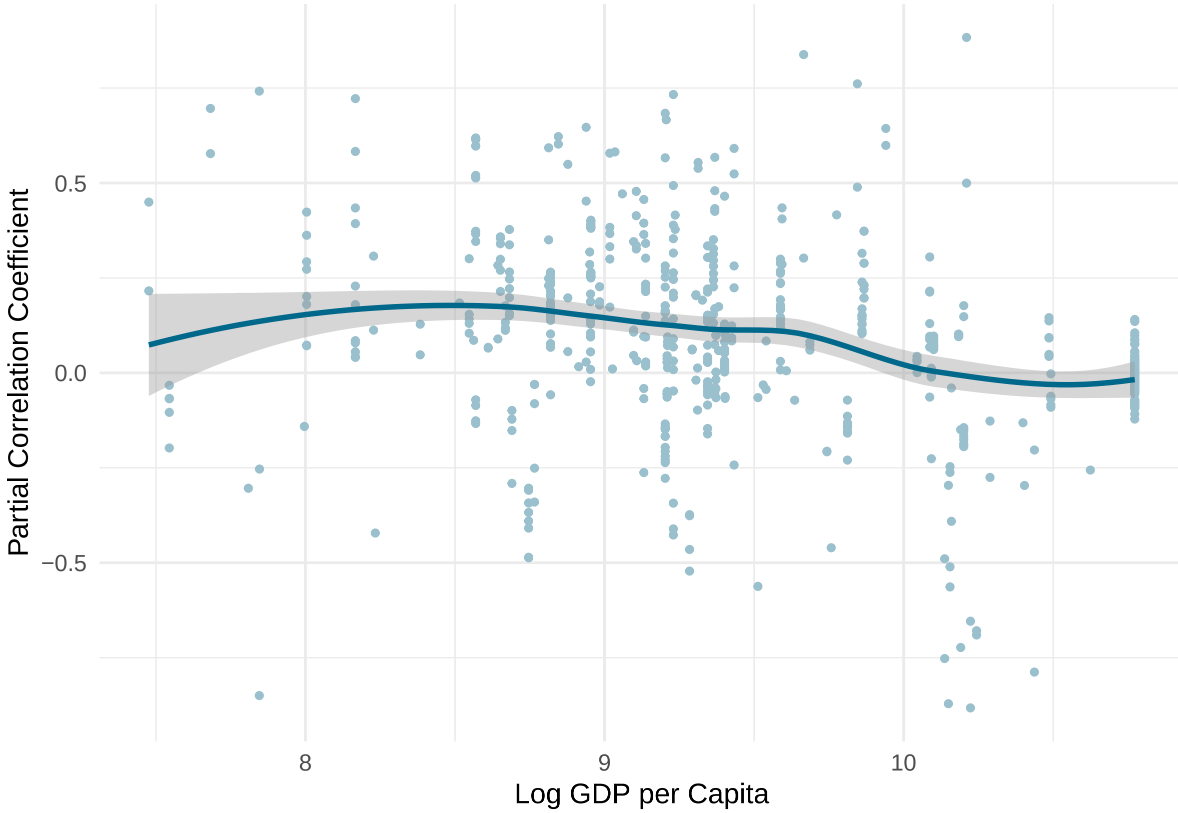

```
ggplot(data = data.frame(Income_Data$avg_GDPPC_pc, FDII$developed_average)) +
  geom_point(aes(x = Income_Data$avg_GDPPC_pc, y = FDII$developed_average), color =
  ↪   "lightblue3", size = 1) +
  geom_smooth(aes(x = Income_Data$avg_GDPPC_pc, y = FDII$developed_average), method
  ↪   = "lm", color = "deepskyblue4", se = FALSE) +
  labs(x = "Average GDP per Capita\n(Maddison Project data)",
       y = "Developed Average\n(World Bank categorisation)") +
  theme_minimal()
```

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



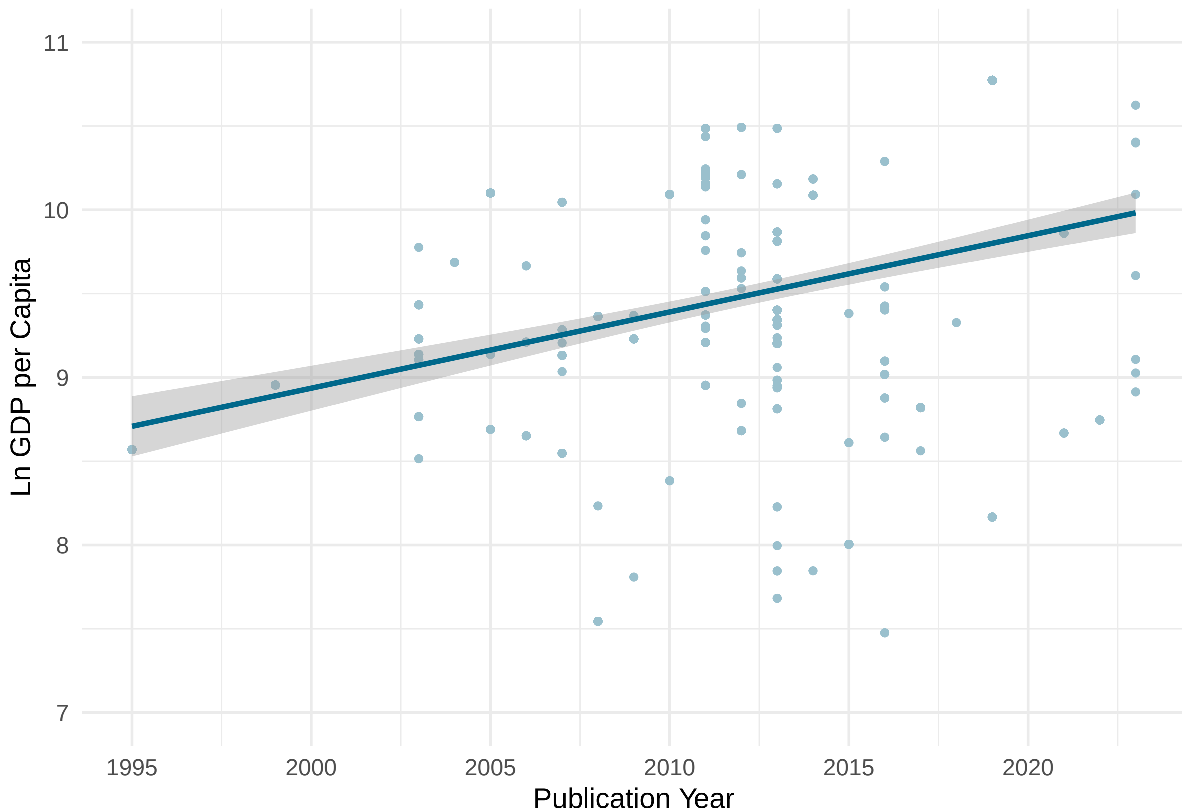

```
ggplot(data = Income_Data, aes(x = avg_GDPPC_pc, y = FDII$PCC)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "loess", color = "deepskyblue4") +
  labs(x = "Log GDP per Capita",
       y = "Partial Correlation Coefficient") +
  theme_minimal()
```

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




```r
ggplot(data = FDII, aes(x = publication_time, y = Income_Data$avg_GDPPC_pc)) +
  geom_point(color = "lightblue3", size = 1) +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  ylim(7, 11) +
  labs(x = "Publication Year",
       y = "Ln GDP per Capita")+
  theme_minimal()
```

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

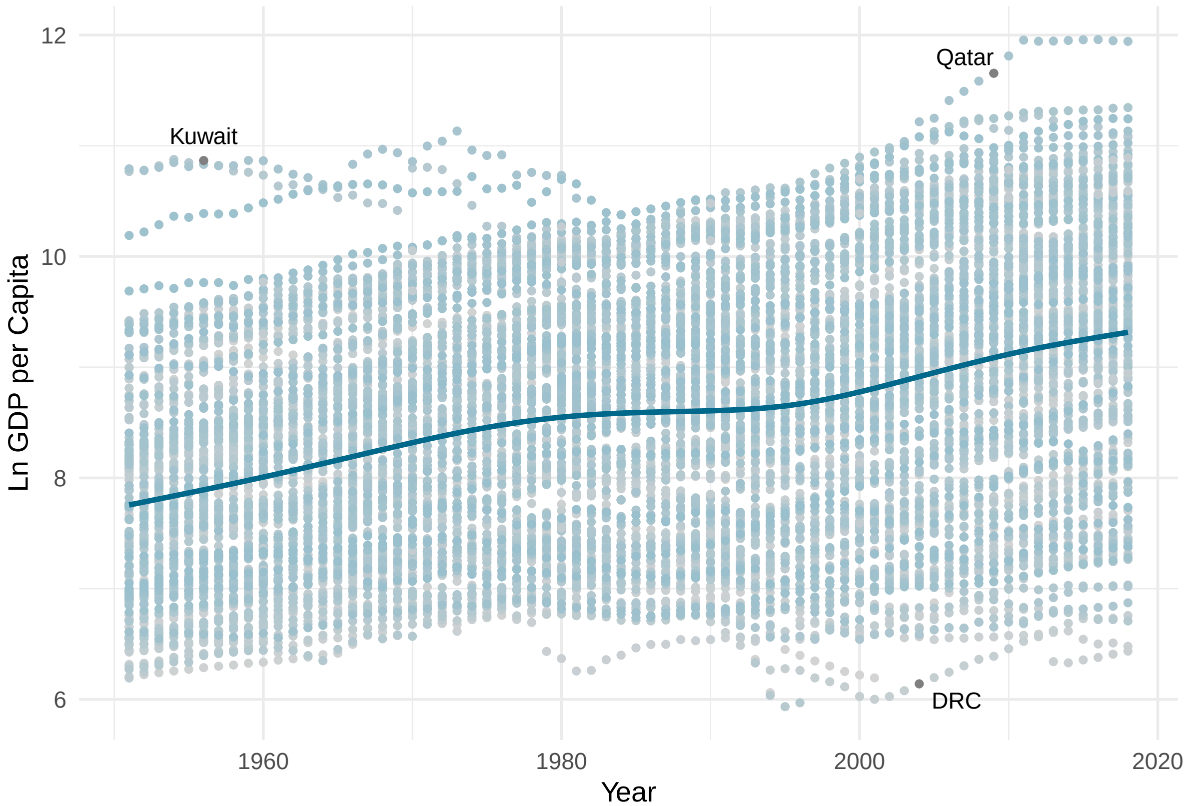



```r
maddison_df <- read.csv("Maddison_Data.csv")
```

```r
maddison_df$LogGDPpc <- log(maddison_df$GDP.per.capita)
maddison_df_recent <- maddison_df[maddison_df$Year > 1950,]

unique_codes <- unique(maddison_df_recent$Code)
grey_blue_colors <- colorRampPalette(c("lightgrey",
↪  "lightblue3"))(length(unique_codes))
color_mapping <- setNames(grey_blue_colors, unique_codes)

ggplot(data = maddison_df_recent, aes(x = Year, y = LogGDPpc, color = Code)) +
  geom_point(size = 1) +
  geom_smooth(method = "gam", se = FALSE, color = "deepskyblue4") +
  labs(x = "Year",
       y = "Ln GDP per Capita") +
  scale_color_manual(values = color_mapping, breaks = NULL) +
  theme_minimal() +
  geom_text(data = subset(maddison_df_recent, Code == "QAT" & Year == 2009),
            aes(label = "Qatar"),
            vjust = -0.5,
            hjust = 1,
            color = "black",
            size = 3) +
  geom_text(data = subset(maddison_df_recent, Code == "KWT" & Year == 1956),
            aes(label = "Kuwait"),
            vjust = -1,
            color = "black",
            size = 3) +
  geom_text(data = subset(maddison_df_recent, Code == "COD" & Year == 2004),
            aes(label = "DRC"),
```

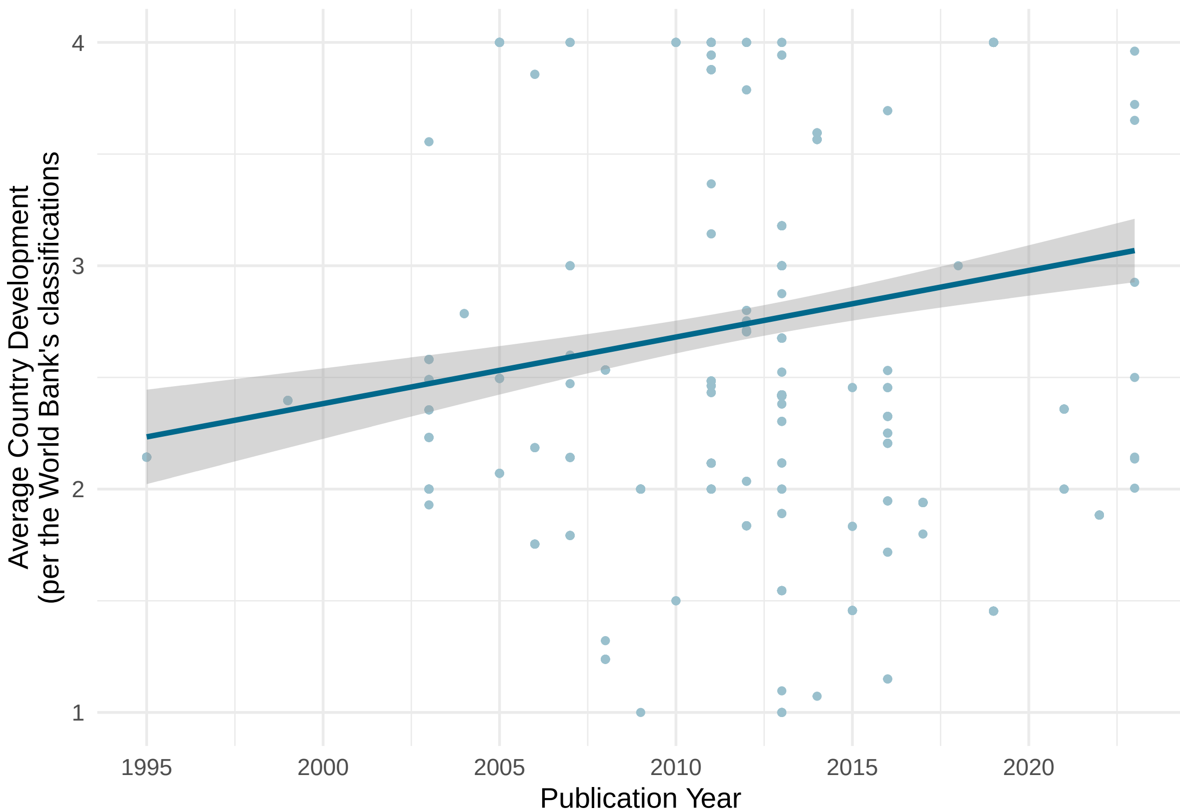
```
            vjust = 1.5,
            hjust = -0.25,
            color = "black",
            size = 3) +
  geom_point(data = subset(maddison_df_recent, (Code == "QAT" & Year == 2009) |
                               (Code == "KWT" & Year == 1956) |
                               (Code == "COD" & Year == 2004)),
             aes(color = "black"),
             size = 1)
```

```
## `geom_smooth()` using formula = 'y ~ s(x, bs = "cs")'
```



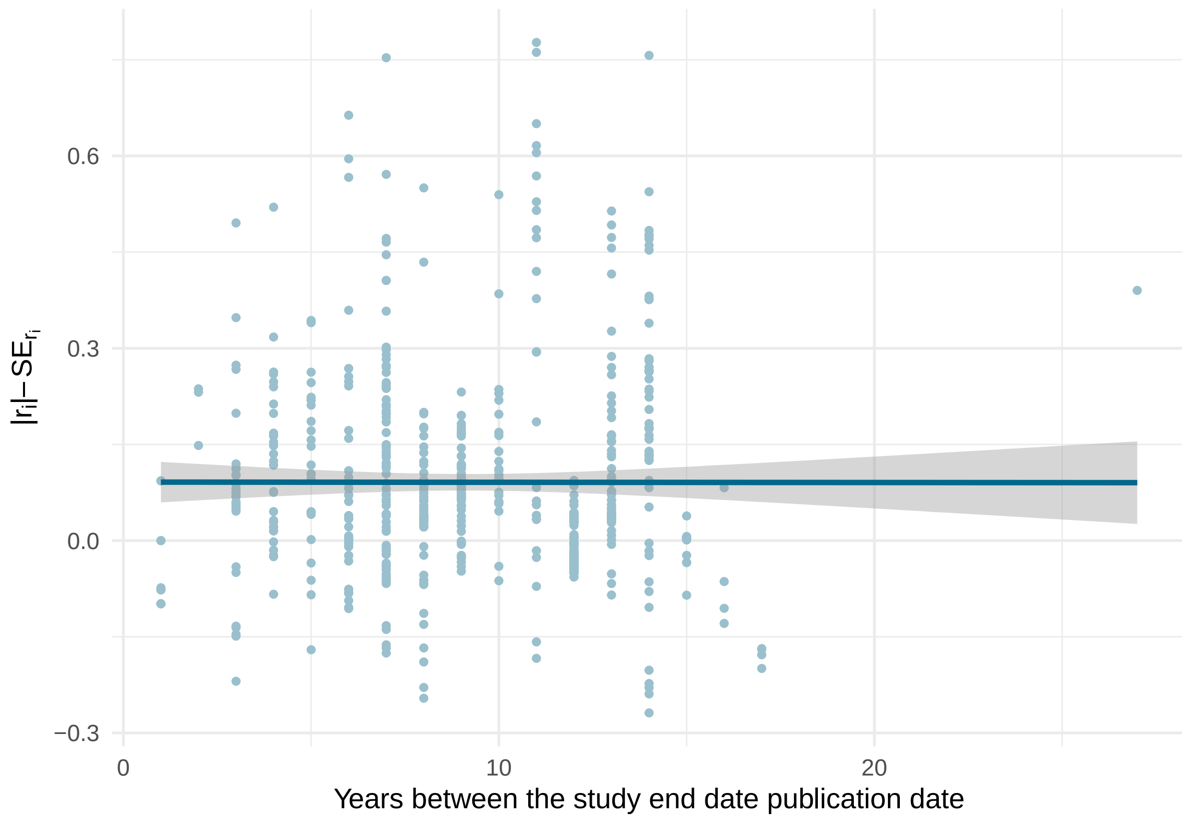

```
ggplot(data = FDII, aes(x = publication_time, y = developed_average)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  labs(x = "Publication Year",
       y = "Average Country Development\n(per the World Bank's classifications") +
  theme_minimal()
```

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




```r
ggplot(data = FDII, aes(x = publication_time-sample_period_end, y =
↪  abs(PCC)-PCC_se)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "lm", color = "deepskyblue4") +
  labs(x = "Years between the study end date publication date",
       y = expression(paste("|", r[i],"|", - SE[r[i]]))) +
  theme_minimal()
```

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



Setup for meta regression analysis by drawing out control variables into binary data

```
# Step 1: Pivot longer
FDII_long <- FDII %>%
  pivot_longer(cols = starts_with("control_variables_"),
               names_to = "Control_Variable_Type",
               values_to = "Control_Variable_Name") %>%
  filter(!is.na(Control_Variable_Name) & Control_Variable_Name != "")

# Step 2: Pivot wider and create indicator variables
FDII_wide <- FDII_long %>%
  mutate(Presence = 1) %>%
  pivot_wider(names_from = Control_Variable_Name, values_from = Presence,
  ↪   values_fill = list(Presence = 0))

# Step 3: Group by and summarise only numeric columns
FDII_merged <- FDII_wide %>%
  group_by(citation, coefficient, t_value_calculated, study_country, journal_name,
  ↪   estimation_methods, journal_rank) %>%
  summarise(across(where(is.numeric), ~if(all(is.na(.))) 0 else max(., na.rm =
  ↪   TRUE)), .groups = 'drop')

# Left join to put the control variables back into FDII
FDII <- left_join(FDII, FDII_merged, by = c("citation", "coefficient",
↪   "t_value_calculated", "study_country", "journal_name", "estimation_methods",
↪   "journal_rank"))

# Remove duplicate columns
```

```
FDII <- FDII %>% dplyr::select(-matches("\\.y$"))
FDII <- FDII %>% dplyr::rename_with(~str_replace(., "\\.x$", ""), .cols =
↪  ends_with(".x"))

FDII$GDPpc_2 <- FDII$`GDPpc^2`
```

Filters for meta analyses

```
# Measures of FDI and Inequality
FDII$stock <- ifelse(FDII$FDI %in% c(1, 3, 9), 1, 0)
FDII$GINI_control <- ifelse(FDII$Inequality == 1, 1, 0)
FDII$W_I_Disp <- ifelse(FDII$Inequality %in% c(4, 7, 8, 11, 12, 13, 14, 16, 17, 18),
↪  1, 0)
FDII$I_Share <- ifelse(FDII$Inequality %in% c(2, 5, 6), 1, 0)
FDII$Other_I_measure <- ifelse(FDII$Inequality %in% c(3, 9, 10, 15), 1, 0)

# Measures of controlling for endogeneity
FDII$est_method <- ifelse(FDII$estimation_methods %in% c('GMM', 'IV', '2SLS'), 1, 0)
# Single Country
FDII$single_country <- ifelse(FDII$level %in% c(1, 2), 1, 0)
# Control(s)
FDII$GDP_control <- ifelse(FDII$control_variables_1 %in% c('GDPpc', 'GDP',
↪  'GDPpc^2','GDP^2') | FDII$control_variables_2 %in% c('GDPpc', 'GDP',
↪  'GDPpc^2','GDP^2') | FDII$control_variables_3 %in% c('GDPpc', 'GDP',
↪  'GDPpc^2','GDP^2') | FDII$control_variables_4 %in% c('GDPpc', 'GDP',
↪  'GDPpc^2','GDP^2') | FDII$control_variables_5 %in% c('GDPpc', 'GDP',
↪  'GDPpc^2','GDP^2'), 1, 0)
FDII$GDP_control <- ifelse(!is.na(FDII$GDP_control), FDII$GDP_control, 0)

FDII$education_control <- ifelse(FDII$control_variables_1 %in%
↪  c('education_secondary_school_enrollment', 'education_middle_school_enrollment')
↪  | FDII$control_variables_2 %in% c('education_secondary_school_enrollment',
↪  'education_middle_school_enrollment') | FDII$control_variables_3 %in%
↪  c('education_secondary_school_enrollment', 'education_middle_school_enrollment')
↪  | FDII$control_variables_4 %in% c('education_secondary_school_enrollment',
↪  'education_middle_school_enrollment') | FDII$control_variables_5 %in%
↪  c('education_secondary_school_enrollment',
↪  'education_middle_school_enrollment'), 1, 0)
FDII$education_control <- ifelse(!is.na(FDII$education_control),
↪  FDII$education_control, 0)

FDII$trade_control <- ifelse(FDII$control_variables_1 == 'trade' |
↪  FDII$control_variables_2 == 'trade' | FDII$control_variables_3 == 'trade' |
↪  FDII$control_variables_4 == 'trade' | FDII$control_variables_5 == 'trade', 1, 0)
FDII$trade_control <- ifelse(!is.na(FDII$trade_control), FDII$trade_control, 0)

FDII$inflation_control <- ifelse(FDII$control_variables_1 == 'inflation' |
↪  FDII$control_variables_2 == 'inflation' | FDII$control_variables_3 ==
↪  'inflation' | FDII$control_variables_4 == 'inflation' | FDII$control_variables_5
↪  == 'inflation', 1, 0)
FDII$inflation_control <- ifelse(!is.na(FDII$inflation_control),
↪  FDII$inflation_control, 0)

FDII$population_control <- ifelse(FDII$control_variables_1 %in% c('pop') |
↪  FDII$control_variables_2 %in% c('pop') | FDII$control_variables_3 %in% c('pop')
↪  | FDII$control_variables_4 %in% c('pop') | FDII$control_variables_5 %in%
↪  c('pop'), 1, 0)
```

```
FDII$population_control <- ifelse(!is.na(FDII$population_control),
↪  FDII$population_control, 0)

FDII$gov_inequality_effort_control <- ifelse(FDII$control_variables_1 %in%
↪  c('gov_exp', 'qov_quality') | FDII$control_variables_2 %in% c('gov_exp',
↪  'qov_quality') | FDII$control_variables_3 %in% c('gov_exp', 'qov_quality') |
↪  FDII$control_variables_4 %in% c('gov_exp', 'qov_quality') |
↪  FDII$control_variables_5 %in% c('gov_exp', 'qov_quality'), 1, 0)
FDII$gov_inequality_effort_control <-
↪  ifelse(!is.na(FDII$gov_inequality_effort_control),
↪  FDII$gov_inequality_effort_control, 0)

FDII$unemployment_control <- ifelse(FDII$control_variables_1 == 'unemployment' |
↪  FDII$control_variables_2 == 'unemployment' | FDII$control_variables_3 ==
↪  'unemployment' | FDII$control_variables_4 == 'unemployment' |
↪  FDII$control_variables_5 == 'unemployment', 1, 0)
FDII$unemployment_control <- ifelse(!is.na(FDII$unemployment_control),
↪  FDII$unemployment_control, 0)

save(FDII, file = "FDI_Inequality_R.rda")
```

Split data into income buckets

```
FDII_higher <- FDII[FDII$developed_average > 3,]
FDII_middle <- FDII[FDII$developed_average > 2 & FDII$developed_average <= 3,]
FDII_lower <- FDII[FDII$developed_average <= 2,]
```

Means for the different income levels

```
mean(FDII_higher$PCC)
```

```
## [1] -0.02479306
```

```
mean(FDII_higher$PCC_se)
```

```
## [1] 0.06929961
```

```
mean(FDII_middle$PCC)
```

```
## [1] 0.143201
```

```
mean(FDII_middle$PCC_se)
```

```
## [1] 0.1014354
```

```
mean(FDII_lower$PCC)
```

```
## [1] 0.143991
```

```
mean(FDII_lower$PCC_se)
```

```
## [1] 0.1421433
```

Meta generation setup

```r
# Meta generation set-up
m.gen <- metagen(TE = FDII$PCC,
                 seTE = FDII$PCC_se,
                 studlab = FDII$citation,
                 data = FDII,
                 sm = "SMD",
                 comb.fixed = FALSE,
                 comb.random = TRUE,
                 method.tau = "ML",
                 weighted = TRUE,
                 hakn = TRUE,
                 test = "knha",
                 title = "FDI Impacts on Income Inequality")
m.gen
```

```
## Review:     FDI Impacts on Income Inequality
##
## Number of studies: k = 616
##
##                                SMD           95%-CI    t  p-value
## Random effects model (HK) 0.0822 [0.0642; 0.1001] 8.99 < 0.0001
##
## Quantifying heterogeneity:
##  tau^2 = 0.0356 [0.0360; 0.0492]; tau = 0.1886 [0.1899; 0.2217]
##  I^2 = 84.9% [83.8%; 85.9%]; H = 2.57 [2.49; 2.66]
##
## Test of heterogeneity:
##        Q d.f. p-value
##  4070.12  615       0
##
## Details on meta-analytical method:
## - Inverse variance method
## - Maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp adjustment for random effects model (df = 615)
```

Meta generation for different income levels

```r
m.higher <- metagen(TE = FDII_higher$PCC,
                 seTE = FDII_higher$PCC_se,
                 studlab = FDII_higher$citation,
                 data = FDII_higher,
                 sm = "SMD",
                 comb.fixed = FALSE,
                 comb.random = TRUE,
                 method.tau = "ML",
                 weighted = TRUE,
                 hakn = TRUE,
                 test = "knha",
                 title = "FDI Impacts on Income Inequality (Higher Income)")
m.higher
```

```
## Review:     FDI Impacts on Income Inequality (Higher Income)
##
## Number of studies: k = 199
##
```

```
##                               SMD            95%-CI     t p-value
## Random effects model (HK) -0.0170 [-0.0458; 0.0117] -1.17  0.2434
##
## Quantifying heterogeneity:
##  tau^2 = 0.0279 [0.0305; 0.0526]; tau = 0.1670 [0.1747; 0.2294]
##  I^2 = 83.6% [81.5%; 85.5%]; H = 2.47 [2.33; 2.62]
##
## Test of heterogeneity:
##        Q d.f.  p-value
##  1208.42  198 < 0.0001
##
## Details on meta-analytical method:
## - Inverse variance method
## - Maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp adjustment for random effects model (df = 198)
```

```
m.middle <- metagen(TE = FDII_middle$PCC,
                 seTE = FDII_middle$PCC_se,
                 studlab = FDII_middle$citation,
                 data = FDII_middle,
                 sm = "SMD",
                 comb.fixed = FALSE,
                 comb.random = TRUE,
                 method.tau = "ML",
                 weighted = TRUE,
                 hakn = TRUE,
                 test = "knha",
                 title = "FDI Impacts on Income Inequality (Middle Income)")
m.middle
```

```
## Review:     FDI Impacts on Income Inequality (Middle Income)
##
## Number of studies: k = 260
##
##                              SMD           95%-CI     t  p-value
## Random effects model (HK) 0.1360 [0.1113; 0.1607] 10.83 < 0.0001
##
## Quantifying heterogeneity:
##  tau^2 = 0.0269 [0.0247; 0.0408]; tau = 0.1642 [0.1571; 0.2019]
##  I^2 = 79.0% [76.5%; 81.2%]; H = 2.18 [2.06; 2.31]
##
## Test of heterogeneity:
##        Q d.f.  p-value
##  1233.00  259 < 0.0001
##
## Details on meta-analytical method:
## - Inverse variance method
## - Maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp adjustment for random effects model (df = 259)
```

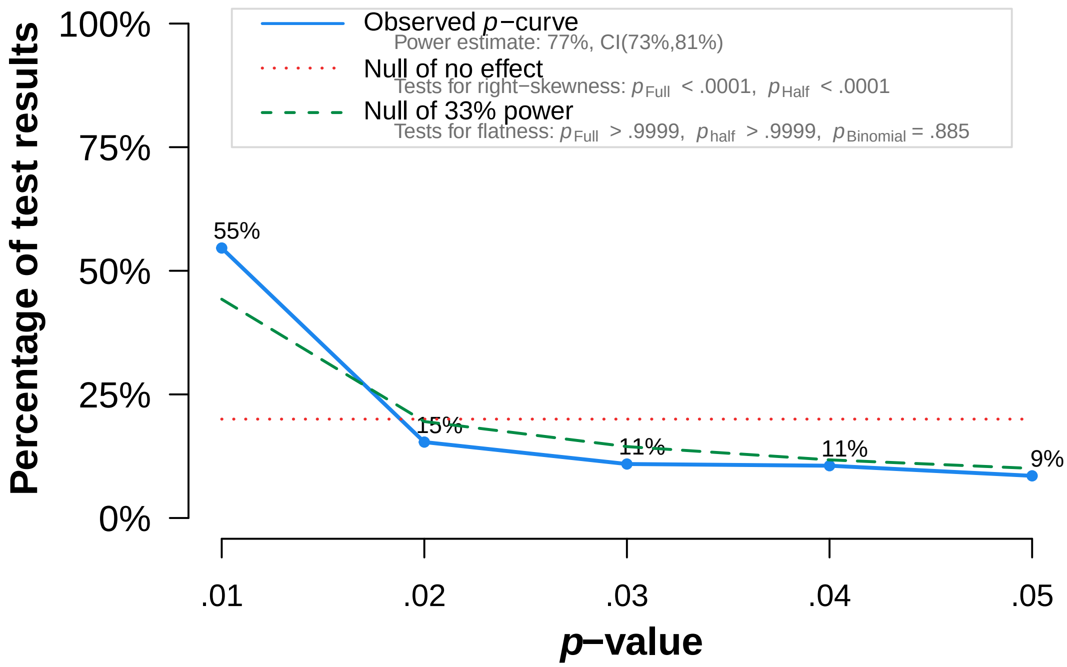
```
m.lower <- metagen(TE = FDII_lower$PCC,
                 seTE = FDII_lower$PCC_se,
                 studlab = FDII_lower$citation,
                 data = FDII_lower,
                 sm = "SMD",
```

```r
                 comb.fixed = FALSE,
                 comb.random = TRUE,
                 method.tau = "ML",
                 weighted = TRUE,
                 hakn = TRUE,
                 test = "knha",
                 title = "FDI Impacts on Income Inequality (Lower Income)")
m.lower
```

```
## Review:     FDI Impacts on Income Inequality (Lower Income)
##
## Number of studies: k = 157
##
##                              SMD           95%-CI    t  p-value
## Random effects model (HK) 0.1413 [0.1042; 0.1783] 7.53 < 0.0001
##
## Quantifying heterogeneity:
##  tau^2 = 0.0341 [0.0266; 0.0534]; tau = 0.1846 [0.1631; 0.2312]
##  I^2 = 85.4% [83.4%; 87.2%]; H = 2.62 [2.45; 2.80]
##
## Test of heterogeneity:
##        Q d.f.  p-value
##  1069.09  156 < 0.0001
##
## Details on meta-analytical method:
## - Inverse variance method
## - Maximum-likelihood estimator for tau^2
## - Q-Profile method for confidence interval of tau^2 and tau
## - Hartung-Knapp adjustment for random effects model (df = 156)
```

p-value test

```r
pcurve(m.gen)
```



Note: The observed p−curve includes 293 statistically significant ($p < .05$) results, of which 218 are $p < .025$. There were 323 additional results entered but excluded from p−curve because they were $p > .05$.

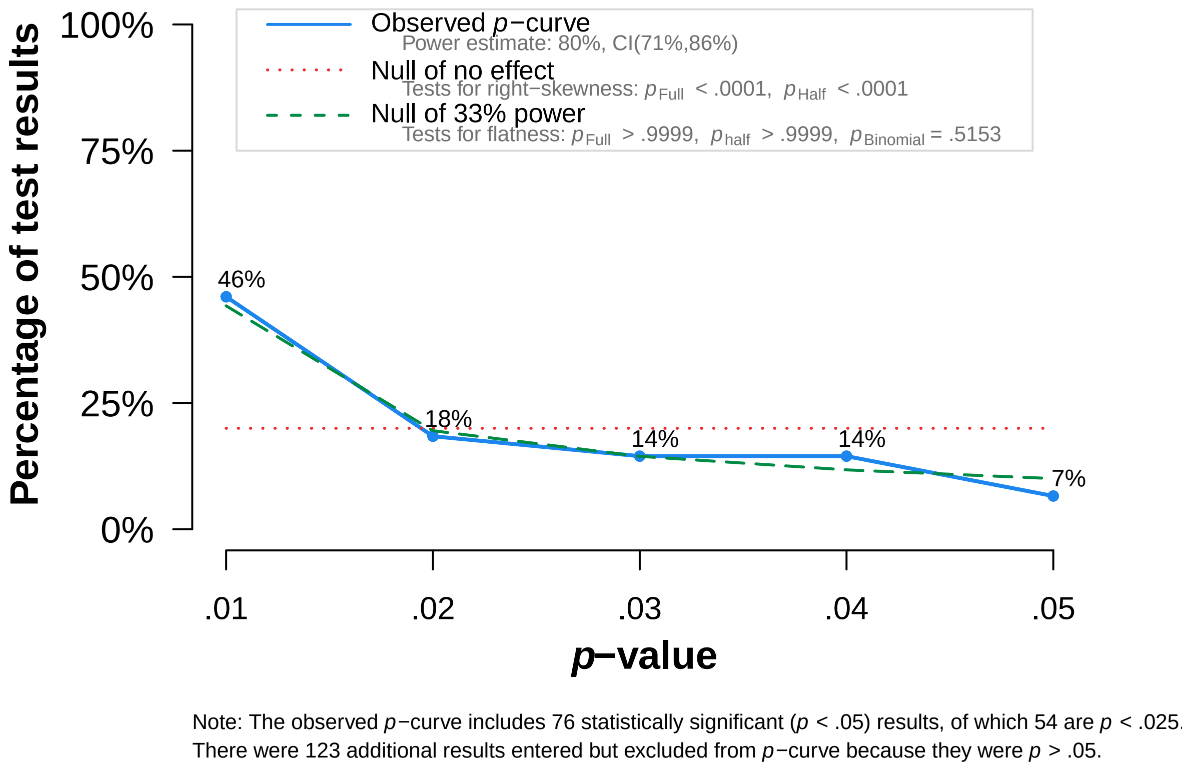
```
## P-curve analysis
##  -----------------------
## - Total number of provided studies: k = 616
## - Total number of p<0.05 studies included into the analysis: k = 293 (47.56%)
## - Total number of studies with p<0.025: k = 218 (35.39%)
##
## Results
##  -----------------------
##                      pBinomial   zFull pFull   zHalf pHalf
## Right-skewness test      0.000 -24.912     0 -25.329     0
## Flatness test            0.885  12.531     1  27.610     1
## Note: p-values of 0 or 1 correspond to p<0.001 and p>0.999, respectively.
## Power Estimate: 77% (72.9%-81.2%)
##
## Evidential value
##  -----------------------
## - Evidential value present: yes
## - Evidential value absent/inadequate: no
```

Talk about how this protects against not just publication bias, but also against p-hacking.

This p-curve test plots statistically-significant ($p < 0.05$) values and assesses whether there is 'evidential' value in the underlying impact of FDI on income inequality. C.47% of results being statistically significant. The flatness test (the graph) and the right-skewness test suggest that our data is not flat, but is skewed to the right – this is what we expect to witness if there is a true underlying effect. There is a 77% chance of detecting that true underlying effect should it exist.

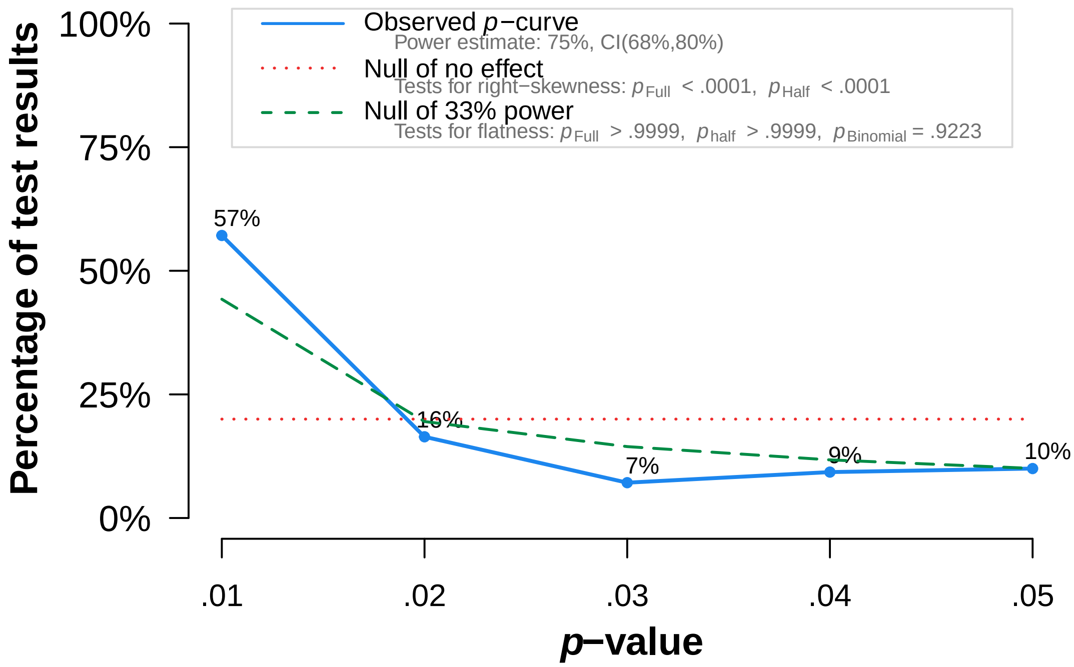
Other studies need to be considered (heterogeneity, study quality, etc.), but this indicates that there is an effect.

```r
pcurve(m.higher)
```




```
## P-curve analysis
##  -----------------------
## - Total number of provided studies: k = 199
## - Total number of p<0.05 studies included into the analysis: k = 76 (38.19%)
## - Total number of studies with p<0.025: k = 54 (27.14%)
##
## Results
##  -----------------------
##                        pBinomial   zFull pFull   zHalf pHalf
## Right-skewness test        0.000 -12.792     0 -12.992     0
## Flatness test              0.515   6.771     1  14.872     1
## Note: p-values of 0 or 1 correspond to p<0.001 and p>0.999, respectively.
## Power Estimate: 80% (71%-86.2%)
##
## Evidential value
##  -----------------------
## - Evidential value present: yes
## - Evidential value absent/inadequate: no
```

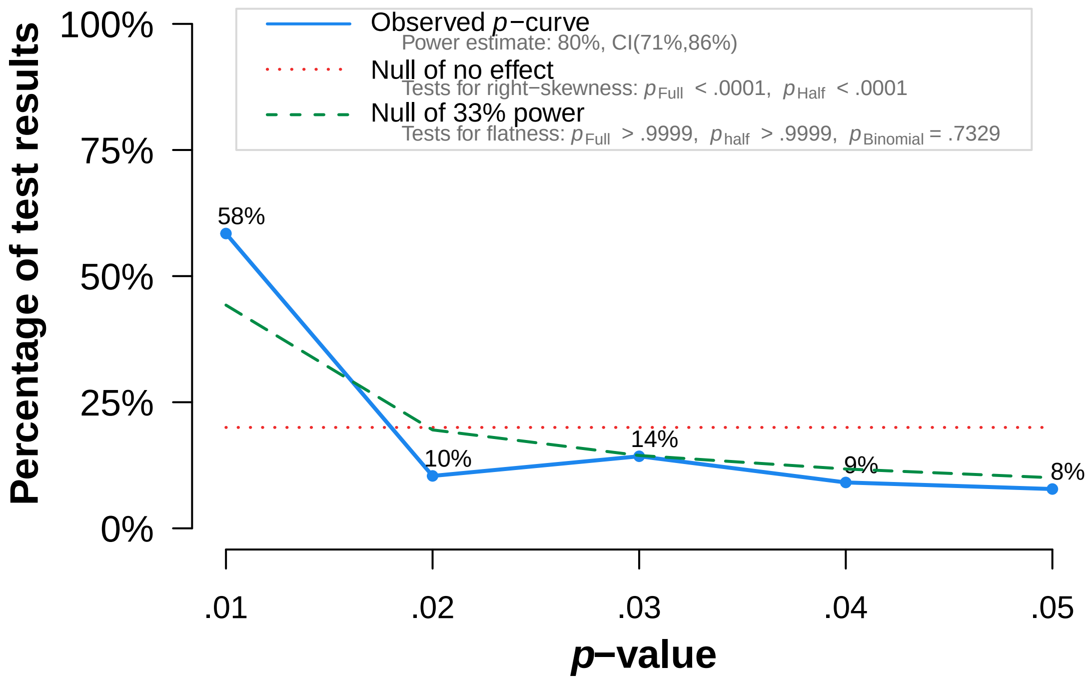
```r
pcurve(m.middle)
```



Note: The observed p−curve includes 140 statistically significant ($p < .05$) results, of which 107 are $p < .025$.
There were 120 additional results entered but excluded from p−curve because they were $p > .05$.

```
## P-curve analysis
##  -----------------------
## - Total number of provided studies: k = 260
## - Total number of p<0.05 studies included into the analysis: k = 140 (53.85%)
## - Total number of studies with p<0.025: k = 107 (41.15%)
##
## Results
##  -----------------------
##                     pBinomial   zFull pFull   zHalf pHalf
## Right-skewness test     0.000 -16.917     0 -16.878     0
## Flatness test           0.922   8.057     1  18.081     1
## Note: p-values of 0 or 1 correspond to p<0.001 and p>0.999, respectively.
## Power Estimate: 75% (67.6%-80.5%)
##
## Evidential value
##  -----------------------
## - Evidential value present: yes
## - Evidential value absent/inadequate: no
```

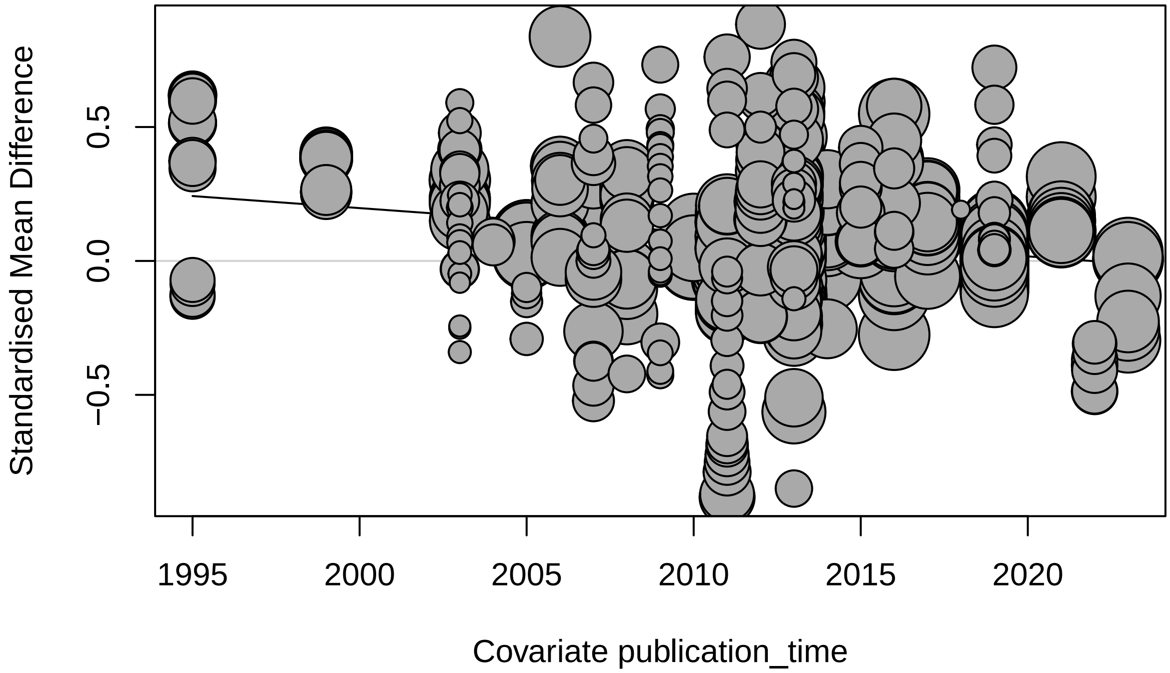
```
pcurve(m.lower)
```



Note: The observed p−curve includes 77 statistically significant (p < .05) results, of which 57 are p < .025. There were 80 additional results entered but excluded from p−curve because they were p > .05.

```
## P-curve analysis
##  -----------------------
## - Total number of provided studies: k = 157
## - Total number of p<0.05 studies included into the analysis: k = 77 (49.04%)
## - Total number of studies with p<0.025: k = 57 (36.31%)
##
## Results
##  -----------------------
##                        pBinomial   zFull pFull   zHalf pHalf
## Right-skewness test        0.000 -13.077     0 -13.764     0
## Flatness test              0.733   6.853     1  14.747     1
## Note: p-values of 0 or 1 correspond to p<0.001 and p>0.999, respectively.
## Power Estimate: 80% (71.2%-86.3%)
##
## Evidential value
##  -----------------------
## - Evidential value present: yes
## - Evidential value absent/inadequate: no
```

```
mgen.test <- metareg(m.gen, ~publication_time)
bubble(mgen.test, studlab = FALSE)
```




```r
FDII[c("if_published", "publication_time", "PCC", "developed_average")] %>% cor()
```

```
##                  if_published publication_time        PCC developed_average
## if_published        1.0000000        0.1998576  0.1814428        -0.2568392
## publication_time    0.1998576        1.0000000 -0.2115030         0.1997580
## PCC                 0.1814428       -0.2115030  1.0000000        -0.3075202
## developed_average  -0.2568392        0.1997580 -0.3075202         1.0000000
```

```r
FDII_higher[c("if_published", "publication_time", "PCC", "developed_average")] %>%
↪  cor()
```

```
##                  if_published publication_time         PCC developed_average
## if_published       1.00000000        0.3913788  0.24122324       -0.09181083
## publication_time   0.39137881        1.0000000 -0.05010500        0.13128095
## PCC                0.24122324       -0.0501050  1.00000000        0.03194525
## developed_average -0.09181083        0.1312810  0.03194525        1.00000000
```

```r
FDII_middle[c("if_published", "publication_time", "PCC", "developed_average")] %>%
↪  cor()
```

```
##                  if_published publication_time         PCC developed_average
## if_published       1.00000000        0.2634164 -0.01485552        0.06354018
## publication_time   0.26341641        1.0000000 -0.23807062        0.23534066
## PCC               -0.01485552       -0.2380706  1.00000000       -0.20033693
## developed_average  0.06354018        0.2353407 -0.20033693        1.00000000
```

```r
FDII_lower[c("if_published", "publication_time", "PCC", "developed_average")] %>%
↪  cor()
```

```
## Warning in cor(.): the standard deviation is zero
```

```
##                  if_published publication_time          PCC developed_average
## if_published                1               NA           NA                NA
## publication_time           NA       1.00000000 -0.098600583      -0.057941622
## PCC                        NA      -0.09860058  1.000000000       0.002377134
## developed_average          NA      -0.05794162  0.002377134       1.000000000
```

PET-PEESE tests

```
data.petpeese <- data.frame(TE = m.gen$TE, seTE = m.gen$seTE, varTE = m.gen$seTE^2)
data.petpeese$w_k <- 1/(data.petpeese$varTE)

pet <- lm(TE ~ seTE, weights = w_k, data = data.petpeese)
pet_summary <- summary(pet)$coefficients
print(round(pet_summary, 3))
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.003      0.004   0.767    0.443
## seTE           0.924      0.124   7.434    0.000
```

```
peese <- lm(TE ~ varTE, weights = w_k, data = data.petpeese)
peese_summary <- summary(peese)$coefficients
print(round(peese_summary, 3))
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.017      0.004   4.529        0
## varTE          4.813      0.864   5.571        0
```

```
### m.gen.inf <- InfluenceAnalysis(m.gen, random = TRUE) # This takes a long time to
↪  run!
### plot(m.gen.inf, "baujat")
```

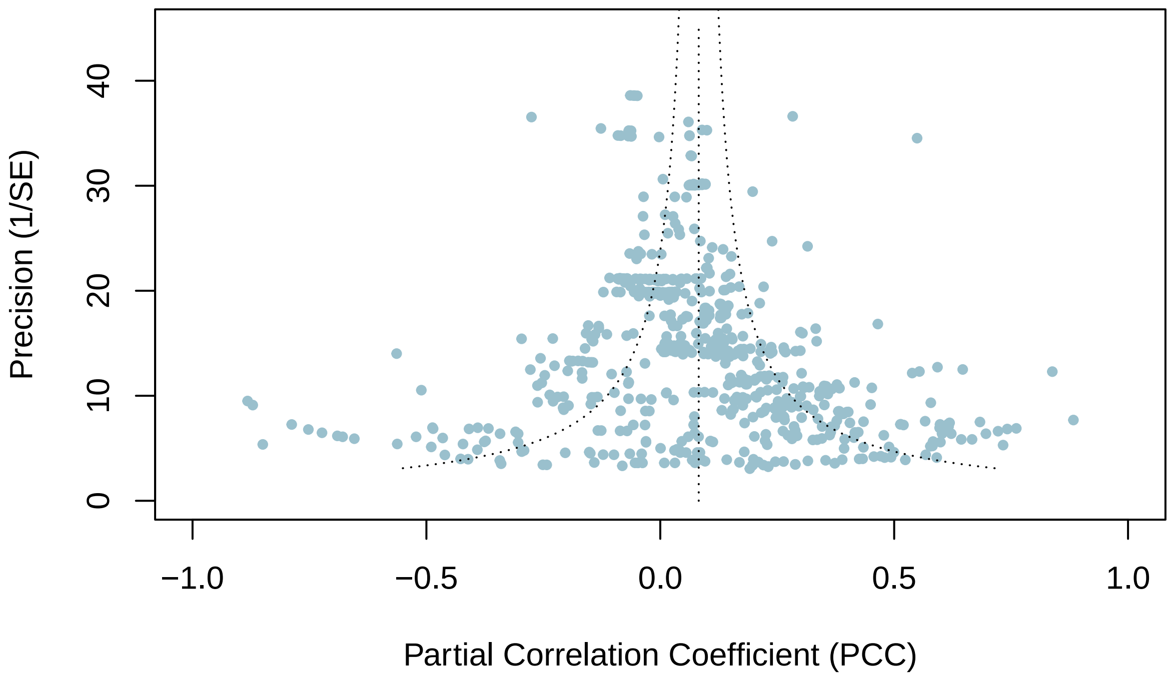
Only run the noted aspects when necessary as the influence analysis takes forever to run!

Inputs to run meta regression

```
# Set it such that the control variables are put into one column
FDII_long <- FDII %>%
  pivot_longer(cols = starts_with("control_variables_"),
               names_to = "Control_Variable_Type",
               values_to = "Control_Variable_Name",
               names_repair = "unique") %>%
  filter(!is.na(Control_Variable_Name) & Control_Variable_Name != "")

# Turn that one column into a grid of binary variables
FDII_wide <- FDII_long %>%
  mutate(Presence = 1) %>%
  pivot_wider(names_from = Control_Variable_Name, values_from = Presence,
   ↪  values_fill = list(Presence = 0), names_repair = "unique")

# Flatten to remove duplicate rows and merge the binary variables
FDII_merged <- FDII_wide %>%
  group_by(citation, coefficient, t_value_calculated, study_country, journal_name,
   ↪  estimation_methods) %>%
  summarise(across(where(is.numeric), ~if(all(is.na(.))) 0 else max(., na.rm =
   ↪  TRUE)), .groups = 'drop')

# Perform a left join to integrate the new columns from FDII_merged back into FDII
FDII <- left_join(FDII, FDII_merged, by = c("citation", "coefficient",
↪  "t_value_calculated", "study_country", "journal_name", "estimation_methods",
↪  "journal_rank"))
FDII <- FDII %>% dplyr::select(-matches("\\.y$"))
FDII <- FDII %>% dplyr::rename_with(~str_replace(., "\\.x$", ""), .cols =
↪  ends_with(".x"))
```

Funnel plot for meta analysis (FAT)

```r
funnel_data <- metafor::funnel(m.gen,
                      title = "FDI Impacts on Income Inequality",
                      xlim = c(-1, 1),
                      ylim = rev(c(45, 0)),
                      studlab = FALSE,
                      xlab = "Partial Correlation Coefficient (PCC)",
                      ylab = "Precision (1/SE)",
                      yaxis = "invse",
                      col = "lightblue3",
                      bg = "lightblue3",
                      cex = 0.6)
```




```r
eggers.test(m.gen)
```

```
## Eggers' test of the intercept
## ==============================
##
##  intercept       95% CI      t                       p
##      0.924 0.68 - 1.17 7.434 0.0000000000003562739
##
## Eggers' test indicates the presence of funnel plot asymmetry.
```

The presence of funnel plot asymmetry could suggest that we have a publication bias. It could also be indicative of (Page et al., 2020):

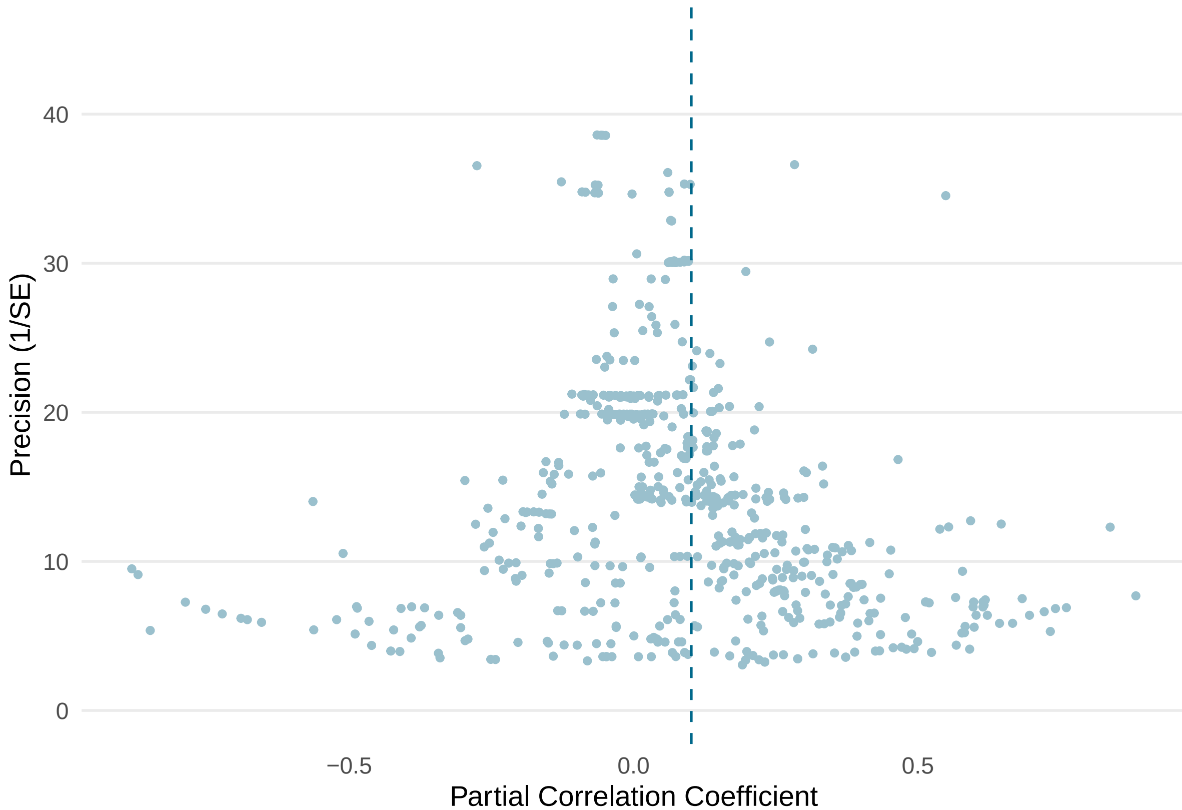
- Between-study heterogeneity - the plot assumes effect sizes are the result of the studies' sampling error, but it could be different true sizes
- Different study methods, leading to greater effects
- Lower-quality studies could show greater effect sizes as a result of greater bias risk
- Could occur by chance (unlikely)

```r
ggplot(data = FDII, aes(x = PCC, y = 1/PCC_se)) +
  geom_point(color = "lightblue3", size = 1) +
  geom_vline(aes(xintercept = mean(FDII$PCC_se)), linetype = "dashed", color =
  ↪  "deepskyblue4") +
```

```r
  ylim(0, 45) +
  labs(x = "Partial Correlation Coefficient",
       y = "Precision (1/SE)") +
  theme_minimal() +
  theme(panel.grid.minor.y = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.major.x = element_blank())
```

```
## Warning: Use of `FDII$PCC_se` is discouraged.
## i Use `PCC_se` instead.
```

```
## Warning: Removed 12 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



FAT-PET test

```r
res <- rma(FDII$PCC, sei = FDII$PCC_se)
#FATPET test
regtest(res, model = "rma", ret.fit = TRUE)
```

```
##
## Regression Test for Funnel Plot Asymmetry
##
## Model:     mixed-effects meta-regression model
## Predictor: standard error
##
## Mixed-Effects Model (k = 616; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0341 (SE = 0.0024)
```

```
## tau (square root of estimated tau^2 value):             0.1847
## I^2 (residual heterogeneity / unaccounted variability): 96.26%
## H^2 (unaccounted variability / sampling variability):   26.72
## R^2 (amount of heterogeneity accounted for):            4.36%
##
## Test for Residual Heterogeneity:
## QE(df = 614) = 3734.0544, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 14.1993, p-val = 0.0002
##
## Model Results:
##
##          estimate      se    zval    pval   ci.lb   ci.ub
## intrcpt    0.0330  0.0155  2.1293  0.0332  0.0026  0.0634    *
## sei        0.5828  0.1547  3.7682  0.0002  0.2797  0.8859  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Test for Funnel Plot Asymmetry: z = 3.7682, p = 0.0002
## Limit Estimate (as sei -> 0):   b = 0.0330 (CI: 0.0026, 0.0634)
```

FAT-PET test – Higher income

```r
res_higher <- rma(FDII_higher$PCC, sei = FDII_higher$PCC_se)
#FATPET test
regtest(res_higher, model = "rma", ret.fit = TRUE)
```

```
##
## Regression Test for Funnel Plot Asymmetry
##
## Model:     mixed-effects meta-regression model
## Predictor: standard error
##
## Mixed-Effects Model (k = 199; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0268 (SE = 0.0031)
## tau (square root of estimated tau^2 value):             0.1638
## I^2 (residual heterogeneity / unaccounted variability): 97.50%
## H^2 (unaccounted variability / sampling variability):   39.94
## R^2 (amount of heterogeneity accounted for):            4.67%
##
## Test for Residual Heterogeneity:
## QE(df = 197) = 1200.1232, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 7.0943, p-val = 0.0077
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt    0.0353  0.0233   1.5142  0.1300  -0.0104   0.0810
## sei       -0.8783  0.3298  -2.6635  0.0077  -1.5247  -0.2320  **
##
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Test for Funnel Plot Asymmetry: z = -2.6635, p = 0.0077
## Limit Estimate (as sei -> 0):   b =  0.0353 (CI: -0.0104, 0.0810)
```

FAT-PET test – Middle income

```
res_middle <- rma(FDII_middle$PCC, sei = FDII_middle$PCC_se)
#FATPET test
regtest(res_middle, model = "rma", ret.fit = TRUE)
```

```
##
## Regression Test for Funnel Plot Asymmetry
##
## Model:     mixed-effects meta-regression model
## Predictor: standard error
##
## Mixed-Effects Model (k = 260; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0246 (SE = 0.0029)
## tau (square root of estimated tau^2 value):             0.1569
## I^2 (residual heterogeneity / unaccounted variability): 83.94%
## H^2 (unaccounted variability / sampling variability):   6.22
## R^2 (amount of heterogeneity accounted for):            9.19%
##
## Test for Residual Heterogeneity:
## QE(df = 258) = 1104.7160, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 11.6317, p-val = 0.0006
##
## Model Results:
##
##          estimate      se    zval    pval   ci.lb   ci.ub
## intrcpt    0.0583  0.0253  2.3011  0.0214  0.0086  0.1080    *
## sei        0.8792  0.2578  3.4105  0.0006  0.3739  1.3844  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Test for Funnel Plot Asymmetry: z = 3.4105, p = 0.0006
## Limit Estimate (as sei -> 0):   b = 0.0583 (CI: 0.0086, 0.1080)
```

FAT-PET test – Lower income

```
res_lower <- rma(FDII_lower$PCC, sei = FDII_lower$PCC_se)
#FATPET test
regtest(res_lower, model = "rma", ret.fit = TRUE)
```

```
##
## Regression Test for Funnel Plot Asymmetry
##
## Model:     mixed-effects meta-regression model
```

```
## Predictor: standard error
##
## Mixed-Effects Model (k = 157; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0345 (SE = 0.0056)
## tau (square root of estimated tau^2 value):             0.1858
## I^2 (residual heterogeneity / unaccounted variability): 95.59%
## H^2 (unaccounted variability / sampling variability):   22.67
## R^2 (amount of heterogeneity accounted for):            0.00%
##
## Test for Residual Heterogeneity:
## QE(df = 155) = 941.0447, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.3828, p-val = 0.5361
##
## Model Results:
##
##          estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt    0.1240  0.0335  3.7053  0.0002   0.0584  0.1895  ***
## sei        0.1548  0.2502  0.6187  0.5361  -0.3356  0.6452
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Test for Funnel Plot Asymmetry: z = 0.6187, p = 0.5361
## Limit Estimate (as sei -> 0):   b = 0.1240 (CI: 0.0584, 0.1895)
```

Meta Regression Analyses (MRAs)

Basic meta analysis test with no moderators

```
meta.anal <- rma(yi = PCC,
                 sei = PCC_se,
                 data = FDII,
                 method = "REML",
                 mods = ~ PCC_se,
                 test = "knha")
meta.anal
```

```
##
## Mixed-Effects Model (k = 616; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0341 (SE = 0.0024)
## tau (square root of estimated tau^2 value):             0.1847
## I^2 (residual heterogeneity / unaccounted variability): 96.26%
## H^2 (unaccounted variability / sampling variability):   26.72
## R^2 (amount of heterogeneity accounted for):            4.36%
##
## Test for Residual Heterogeneity:
## QE(df = 614) = 3734.0544, p-val < .0001
##
## Test of Moderators (coefficient 2):
## F(df1 = 1, df2 = 614) = 12.4933, p-val = 0.0004
##
## Model Results:
```

```
##
##          estimate      se    tval   df    pval   ci.lb   ci.ub
## intrcpt    0.0330  0.0165  1.9973  614  0.0462  0.0006  0.0655    *
## PCC_se     0.5828  0.1649  3.5346  614  0.0004  0.2590  0.9066  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
selmodel(meta.anal, type = "stepfun", steps = 0.025, alternative = "two.sided")
```

```
##
## Mixed-Effects Model (k = 616; tau^2 estimator: ML)
##
## tau^2 (estimated amount of residual heterogeneity): 0.0462 (SE = 0.0050)
## tau (square root of estimated tau^2 value):         0.2148
##
## Test for Residual Heterogeneity:
## LRT(df = 1) = 1396.7982, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 12.5437, p-val = 0.0004
##
## Model Results:
##
##          estimate      se    zval    pval   ci.lb   ci.ub
## intrcpt    0.0476  0.0195  2.4424  0.0146  0.0094  0.0857    *
## PCC_se     0.6530  0.1844  3.5417  0.0004  0.2916  1.0143  ***
##
## Test for Selection Model Parameters:
## LRT(df = 1) = 20.5873, p-val < .0001
##
## Selection Model Results:
##
##                      k  estimate      se    zval    pval   ci.lb   ci.ub
## 0     < p <= 0.025  218    1.0000     ---     ---     ---     ---     ---
## 0.025 < p <= 1      398    1.7544  0.2174  3.4699  0.0005  1.3283  2.1805  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
meta.anal.full.model <- rma(yi = PCC,
                            sei = PCC_se,
                            data = FDII,
                            method = "REML",
                            mods = ~ stock + GINI_control + I_Share + W_I_Disp + Other_I_measure + mean_log_GDPPC + est_method +
                            ↪  GDP_control + education_control + single_country + trade_control + population_control +
                            ↪  inflation_control + unemployment_control + gov_inequality_effort_control + if_published + PCC_se,
                            test = "knha",
                            digits = 3)
```

```
## Warning: Redundant predictors dropped from the model.
```

```
meta.anal.full.model
```

```
##
## Mixed-Effects Model (k = 616; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.027 (SE = 0.002)
## tau (square root of estimated tau^2 value):             0.164
## I^2 (residual heterogeneity / unaccounted variability): 94.92%
## H^2 (unaccounted variability / sampling variability):   19.67
## R^2 (amount of heterogeneity accounted for):            24.96%
##
## Test for Residual Heterogeneity:
## QE(df = 599) = 3073.598, p-val < .001
##
## Test of Moderators (coefficients 2:17):
## F(df1 = 16, df2 = 599) = 8.055, p-val < .001
##
## Model Results:
##
##                                estimate     se    tval   df   pval   ci.lb
## intrcpt                           0.348  0.200   1.743  599  0.082  -0.044
## stock                             0.019  0.027   0.719  599  0.472  -0.034
## GINI_control                      0.144  0.033   4.368  599  <.001   0.079
## I_Share                           0.250  0.067   3.755  599  <.001   0.119
## W_I_Disp                          0.197  0.041   4.792  599  <.001   0.116
```

```
## mean_log_GDPPC                   -0.056  0.018  -3.022  599  0.003  -0.092
## est_method                        0.018  0.021   0.824  599  0.410  -0.024
## GDP_control                       0.019  0.022   0.900  599  0.369  -0.023
## education_control                -0.008  0.024  -0.342  599  0.732  -0.054
## single_country                   -0.106  0.026  -4.094  599  <.001  -0.156
## trade_control                    -0.007  0.023  -0.321  599  0.749  -0.053
## population_control                0.000  0.033   0.013  599  0.990  -0.065
## inflation_control                -0.029  0.034  -0.837  599  0.403  -0.096
## unemployment_control             -0.036  0.043  -0.833  599  0.405  -0.120
## gov_inequality_effort_control    -0.015  0.030  -0.494  599  0.622  -0.074
## if_published                      0.128  0.040   3.169  599  0.002   0.049
## PCC_se                            0.323  0.198   1.630  599  0.104  -0.066
##                                 ci.ub
## intrcpt                         0.740    .
## stock                           0.072
## GINI_control                    0.208  ***
## I_Share                         0.381  ***
## W_I_Disp                        0.278  ***
## mean_log_GDPPC                 -0.020   **
## est_method                      0.060
## GDP_control                     0.062
## education_control               0.038
## single_country                 -0.055  ***
## trade_control                   0.038
## population_control              0.066
## inflation_control               0.039
## unemployment_control            0.049
## gov_inequality_effort_control   0.044
## if_published                    0.208   **
## PCC_se                          0.712
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
meta.anal.select.model <- rma(yi = PCC,
                             sei = PCC_se,
                             data = FDII,
```

```r
                              method = "REML",
                              mods = ~ GINI_control + I_Share + W_I_Disp + mean_log_GDPPC + single_country + if_published + PCC_se,
                              test = "knha",
                              digits = 3)

meta.anal.select.model
```

```
##
## Mixed-Effects Model (k = 616; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.026 (SE = 0.002)
## tau (square root of estimated tau^2 value):             0.162
## I^2 (residual heterogeneity / unaccounted variability): 94.98%
## H^2 (unaccounted variability / sampling variability):   19.90
## R^2 (amount of heterogeneity accounted for):            26.09%
##
## Test for Residual Heterogeneity:
## QE(df = 608) = 3207.924, p-val < .001
##
## Test of Moderators (coefficients 2:8):
## F(df1 = 7, df2 = 608) = 17.978, p-val < .001
##
## Model Results:
##
##                 estimate     se    tval   df    pval   ci.lb   ci.ub
## intrcpt            0.392  0.180   2.174  608  0.030   0.038   0.746    *
## GINI_control       0.140  0.031   4.508  608  <.001   0.079   0.202  ***
## I_Share            0.194  0.043   4.538  608  <.001   0.110   0.278  ***
## W_I_Disp           0.192  0.039   4.949  608  <.001   0.116   0.268  ***
## mean_log_GDPPC    -0.059  0.017  -3.546  608  <.001  -0.091  -0.026  ***
## single_country    -0.102  0.023  -4.489  608  <.001  -0.147  -0.057  ***
## if_published       0.124  0.038   3.265  608  0.001   0.050   0.199   **
## PCC_se             0.366  0.182   2.018  608  0.044   0.010   0.723    *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
meta.anal.select.model.median <- rma(yi = PCC,
                                      sei = PCC_se,
                                      data = FDII,
                                      method = "REML",
                                      mods = ~ GINI_control + I_Share + W_I_Disp + median_log_GDPPC + single_country + if_published + PCC_se,
                                      test = "knha",
                                      digits = 3)

meta.anal.select.model.median
```

```
##
## Mixed-Effects Model (k = 616; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.027 (SE = 0.002)
## tau (square root of estimated tau^2 value):             0.164
## I^2 (residual heterogeneity / unaccounted variability): 95.05%
## H^2 (unaccounted variability / sampling variability):   20.21
## R^2 (amount of heterogeneity accounted for):            24.89%
##
## Test for Residual Heterogeneity:
## QE(df = 608) = 3218.079, p-val < .001
##
## Test of Moderators (coefficients 2:8):
## F(df1 = 7, df2 = 608) = 17.169, p-val < .001
##
## Model Results:
##
##                   estimate     se    tval   df    pval   ci.lb   ci.ub
## intrcpt              0.224  0.168   1.333  608   0.183  -0.106   0.553
## GINI_control         0.144  0.031   4.578  608   <.001   0.082   0.205  ***
## I_Share              0.177  0.042   4.189  608   <.001   0.094   0.260  ***
## W_I_Disp             0.192  0.039   4.894  608   <.001   0.115   0.268  ***
## median_log_GDPPC    -0.043  0.015  -2.789  608   0.005  -0.073  -0.013   **
## single_country      -0.094  0.024  -3.981  608   <.001  -0.141  -0.048  ***
## if_published         0.132  0.038   3.440  608   <.001   0.057   0.208  ***
```

```
## PCC_se               0.410  0.182   2.259  608  0.024   0.054   0.767    *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Meta tests – Higher income

```
meta.anal.higher <- rma(yi = PCC,
                 sei = PCC_se,
                 data = FDII_higher,
                 method = "REML",
                 mods = ~ PCC_se,
                 test = "knha",
                 digits = 3)
meta.anal.higher
```

```
##
## Mixed-Effects Model (k = 199; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.027 (SE = 0.003)
## tau (square root of estimated tau^2 value):             0.164
## I^2 (residual heterogeneity / unaccounted variability): 97.50%
## H^2 (unaccounted variability / sampling variability):   39.94
## R^2 (amount of heterogeneity accounted for):            4.67%
##
## Test for Residual Heterogeneity:
## QE(df = 197) = 1200.123, p-val < .001
##
## Test of Moderators (coefficient 2):
## F(df1 = 1, df2 = 197) = 5.520, p-val = 0.020
##
## Model Results:
##
##          estimate     se    tval   df    pval   ci.lb   ci.ub
## intrcpt     0.035  0.026   1.336  197  0.183  -0.017   0.087
## PCC_se     -0.878  0.374  -2.349  197  0.020  -1.616  -0.141  *
```

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

```
meta.anal.full.model.higher <- rma(yi = PCC,
                             sei = PCC_se,
                             data = FDII_higher,
                             method = "REML",
                             mods = ~ stock + GINI_control + I_Share + W_I_Disp + est_method + GDP_control + education_control +
                             ↪  single_country + trade_control + population_control + inflation_control + unemployment_control +
                             ↪  gov_inequality_effort_control + if_published + PCC_se,
                             test = "knha",
                             digits = 3)

meta.anal.full.model.higher
```

```
##
## Mixed-Effects Model (k = 199; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.006 (SE = 0.001)
## tau (square root of estimated tau^2 value):             0.074
## I^2 (residual heterogeneity / unaccounted variability): 89.34%
## H^2 (unaccounted variability / sampling variability):   9.38
## R^2 (amount of heterogeneity accounted for):            80.35%
##
## Test for Residual Heterogeneity:
## QE(df = 183) = 589.272, p-val < .001
##
## Test of Moderators (coefficients 2:16):
## F(df1 = 15, df2 = 183) = 9.827, p-val < .001
##
## Model Results:
##
##                                estimate     se    tval   df    pval   ci.lb
## intrcpt                          -0.339  0.074  -4.599  183  <.001  -0.484
## stock                             0.089  0.049   1.826  183  0.069  -0.007
## GINI_control                      0.009  0.043   0.204  183  0.838  -0.075
```

```
## I_Share                        0.110  0.086   1.274  183  0.204  -0.060
## W_I_Disp                       0.357  0.076   4.706  183  <.001   0.207
## est_method                     0.005  0.021   0.263  183  0.793  -0.035
## GDP_control                   -0.206  0.059  -3.494  183  <.001  -0.322
## education_control             -0.313  0.080  -3.931  183  <.001  -0.471
## single_country                -0.047  0.069  -0.686  183  0.494  -0.183
## trade_control                  0.559  0.121   4.623  183  <.001   0.320
## population_control             0.667  0.112   5.944  183  <.001   0.445
## inflation_control              0.489  0.094   5.188  183  <.001   0.303
## unemployment_control          -0.480  0.081  -5.950  183  <.001  -0.639
## gov_inequality_effort_control  0.654  0.096   6.839  183  <.001   0.465
## if_published                   0.116  0.037   3.161  183  0.002   0.044
## PCC_se                        -0.029  0.451  -0.064  183  0.949  -0.919
##                              ci.ub
## intrcpt                     -0.194  ***
## stock                        0.186    .
## GINI_control                 0.093
## I_Share                      0.280
## W_I_Disp                     0.507  ***
## est_method                   0.046
## GDP_control                 -0.090  ***
## education_control           -0.156  ***
## single_country               0.088
## trade_control                0.798  ***
## population_control           0.888  ***
## inflation_control            0.675  ***
## unemployment_control        -0.321  ***
## gov_inequality_effort_control  0.843  ***
## if_published                 0.189   **
## PCC_se                       0.861
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
meta.anal.select.model.higher <- rma(yi = PCC,
                            sei = PCC_se,
                            data = FDII_higher,
```

```r
                              method = "REML",
                              mods = ~ stock + W_I_Disp + GDP_control + education_control + trade_control + population_control +
                              ↪  inflation_control + unemployment_control + gov_inequality_effort_control + if_published + PCC_se,
                              test = "knha",
                              digits = 3)

meta.anal.select.model.higher
```

```
##
## Mixed-Effects Model (k = 199; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.005 (SE = 0.001)
## tau (square root of estimated tau^2 value):             0.070
## I^2 (residual heterogeneity / unaccounted variability): 88.08%
## H^2 (unaccounted variability / sampling variability):   8.39
## R^2 (amount of heterogeneity accounted for):            82.47%
##
## Test for Residual Heterogeneity:
## QE(df = 187) = 593.338, p-val < .001
##
## Test of Moderators (coefficients 2:12):
## F(df1 = 11, df2 = 187) = 13.568, p-val < .001
##
## Model Results:
##
##                                estimate     se    tval   df    pval    ci.lb
## intrcpt                          -0.325  0.061  -5.352  187  <.001  -0.445
## stock                             0.089  0.046   1.942  187  0.054  -0.001
## W_I_Disp                          0.303  0.053   5.712  187  <.001   0.198
## GDP_control                      -0.206  0.054  -3.827  187  <.001  -0.312
## education_control                -0.306  0.071  -4.304  187  <.001  -0.446
## trade_control                     0.555  0.093   5.948  187  <.001   0.371
## population_control                0.688  0.107   6.450  187  <.001   0.478
## inflation_control                 0.504  0.075   6.765  187  <.001   0.357
## unemployment_control             -0.465  0.073  -6.346  187  <.001  -0.609
## gov_inequality_effort_control     0.656  0.080   8.243  187  <.001   0.499
```

```
## if_published                        0.121  0.033   3.675  187  <.001   0.056
## PCC_se                             -0.298  0.399  -0.747  187  0.456  -1.085
##                                  ci.ub
## intrcpt                         -0.205  ***
## stock                            0.179    .
## W_I_Disp                         0.407  ***
## GDP_control                     -0.100  ***
## education_control               -0.166  ***
## trade_control                    0.739  ***
## population_control               0.899  ***
## inflation_control                0.651  ***
## unemployment_control            -0.320  ***
## gov_inequality_effort_control    0.813  ***
## if_published                     0.187  ***
## PCC_se                           0.489
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Meta tests – Middle income

```
meta.anal.full.model.middle <- rma(yi = PCC,
                             sei = PCC_se,
                             data = FDII_middle,
                             method = "REML",
                             mods = ~ stock + GINI_control + I_Share + W_I_Disp + est_method + GDP_control + education_control +
                             ↪  single_country + trade_control + population_control + inflation_control + unemployment_control +
                             ↪  gov_inequality_effort_control + if_published + PCC_se,
                             test = "knha",
                             digits = 3)
```

```
## Warning: Redundant predictors dropped from the model.
```

```
meta.anal.full.model.middle
```

```
##
## Mixed-Effects Model (k = 260; tau^2 estimator: REML)
```

```
##
## tau^2 (estimated amount of residual heterogeneity):     0.023 (SE = 0.003)
## tau (square root of estimated tau^2 value):             0.152
## I^2 (residual heterogeneity / unaccounted variability): 82.53%
## H^2 (unaccounted variability / sampling variability):   5.72
## R^2 (amount of heterogeneity accounted for):            14.74%
##
## Test for Residual Heterogeneity:
## QE(df = 245) = 978.155, p-val < .001
##
## Test of Moderators (coefficients 2:15):
## F(df1 = 14, df2 = 245) = 2.622, p-val = 0.001
##
## Model Results:
##
##                                estimate     se    tval   df   pval   ci.lb
## intrcpt                          -0.026  0.123  -0.212  245  0.832  -0.268
## stock                             0.075  0.041   1.847  245  0.066  -0.005
## GINI_control                      0.057  0.050   1.135  245  0.257  -0.042
## W_I_Disp                          0.067  0.061   1.091  245  0.276  -0.054
## est_method                       -0.084  0.039  -2.130  245  0.034  -0.162
## GDP_control                      -0.047  0.030  -1.563  245  0.119  -0.105
## education_control                -0.051  0.035  -1.435  245  0.153  -0.120
## single_country                   -0.053  0.058  -0.913  245  0.362  -0.167
## trade_control                     0.001  0.030   0.033  245  0.973  -0.058
## population_control               -0.022  0.072  -0.311  245  0.756  -0.163
## inflation_control                -0.016  0.056  -0.285  245  0.776  -0.127
## unemployment_control             -0.057  0.060  -0.948  245  0.344  -0.176
## gov_inequality_effort_control     0.054  0.047   1.164  245  0.246  -0.038
## if_published                      0.080  0.089   0.899  245  0.369  -0.095
## PCC_se                            0.486  0.371   1.310  245  0.191  -0.244
##                                 ci.ub
## intrcpt                         0.216
## stock                           0.155  .
## GINI_control                    0.156
## W_I_Disp                        0.187
## est_method                     -0.006  *
```

```
## GDP_control                     0.012
## education_control               0.019
## single_country                  0.061
## trade_control                   0.060
## population_control              0.119
## inflation_control               0.095
## unemployment_control            0.062
## gov_inequality_effort_control   0.146
## if_published                    0.255
## PCC_se                          1.216
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
meta.anal.select.model.middle <- rma(yi = PCC,
                            sei = PCC_se,
                            data = FDII_middle,
                            method = "REML",
                            mods = ~ stock + est_method + if_published + PCC_se,
                            test = "knha",
                            digits = 3)

meta.anal.select.model.middle
```

```
##
## Mixed-Effects Model (k = 260; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.024 (SE = 0.003)
## tau (square root of estimated tau^2 value):             0.156
## I^2 (residual heterogeneity / unaccounted variability): 83.67%
## H^2 (unaccounted variability / sampling variability):   6.12
## R^2 (amount of heterogeneity accounted for):            10.09%
##
## Test for Residual Heterogeneity:
## QE(df = 255) = 1093.975, p-val < .001
##
## Test of Moderators (coefficients 2:5):
```

```
## F(df1 = 4, df2 = 255) = 4.166, p-val = 0.003
##
## Model Results:
##
##                estimate     se    tval   df    pval   ci.lb   ci.ub
## intrcpt          -0.027  0.100  -0.266  255  0.790  -0.223   0.170
## stock             0.058  0.039   1.493  255  0.137  -0.019   0.135
## est_method       -0.060  0.030  -1.983  255  0.048  -0.120  -0.000   *
## if_published      0.049  0.083   0.593  255  0.554  -0.114   0.212
## PCC_se            0.856  0.307   2.790  255  0.006   0.252   1.461  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Meta tests – Lower income

```
meta.anal.full.model.lower <- rma(yi = PCC,
                             sei = PCC_se,
                             data = FDII_lower,
                             method = "REML",
                             mods = ~ stock + GINI_control + I_Share + W_I_Disp + est_method + GDP_control + education_control +
                              ↪  single_country + trade_control + population_control + inflation_control + unemployment_control +
                              ↪  gov_inequality_effort_control + if_published + PCC_se,
                             test = "knha",
                             digits = 3)
```

```
## Warning: Redundant predictors dropped from the model.
```

```
meta.anal.full.model.lower
```

```
##
## Mixed-Effects Model (k = 157; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.022 (SE = 0.004)
## tau (square root of estimated tau^2 value):             0.149
## I^2 (residual heterogeneity / unaccounted variability): 91.77%
## H^2 (unaccounted variability / sampling variability):   12.15
```

```
## R^2 (amount of heterogeneity accounted for):            35.20%
##
## Test for Residual Heterogeneity:
## QE(df = 143) = 518.648, p-val < .001
##
## Test of Moderators (coefficients 2:14):
## F(df1 = 13, df2 = 143) = 4.260, p-val < .001
##
## Model Results:
##
##                               estimate     se    tval   df   pval   ci.lb
## intrcpt                          0.112  0.103   1.092  143  0.277  -0.091
## stock                            0.074  0.074   1.008  143  0.315  -0.071
## GINI_control                    -0.024  0.113  -0.212  143  0.832  -0.246
## W_I_Disp                        -0.048  0.104  -0.464  143  0.643  -0.255
## est_method                       0.088  0.058   1.514  143  0.132  -0.027
## GDP_control                      0.126  0.045   2.785  143  0.006   0.036
## education_control                0.004  0.062   0.058  143  0.954  -0.119
## single_country                  -0.075  0.059  -1.264  143  0.208  -0.191
## trade_control                    0.035  0.043   0.811  143  0.419  -0.050
## population_control              -0.073  0.063  -1.159  143  0.248  -0.198
## inflation_control               -0.339  0.072  -4.692  143  <.001  -0.482
## unemployment_control             0.146  0.115   1.272  143  0.205  -0.081
## gov_inequality_effort_control   -0.054  0.062  -0.866  143  0.388  -0.176
## PCC_se                          -0.164  0.406  -0.404  143  0.687  -0.966
##                                 ci.ub
## intrcpt                         0.315
## stock                           0.220
## GINI_control                    0.199
## W_I_Disp                        0.158
## est_method                      0.203
## GDP_control                     0.215   **
## education_control               0.126
## single_country                  0.042
## trade_control                   0.121
## population_control              0.052
## inflation_control              -0.196  ***
```

```
## unemployment_control            0.373
## gov_inequality_effort_control   0.069
## PCC_se                          0.638
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
meta.anal.select.model.lower <- rma(yi = PCC,
                            sei = PCC_se,
                            data = FDII_lower,
                            method = "REML",
                            mods = ~ GDP_control + inflation_control + if_published + PCC_se,
                            test = "knha",
                            digits = 3)
```

```
## Warning: Redundant predictors dropped from the model.
```

```
meta.anal.select.model.lower
```



```
##
## Mixed-Effects Model (k = 157; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.025 (SE = 0.004)
## tau (square root of estimated tau^2 value):             0.159
## I^2 (residual heterogeneity / unaccounted variability): 94.08%
## H^2 (unaccounted variability / sampling variability):   16.89
## R^2 (amount of heterogeneity accounted for):            26.36%
##
## Test for Residual Heterogeneity:
## QE(df = 153) = 649.298, p-val < .001
##
## Test of Moderators (coefficients 2:4):
## F(df1 = 3, df2 = 153) = 11.163, p-val < .001
##
## Model Results:
##
##                    estimate     se    tval   df    pval   ci.lb   ci.ub
## intrcpt               0.083  0.034   2.404  153   0.017   0.015   0.151    *
```

```
## GDP_control           0.149  0.036   4.098  153  <.001   0.077   0.220  ***
## inflation_control    -0.229  0.058  -3.974  153  <.001  -0.343  -0.115  ***
## PCC_se               -0.124  0.248  -0.500  153  0.618  -0.615   0.366
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Meta analysis test based on journal rank

```r
meta.anal.journ <- rma(yi = PCC,
                 sei = PCC_se,
                 data = FDII,
                 method = "REML",
                 mods = ~ journal_rank,
                 test = "knha")
```

```
## Warning: 71 studies with NAs omitted from model fitting.
```

```r
meta.anal.journ
```

```
##
## Mixed-Effects Model (k = 545; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0263 (SE = 0.0020)
## tau (square root of estimated tau^2 value):             0.1620
## I^2 (residual heterogeneity / unaccounted variability): 91.75%
## H^2 (unaccounted variability / sampling variability):   12.12
## R^2 (amount of heterogeneity accounted for):            0.85%
##
## Test for Residual Heterogeneity:
## QE(df = 543) = 3150.5280, p-val < .0001
##
## Test of Moderators (coefficient 2):
## F(df1 = 1, df2 = 543) = 3.5667, p-val = 0.0595
##
## Model Results:
##
```

```
##                estimate      se     tval   df    pval    ci.lb   ci.ub
## intrcpt          0.1146  0.0162   7.0765  543  <.0001   0.0828  0.1464  ***
## journal_rank   -0.0000  0.0000  -1.8886  543  0.0595  -0.0000  0.0000    .
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Meta analysis test based on log journal rank

```r
meta.anal.lnjourn <- rma(yi = PCC,
                       sei = PCC_se,
                       data = FDII,
                       method = "ML",
                       mods = ~ log_journal_rank,
                       test = "knha")
```

```
## Warning: 71 studies with NAs omitted from model fitting.
```

```r
meta.anal.lnjourn
```

```
##
## Mixed-Effects Model (k = 545; tau^2 estimator: ML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0251 (SE = 0.0019)
## tau (square root of estimated tau^2 value):             0.1585
## I^2 (residual heterogeneity / unaccounted variability): 91.39%
## H^2 (unaccounted variability / sampling variability):   11.62
## R^2 (amount of heterogeneity accounted for):            4.89%
##
## Test for Residual Heterogeneity:
## QE(df = 543) = 3043.2150, p-val < .0001
##
## Test of Moderators (coefficient 2):
## F(df1 = 1, df2 = 543) = 14.8951, p-val = 0.0001
##
## Model Results:
##
```

```
##                    estimate      se     tval   df    pval    ci.lb    ci.ub     
## intrcpt               0.3231  0.0615   5.2575  543  <.0001   0.2024   0.4438  *** 
## log_journal_rank     -0.0270  0.0070  -3.8594  543  0.0001  -0.0407  -0.0132  *** 
## 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Meta analysis test without control variables

```r
options(width = 200)

meta.anal.woCV <- rma(yi = PCC,
                      sei = PCC_se,
                      data = FDII,
                      method = "ML",
                      mods = ~ `study_country` + `estimation_methods` + `level` + `publication_time` + `avg_sample_year` +
                      ↪  `data_type`,
                      test = "t")
```

```
## Warning: Redundant predictors dropped from the model.
```

```r
meta.anal.woCV
```

```
## 
## Mixed-Effects Model (k = 616; tau^2 estimator: ML)
## 
## tau^2 (estimated amount of residual heterogeneity):     0.0112 (SE = 0.0010)
## tau (square root of estimated tau^2 value):             0.1060
## I^2 (residual heterogeneity / unaccounted variability): 89.60%
## H^2 (unaccounted variability / sampling variability):   9.61
## R^2 (amount of heterogeneity accounted for):            68.39%
## 
## Test for Residual Heterogeneity:
## QE(df = 549) = 2051.5311, p-val < .0001
## 
## Test of Moderators (coefficients 2:67):
## F(df1 = 66, df2 = 549) = 10.1044, p-val < .0001
```

```
##
## Model Results:
##
##                                                      estimate      se     tval   df    pval    ci.lb    ci.ub
## intrcpt                                               16.1703  3.0548   5.2934  549  <.0001  10.1698  22.1709  ***
## study_countryArabian                                   0.0535  0.1183   0.4519  549  0.6515  -0.1789   0.2858
## study_countryAsia                                     -0.1362  0.0673  -2.0232  549  0.0435  -0.2683  -0.0040    *
## study_countryBolivia                                  -0.4111  0.1977  -2.0790  549  0.0381  -0.7995  -0.0227    *
## study_countryBrazil                                   -0.1313  0.3107  -0.4226  549  0.6727  -0.7416   0.4790
## study_countryCentral and Eastern Europe               -0.0427  0.0787  -0.5422  549  0.5879  -0.1972   0.1119
## study_countryCentral and Eastern European countries   -0.1775  0.1054  -1.6835  549  0.0929  -0.3845   0.0296    .
## study_countryChile                                    -0.0775  0.1949  -0.3978  549  0.6909  -0.4603   0.3053
## study_countryChina                                    -0.5252  0.1432  -3.6685  549  0.0003  -0.8065  -0.2440  ***
## study_countryColombia                                 -0.3227  0.1975  -1.6341  549  0.1028  -0.7105   0.0652
## study_countryEastern Europe and Central Asia          -0.2528  0.0946  -2.6716  549  0.0078  -0.4387  -0.0669   **
## study_countryEgypt                                    -0.9179  0.1844  -4.9781  549  <.0001  -1.2801  -0.5557  ***
## study_countryEU                                       -0.3471  0.0646  -5.3751  549  <.0001  -0.4740  -0.2203  ***
## study_countryEU;APEC;LAIA                             -0.4166  0.1016  -4.1013  549  <.0001  -0.6161  -0.2171  ***
## study_countryFinland                                  -1.2220  0.2167  -5.6401  549  <.0001  -1.6477  -0.7964  ***
## study_countryGermany                                  -1.0868  0.2222  -4.8910  549  <.0001  -1.5233  -0.6503  ***
## study_countryHungary                                  -1.0902  0.3061  -3.5618  549  0.0004  -1.6914  -0.4890  ***
## study_countryIndia                                     0.0429  0.2242   0.1913  549  0.8484  -0.3976   0.4834
## study_countryIreland                                   0.0468  0.2195   0.2132  549  0.8313  -0.3843   0.4779
## study_countryItaly                                    -0.8021  0.2362  -3.3964  549  0.0007  -1.2660  -0.3382  ***
## study_countryLAC                                      -0.0525  0.0984  -0.5341  549  0.5935  -0.2458   0.1407
## study_countryLatin America                             0.0463  0.0702   0.6597  549  0.5097  -0.0916   0.1842
## study_countryLDC                                      -0.0800  0.0611  -1.3081  549  0.1914  -0.2001   0.0401
## study_countryMalta                                    -0.9254  0.2309  -4.0076  549  <.0001  -1.3790  -0.4718  ***
## study_countryMENA                                      0.3411  0.1611   2.1175  549  0.0347   0.0247   0.6576    *
## study_countryMexico                                   -0.6058  0.1590  -3.8091  549  0.0002  -0.9181  -0.2934  ***
## study_countryNetherlands                              -0.8825  0.1907  -4.6286  549  <.0001  -1.2570  -0.5080  ***
## study_countrynon-OECD                                 -0.1565  0.0694  -2.2557  549  0.0245  -0.2928  -0.0202    *
## study_countryNorway                                   -1.1653  0.2173  -5.3624  549  <.0001  -1.5922  -0.7385  ***
## study_countryOECD                                     -0.2900  0.0671  -4.3226  549  <.0001  -0.4218  -0.1582  ***
## study_countryPakistan                                 -0.5319  0.1848  -2.8786  549  0.0041  -0.8949  -0.1689   **
## study_countrySouth Africa                             -0.3454  0.3821  -0.9038  549  0.3665  -1.0960   0.4052
## study_countrySouth asia                                0.0070  0.0952   0.0737  549  0.9412  -0.1800   0.1940
```

```
## study_countrySouth Korea                               -0.4711  0.1791  -2.6299  549  0.0088  -0.8229  -0.1192   **
## study_countrySpain                                      0.0784  0.2161   0.3629  549  0.7168  -0.3461   0.5030
## study_countrysub-Saharan African                       -0.2843  0.1058  -2.6879  549  0.0074  -0.4920  -0.0765   **
## study_countrySweden                                    -1.3753  0.2038  -6.7488  549  <.0001  -1.7756  -0.9750  ***
## study_countryThailand                                  -0.6953  0.1662  -4.1844  549  <.0001  -1.0217  -0.3689  ***
## study_countryTurkey                                    -0.6166  0.2748  -2.2441  549  0.0252  -1.1563  -0.0769    *
## study_countryUK                                        -0.6723  0.1616  -4.1590  549  <.0001  -0.9898  -0.3548  ***
## study_countryUruguay                                   -0.6669  0.1984  -3.3604  549  0.0008  -1.0567  -0.2771  ***
## study_countryUS                                        -0.4202  0.0945  -4.4488  549  <.0001  -0.6058  -0.2347  ***
## study_countryVietnam                                   -0.2906  0.1159  -2.5073  549  0.0125  -0.5182  -0.0629    *
## study_countryworldwide                                 -0.1578  0.0593  -2.6632  549  0.0080  -0.2742  -0.0414   **
## estimation_methods3SLS                                 -0.4015  0.2441  -1.6446  549  0.1006  -0.8811   0.0780
## estimation_methodsARDL                                  0.2448  0.2551   0.9597  549  0.3376  -0.2562   0.7459
## estimation_methodsBetween effects                       0.0420  0.0946   0.4437  549  0.6574  -0.1438   0.2277
## estimation_methodsCCE                                  -0.1239  0.0663  -1.8685  549  0.0622  -0.2541   0.0063    .
## estimation_methodsDOLS                                 -0.0329  0.0532  -0.6176  549  0.5371  -0.1375   0.0717
## estimation_methodsECM                                  -0.2344  0.1067  -2.1976  549  0.0284  -0.4440  -0.0249    *
## estimation_methodsFM-OLS                                0.0985  0.0937   1.0513  549  0.2936  -0.0856   0.2827
## estimation_methodsGLS                                   0.3277  0.0897   3.6523  549  0.0003   0.1515   0.5040  ***
## estimation_methodsGMM                                  -0.0012  0.0339  -0.0342  549  0.9728  -0.0677   0.0654
## estimation_methodsIV                                   -0.0751  0.0431  -1.7408  549  0.0823  -0.1598   0.0096    .
## estimation_methodsJohansen's cointegration test        -0.0344  0.0629  -0.5473  549  0.5844  -0.1580   0.0891
## estimation_methodsLIML                                  0.0431  0.0704   0.6119  549  0.5408  -0.0953   0.1815
## estimation_methodsLSDV                                 -0.0984  0.1529  -0.6436  549  0.5201  -0.3987   0.2019
## estimation_methodsOLS                                   0.0090  0.0349   0.2580  549  0.7965  -0.0596   0.0776
## estimation_methodsPanel                                -0.0227  0.0308  -0.7360  549  0.4620  -0.0832   0.0379
## estimation_methodsParks                                 0.4284  0.1031   4.1570  549  <.0001   0.2260   0.6309  ***
## estimation_methodsRandom effect                         0.0645  0.0481   1.3410  549  0.1805  -0.0300   0.1590
## estimation_methodsSUR                                  -0.2325  0.2605  -0.8923  549  0.3726  -0.7442   0.2793
## estimation_methodsSURE                                  0.0673  0.0587   1.1460  549  0.2523  -0.0480   0.1825
## level                                                  -0.1944  0.0742  -2.6208  549  0.0090  -0.3401  -0.0487   **
## publication_time                                       -0.0087  0.0025  -3.4919  549  0.0005  -0.0136  -0.0038  ***
## avg_sample_year                                         0.0012  0.0018   0.6372  549  0.5243  -0.0024   0.0048
## data_type                                              -0.0703  0.0287  -2.4529  549  0.0145  -0.1266  -0.0140    *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Meta analysis test with control variables only

```
FDII$FDI_exports <- FDII$`FDI*exports`

meta.anal.cont <- rma(yi = PCC,
                      sei = PCC_se,
                      data = FDII,
                      method = "ML",
                      mods = ~ `GDP` + `GDPpc` + `trade` + `exports` + `openness` + `GNI` + `export_incentives` + `GDPgr` +
                      ↪  `education_secondary_school_enrollment` + `gov_quality` + `inflation` + `unemployment` + `GDPpc^2` +
                      ↪  `export_growth` + `gov_exp` + `education_middle_school_enrollment` + `popgr` + `financial_development`
                      ↪  + `labor_productivity` + `Country_size` + `import` + `GDP^2` + `FDI_exports` + `relative_productivity`,
                      test = "t")
```

```
## Warning: Redundant predictors dropped from the model.
```

```
meta.anal.cont
```



```
##
## Mixed-Effects Model (k = 616; tau^2 estimator: ML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0270 (SE = 0.0020)
## tau (square root of estimated tau^2 value):             0.1642
## I^2 (residual heterogeneity / unaccounted variability): 95.22%
## H^2 (unaccounted variability / sampling variability):   20.93
## R^2 (amount of heterogeneity accounted for):            24.18%
##
## Test for Residual Heterogeneity:
## QE(df = 592) = 3225.2929, p-val < .0001
##
## Test of Moderators (coefficients 2:24):
## F(df1 = 23, df2 = 592) = 5.3464, p-val < .0001
##
## Model Results:
##
##                                        estimate      se     tval   df    pval    ci.lb    ci.ub
## intrcpt                                  0.0217  0.0191   1.1377  592  0.2557  -0.0157   0.0591
```

```
## GDP                                     0.0600  0.0331   1.8110  592  0.0706  -0.0051   0.1251    .
## GDPpc                                   0.0163  0.0224   0.7286  592  0.4665  -0.0277   0.0603
## trade                                   0.0473  0.0216   2.1880  592  0.0291   0.0048   0.0898    *
## exports                                 0.0474  0.0416   1.1390  592  0.2552  -0.0343   0.1291
## openness                                0.0754  0.0317   2.3773  592  0.0178   0.0131   0.1377    *
## GNI                                    -0.0938  0.1843  -0.5088  592  0.6111  -0.4557   0.2682
## export_incentives                      -0.1571  0.1444  -1.0882  592  0.2770  -0.4407   0.1265
## GDPgr                                   0.0317  0.0351   0.9010  592  0.3679  -0.0374   0.1007
## education_secondary_school_enrollment   0.0020  0.0226   0.0904  592  0.9280  -0.0424   0.0465
## gov_quality                             0.1486  0.0634   2.3439  592  0.0194   0.0241   0.2731    *
## inflation                              -0.0502  0.0337  -1.4895  592  0.1369  -0.1165   0.0160
## unemployment                           -0.0207  0.0431  -0.4803  592  0.6312  -0.1054   0.0640
## `GDPpc^2`                               0.1696  0.0321   5.2902  592  <.0001   0.1066   0.2325  ***
## gov_exp                                 0.0174  0.0286   0.6100  592  0.5421  -0.0387   0.0735
## education_middle_school_enrollment      0.0838  0.0941   0.8907  592  0.3735  -0.1010   0.2686
## popgr                                   0.1921  0.0421   4.5642  592  <.0001   0.1094   0.2748  ***
## financial_development                  -0.3388  0.0975  -3.4727  592  0.0006  -0.5303  -0.1472  ***
## labor_productivity                      0.0191  0.1245   0.1538  592  0.8778  -0.2253   0.2636
## Country_size                            0.0454  0.1197   0.3790  592  0.7048  -0.1897   0.2805
## import                                 -0.1818  0.0724  -2.5101  592  0.0123  -0.3240  -0.0395    *
## `GDP^2`                                 0.0337  0.1138   0.2965  592  0.7670  -0.1897   0.2572
## FDI_exports                            -0.1400  0.1621  -0.8641  592  0.3879  -0.4583   0.1782
## relative_productivity                   0.0756  0.1801   0.4200  592  0.6746  -0.2781   0.4294
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Meta analysis all

```r
meta.anal.adv <- rma(yi = PCC,
                     sei = PCC_se,
                     data = FDII,
                     method = "ML",
                     mods = ~ `GDP` + `GDPpc` + `trade` + `exports` + `openness` + `GNI` + `export_incentives` + `GDPgr` +
                      ↪  `education_secondary_school_enrollment` + `gov_quality` + `inflation` + `unemployment` + `GDPpc^2` +
                      ↪  `export_growth` + `gov_exp` + `education_middle_school_enrollment` + `popgr` + `financial_development` +
                      ↪  `labor_productivity` + `Country_size` + `import` + `GDP^2` + `FDI_exports` + `relative_productivity` +
                      ↪  study_country + estimation_methods + study_country + estimation_methods + level + publication_time +
                      ↪  avg_sample_year + data_type,
```

```r
                    test = "t")
```

```
## Warning: Redundant predictors dropped from the model.
```

```r
meta.anal.adv
```

```
##
## Mixed-Effects Model (k = 616; tau^2 estimator: ML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0071 (SE = 0.0007)
## tau (square root of estimated tau^2 value):             0.0842
## I^2 (residual heterogeneity / unaccounted variability): 84.16%
## H^2 (unaccounted variability / sampling variability):   6.31
## R^2 (amount of heterogeneity accounted for):            80.05%
##
## Test for Residual Heterogeneity:
## QE(df = 527) = 1653.6094, p-val < .0001
##
## Test of Moderators (coefficients 2:89):
## F(df1 = 88, df2 = 527) = 11.1984, p-val < .0001
##
## Model Results:
##
##                                             estimate      se     tval   df    pval    ci.lb    ci.ub
## intrcpt                                      10.6245  4.1476   2.5616  527  0.0107   2.4766  18.7724    *
## GDP                                           0.1225  0.0389   3.1494  527  0.0017   0.0461   0.1990   **
## GDPpc                                        -0.0359  0.0210  -1.7038  527  0.0890  -0.0772   0.0055    .
## trade                                        -0.0129  0.0245  -0.5274  527  0.5981  -0.0611   0.0353
## exports                                       0.0739  0.0341   2.1680  527  0.0306   0.0069   0.1409    *
## openness                                     -0.0058  0.0371  -0.1552  527  0.8767  -0.0787   0.0672
## GNI                                           0.1672  0.1835   0.9109  527  0.3628  -0.1933   0.5276
## export_incentives                             0.1590  0.1750   0.9086  527  0.3640  -0.1847   0.5027
## GDPgr                                        -0.0869  0.0346  -2.5122  527  0.0123  -0.1549  -0.0190    *
## education_secondary_school_enrollment        -0.0001  0.0286  -0.0019  527  0.9985  -0.0563   0.0562
## gov_quality                                   0.1608  0.0705   2.2800  527  0.0230   0.0223   0.2993    *
## inflation                                     0.0388  0.0299   1.2955  527  0.1957  -0.0200   0.0976
## unemployment                                  0.0530  0.0714   0.7423  527  0.4583  -0.0873   0.1934
```

```
## `GDPpc^2`                                          0.1495  0.0373   4.0061  527  <.0001   0.0762   0.2229  ***
## gov_exp                                            0.0062  0.0310   0.2007  527  0.8410  -0.0548   0.0672
## education_middle_school_enrollment                -0.3080  0.0962  -3.2030  527  0.0014  -0.4970  -0.1191   **
## popgr                                              0.1148  0.0483   2.3742  527  0.0179   0.0198   0.2098    *
## financial_development                             -0.2685  0.1200  -2.2381  527  0.0256  -0.5042  -0.0328    *
## labor_productivity                                -0.0691  0.0784  -0.8813  527  0.3786  -0.2231   0.0849
## Country_size                                      -0.3127  0.1171  -2.6712  527  0.0078  -0.5426  -0.0827   **
## import                                            -0.3609  0.0620  -5.8217  527  <.0001  -0.4827  -0.2391  ***
## `GDP^2`                                           -0.2353  0.1248  -1.8862  527  0.0598  -0.4805   0.0098    .
## FDI_exports                                        0.4831  0.1782   2.7106  527  0.0069   0.1330   0.8332   **
## relative_productivity                              0.0634  0.1118   0.5675  527  0.5706  -0.1562   0.2831
## study_countryArabian                               0.0496  0.1236   0.4016  527  0.6882  -0.1931   0.2923
## study_countryAsia                                 -0.3346  0.0833  -4.0183  527  <.0001  -0.4982  -0.1710  ***
## study_countryBolivia                              -0.3080  0.1920  -1.6044  527  0.1092  -0.6851   0.0691
## study_countryBrazil                                0.1027  0.3047   0.3369  527  0.7363  -0.4960   0.7013
## study_countryCentral and Eastern Europe           -0.2246  0.1161  -1.9348  527  0.0535  -0.4527   0.0034    .
## study_countryChile                                 0.0249  0.1890   0.1319  527  0.8951  -0.3464   0.3962
## study_countryChina                                -0.5327  0.1496  -3.5609  527  0.0004  -0.8266  -0.2388  ***
## study_countryColombia                             -0.2203  0.1917  -1.1495  527  0.2509  -0.5968   0.1562
## study_countryEastern Europe and Central Asia      -0.6014  0.1370  -4.3888  527  <.0001  -0.8706  -0.3322  ***
## study_countryEgypt                                -0.7331  0.1856  -3.9489  527  <.0001  -1.0978  -0.3684  ***
## study_countryEU                                   -0.3831  0.0740  -5.1771  527  <.0001  -0.5285  -0.2377  ***
## study_countryEU;APEC;LAIA                         -0.5797  0.1661  -3.4907  527  0.0005  -0.9060  -0.2535  ***
## study_countryFinland                              -1.1500  0.2119  -5.4279  527  <.0001  -1.5662  -0.7338  ***
## study_countryGermany                              -1.0199  0.2174  -4.6919  527  <.0001  -1.4469  -0.5929  ***
## study_countryHungary                              -0.9174  0.3018  -3.0401  527  0.0025  -1.5102  -0.3246   **
## study_countryIndia                                 0.2778  0.2205   1.2599  527  0.2083  -0.1554   0.7110
## study_countryIreland                               0.1254  0.2149   0.5834  527  0.5598  -0.2968   0.5476
## study_countryItaly                                -0.7259  0.2319  -3.1302  527  0.0018  -1.1815  -0.2704   **
## study_countryLAC                                  -0.1163  0.1154  -1.0081  527  0.3139  -0.3429   0.1103
## study_countryLatin America                        -0.0240  0.0826  -0.2904  527  0.7716  -0.1861   0.1382
## study_countryLDC                                  -0.1338  0.0757  -1.7664  527  0.0779  -0.2826   0.0150    .
## study_countryMalta                                -0.8483  0.2265  -3.7455  527  0.0002  -1.2932  -0.4034  ***
## study_countryMENA                                  0.4001  0.1589   2.5177  527  0.0121   0.0879   0.7123    *
## study_countryMexico                               -0.4890  0.1653  -2.9573  527  0.0032  -0.8138  -0.1642   **
## study_countryNetherlands                          -1.0222  0.1984  -5.1535  527  <.0001  -1.4119  -0.6326  ***
## study_countrynon-OECD                             -0.1422  0.0775  -1.8350  527  0.0671  -0.2944   0.0100    .
```

```
## study_countryNorway                               -1.1019  0.2122  -5.1927  527  <.0001  -1.5188  -0.6850  ***
## study_countryOECD                                 -0.3200  0.0750  -4.2661  527  <.0001  -0.4673  -0.1726  ***
## study_countryPakistan                             -0.5030  0.1836  -2.7405  527  0.0063  -0.8636  -0.1424   **
## study_countrySouth Africa                         -0.3872  0.3797  -1.0198  527  0.3083  -1.1331   0.3587
## study_countrySouth asia                           -0.0175  0.1094  -0.1600  527  0.8730  -0.2325   0.1975
## study_countrySouth Korea                          -0.0782  0.1842  -0.4247  527  0.6712  -0.4401   0.2836
## study_countrySpain                                 0.1565  0.2113   0.7407  527  0.4592  -0.2586   0.5716
## study_countrysub-Saharan African                  -0.2965  0.1031  -2.8761  527  0.0042  -0.4991  -0.0940   **
## study_countrySweden                               -1.3067  0.1984  -6.5845  527  <.0001  -1.6965  -0.9168  ***
## study_countryThailand                             -0.6454  0.1687  -3.8261  527  0.0001  -0.9768  -0.3140  ***
## study_countryTurkey                               -0.5596  0.2760  -2.0278  527  0.0431  -1.1017  -0.0175    *
## study_countryUK                                   -0.5825  0.1647  -3.5361  527  0.0004  -0.9060  -0.2589  ***
## study_countryUruguay                              -0.5645  0.1927  -2.9295  527  0.0035  -0.9430  -0.1859   **
## study_countryUS                                   -0.6285  0.1097  -5.7313  527  <.0001  -0.8440  -0.4131  ***
## study_countryVietnam                              -0.4066  0.1295  -3.1405  527  0.0018  -0.6609  -0.1523   **
## study_countryworldwide                            -0.1775  0.0715  -2.4839  527  0.0133  -0.3179  -0.0371    *
## estimation_methods3SLS                            -0.2447  0.2412  -1.0145  527  0.3108  -0.7185   0.2291
## estimation_methodsARDL                             0.3819  0.2416   1.5807  527  0.1146  -0.0927   0.8566
## estimation_methodsBetween effects                  0.0483  0.0848   0.5696  527  0.5692  -0.1183   0.2149
## estimation_methodsCCE                              0.1606  0.0846   1.8983  527  0.0582  -0.0056   0.3268    .
## estimation_methodsDOLS                             0.1095  0.0543   2.0145  527  0.0445   0.0027   0.2162    *
## estimation_methodsECM                              0.0683  0.1058   0.6457  527  0.5188  -0.1396   0.2762
## estimation_methodsFM-OLS                           0.2352  0.0873   2.6936  527  0.0073   0.0637   0.4068   **
## estimation_methodsGLS                              0.5458  0.0932   5.8574  527  <.0001   0.3627   0.7288  ***
## estimation_methodsGMM                              0.0025  0.0302   0.0835  527  0.9335  -0.0568   0.0618
## estimation_methodsIV                               0.0481  0.0609   0.7893  527  0.4303  -0.0716   0.1678
## estimation_methodsJohansen's cointegration test    0.1354  0.0633   2.1376  527  0.0330   0.0110   0.2597    *
## estimation_methodsLIML                            -0.0008  0.0632  -0.0123  527  0.9902  -0.1248   0.1233
## estimation_methodsLSDV                            -0.1326  0.1421  -0.9331  527  0.3512  -0.4117   0.1465
## estimation_methodsOLS                              0.0452  0.0315   1.4351  527  0.1518  -0.0167   0.1071
## estimation_methodsPanel                           -0.0149  0.0276  -0.5387  527  0.5903  -0.0692   0.0394
## estimation_methodsParks                            0.3959  0.0974   4.0637  527  <.0001   0.2045   0.5873  ***
## estimation_methodsRandom effect                    0.0527  0.0423   1.2452  527  0.2136  -0.0304   0.1357
## estimation_methodsSUR                             -0.0756  0.2578  -0.2934  527  0.7693  -0.5820   0.4307
## estimation_methodsSURE                             0.1370  0.0560   2.4470  527  0.0147   0.0270   0.2470    *
## level                                             -0.1210  0.0716  -1.6886  527  0.0919  -0.2617   0.0198    .
## publication_time                                  -0.0126  0.0027  -4.7062  527  <.0001  -0.0179  -0.0074  ***
```

```
## avg_sample_year                                        0.0079  0.0024   3.2385  527  0.0013   0.0031   0.0126   **
## data_type                                             -0.1654  0.0325  -5.0866  527  <.0001  -0.2292  -0.1015  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA test for best model (to test against overfitting)

```
# meta.permutation <- permutest(meta.anal.adv)     # This takes c.40 minutes to run. It does not need to be run each time

anova(meta.anal.cont, meta.anal.adv)
```

```
##
##         df       AIC       BIC      AICc   logLik      LRT   pval        QE  tau^2      R^2
## Full    90 -500.2877 -102.1955 -469.0877 340.1438                 1653.6094 0.0071
## Reduced 25 -112.0900   -1.5089 -109.8866  81.0450 518.1977 <.0001 3225.2929 0.0270 73.6848%
```

Further funnel plot asymmetry tests

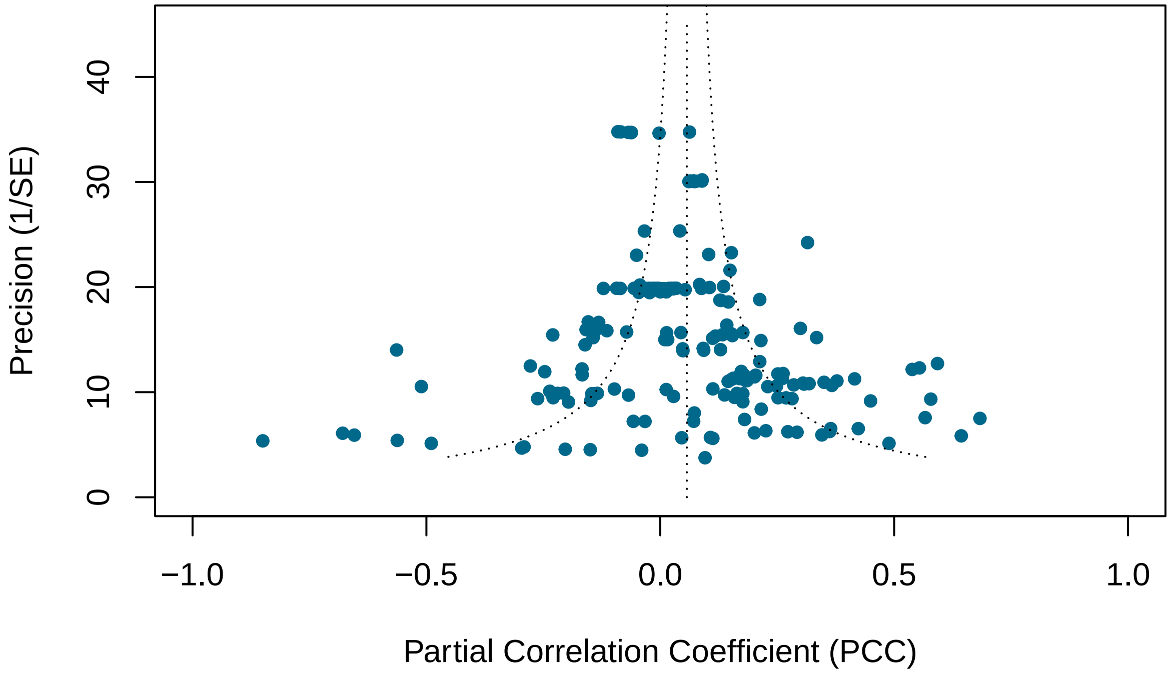
Run the Eggers test for those with endogeneity controls

The tests used that allow for endogeneity controls are: - 2SLS \ - 3SLS \ - IV \ - ARDL \ - LIML \ - LSDV \ - FM-OLS \ - CCE \ - DOLS \ - ECM \ - GMM \

```r
options(width = 70)

FDII_endo <- FDII %>%
  filter(estimation_methods %in% c("2SLS", "3SLS", "IV", "Johansen's cointegration
  ↪  test", "ARDL", "LIML", "LSDV", "FM-OLS", "CCE", "DOLS", "ECM", "GMM"))
m.gen.endo <- metagen(TE = FDII_endo$PCC,
                      seTE = FDII_endo$PCC_se,
                      studlab = FDII_endo$citation,
                      data = FDII_endo,
                      sm = "SMD",
                      comb.fixed = FALSE,
                      comb.random = TRUE,
                      method.tau = "REML",
                      hakn = TRUE,
                      title = "FDI Impacts on Income Inequality (with endogeneity
                       ↪  adjustments)")

eggers.test(m.gen.endo)
```

```
## Eggers' test of the intercept
## =============================
##
##  intercept       95% CI     t              p
##      0.829 0.44 - 1.22 4.13 0.00005361478
##
## Eggers' test indicates the presence of funnel plot asymmetry.
```

```r
PCC_Funnel_Graph_endo <- metafor::funnel(m.gen.endo,
                                    xlim = c(-1, 1),
                                    ylim = rev(c(45, 0)),
                                    studlab = FALSE,
                                    xlab = "Partial Correlation Coefficient (PCC)",
                                    ylab = "Precision (1/SE)",
                                    yaxis = "invse",
                                    col = "deepskyblue4",
                                    bg = "deepskyblue4",
                                    cex = 0.8)
```



We see that when we run a funnel plot asymmetry test on only those which have accounted for endogeneity issues, we still witness funnel ploy asymmetry.

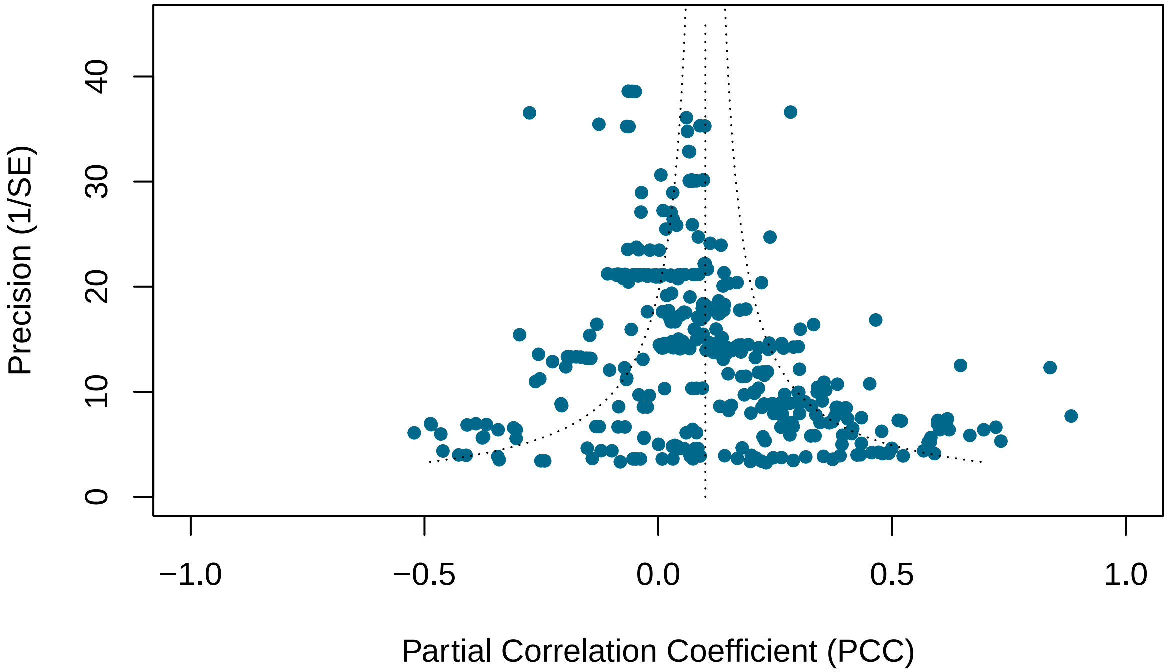
Run the Eggers test for those without endogeneity controls

```
FDII_no_test <- FDII %>%
  filter(estimation_methods %in% c("Between effects", "GLS", "SUR", "SURE", "Random
  ↪  effect", "Probit", "Parks", "Panel", "OLS"))
m.gen.no.test <- metagen(TE = FDII_no_test$PCC,
                         seTE = FDII_no_test$PCC_se,
                         studlab = FDII_no_test$citation,
                         data = FDII_no_test,
                         sm = "SMD",
                         comb.fixed = FALSE,
                         comb.random = TRUE,
                         method.tau = "REML",
                         hakn = TRUE,
                         title = "FDI Impacts on Income Inequality (without
                         ↪  endogeneity adjustments)")

eggers.test(m.gen.no.test)
```

```
## Eggers' test of the intercept
## =============================
##
##  intercept      95% CI     t                 p
##      0.903 0.61 - 1.19 6.081 0.000000002796337
##
## Eggers' test indicates the presence of funnel plot asymmetry.
```

```
PCC_Funnel_Graph_no_test <- metafor::funnel(m.gen.no.test,
                                            xlim = c(-1, 1),
                                            ylim = rev(c(45, 0)),
                                            studlab = FALSE,
                                            xlab = "Partial Correlation Coefficient (PCC)",
```

```
                                ylab = "Precision (1/SE)",
                                yaxis = "invse",
                                col = "deepskyblue4",
                                bg = "deepskyblue4",
                                cex = 0.8,
                                title = "FDI Impacts on Income Inequality
                                ↪  (without endogeneity adjustments)")
```



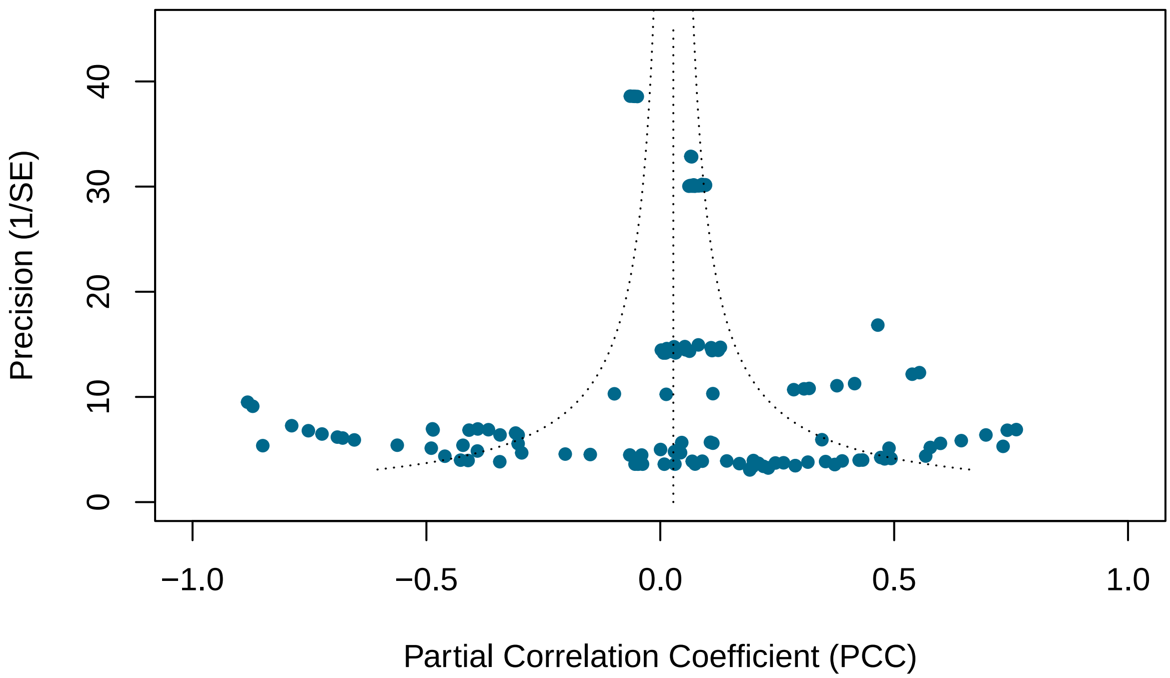
As expected, we again witness funnel plot asymmetry on those who have not accounted for endogeneity

FAT - single-country

```
FDII_single <- subset(FDII, level == 1)

m.gen.single <- metagen(TE = FDII_single$PCC,
                        seTE = FDII_single$PCC_se,
                        studlab = FDII_single$citation,
                        data = FDII_single,
                        sm = "SMD",
                        comb.fixed = FALSE,
                        comb.random = TRUE,
                        method.tau = "REML",
                        hakn = TRUE,
                        title = "FDI Impacts on Income Inequality (single country)")

eggers.test(m.gen.single)
```

```
## Eggers' test of the intercept
## ============================
##
##  intercept       95% CI     t          p
##       0.36 -0.12 - 0.84 1.477 0.1418879
##
## Eggers' test does not indicate the presence of funnel plot asymmetry.
```

```r
PCC_Funnel_Graph_single <- metafor::funnel(m.gen.single,
                                    xlim = c(-1, 1),
                                    ylim = rev(c(45, 0)),
                                    studlab = FALSE,
                                    xlab = "Partial Correlation Coefficient (PCC)",
                                    ylab = "Precision (1/SE)",
                                    yaxis = "invse",
                                    col = "deepskyblue4",
                                    bg = "deepskyblue4",
                                    cex = 0.8)
```



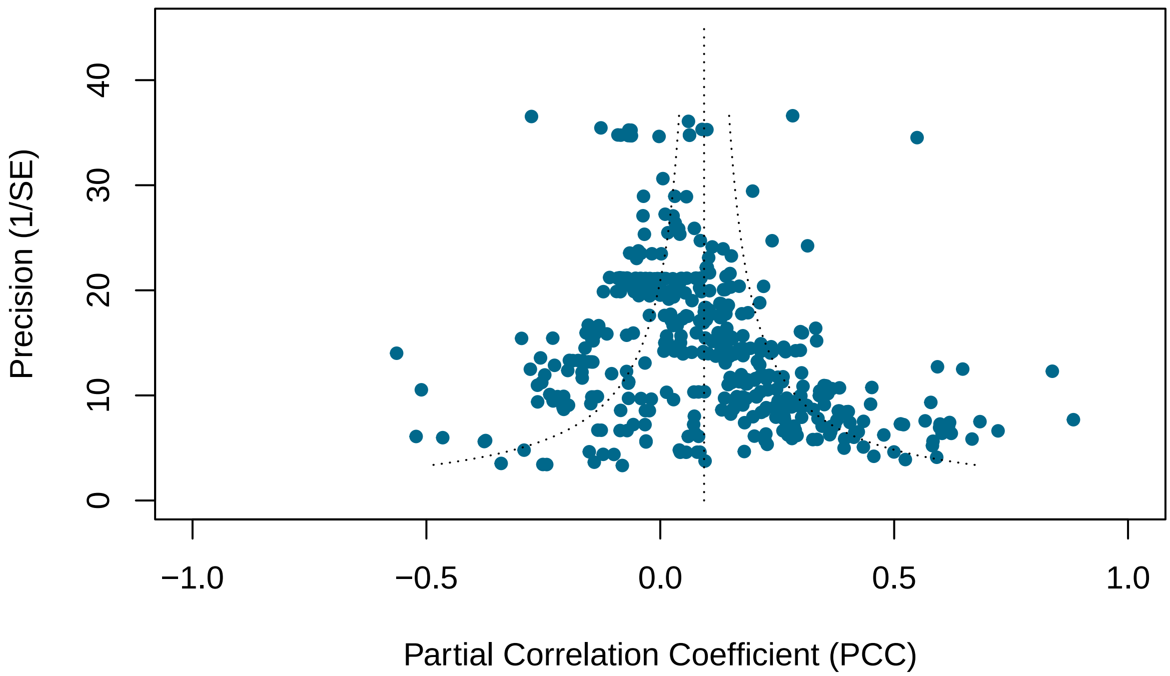
FAT - multiple-country

```r
FDII_multi <- subset(FDII, level != 1)

m.gen.multi <- metagen(TE = FDII_multi$PCC,
                       seTE = FDII_multi$PCC_se,
                       studlab = FDII_multi$citation,
                       data = FDII_multi,
                       sm = "SMD",
                       comb.fixed = FALSE,
                       comb.random = TRUE,
                       method.tau = "REML",
                       hakn = TRUE,
                       title = "FDI Impacts on Income Inequality (multi country)")

eggers.test(m.gen.multi)
```

```
## Eggers' test of the intercept
## ============================
##
##  intercept      95% CI     t                   p
##      1.741 1.24 - 2.24 6.864 0.00000000002149276
##
```

```
## Eggers' test indicates the presence of funnel plot asymmetry.
```

```
PCC_Funnel_Graph_multi <- metafor::funnel(m.gen.multi,
                                          xlim = c(-1, 1),
                                          ylim = rev(c(45, 0)),
                                          studlab = FALSE,
                                          xlab = "Partial Correlation Coefficient
                                          ↪  (PCC)",
                                          ylab = "Precision (1/SE)",
                                          yaxis = "invse",
                                          col = "deepskyblue4",
                                          bg = "deepskyblue4",
                                          cex = 0.8)
```



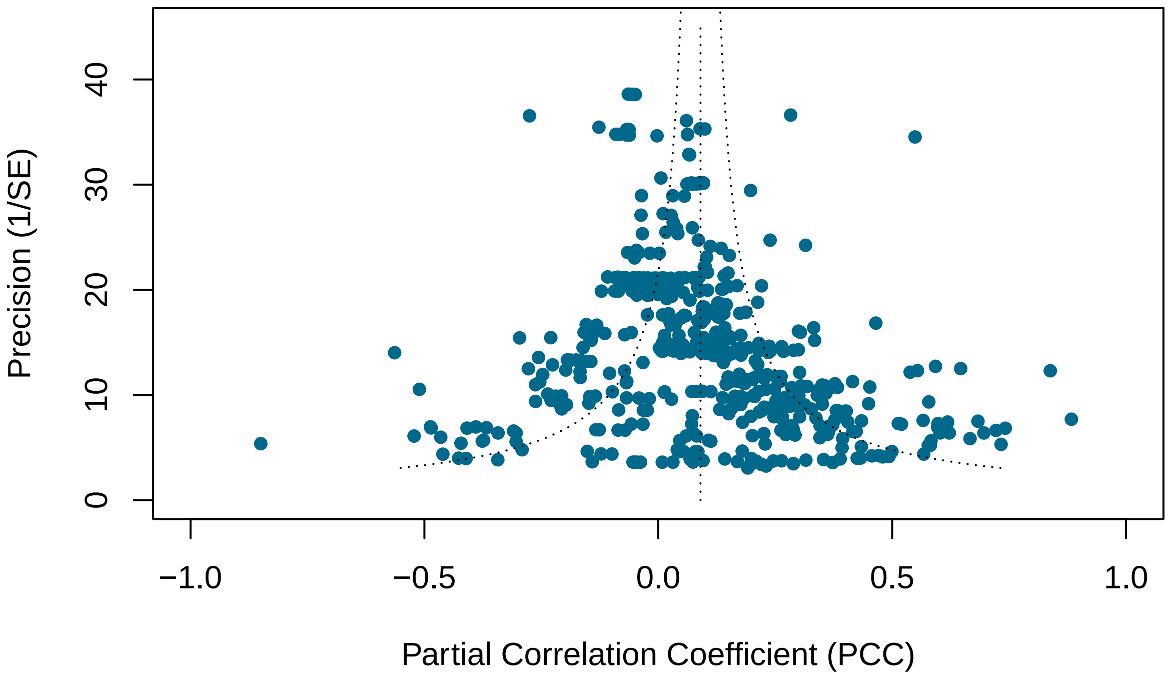
FAT - published papers

```
FDII_published <- subset(FDII, if_published == 1)

m.gen.published <- metagen(TE = FDII_published$PCC,
                        seTE = FDII_published$PCC_se,
                        studlab = FDII_published$citation,
                        data = FDII_published,
                        sm = "SMD",
                        comb.fixed = FALSE,
                        comb.random = TRUE,
                        method.tau = "REML",
                        hakn = TRUE,
                        title = "FDI Impacts on Income Inequality (single country)")

eggers.test(m.gen.published)
```

```
## Eggers' test of the intercept
## ============================
##
##  intercept      95% CI     t                p
```

```
##      0.872 0.58 - 1.17 5.817 0.00000001003444
##
## Eggers' test indicates the presence of funnel plot asymmetry.
```

```
PCC_Funnel_Graph_published <- metafor::funnel(m.gen.published,
                                              xlim = c(-1, 1),
                                              ylim = rev(c(45, 0)),
                                              studlab = FALSE,
                                              xlab = "Partial Correlation Coefficient
                                              ↪  (PCC)",
                                              ylab = "Precision (1/SE)",
                                              yaxis = "invse",
                                              col = "deepskyblue4",
                                              bg = "deepskyblue4",
                                              cex = 0.8)
```



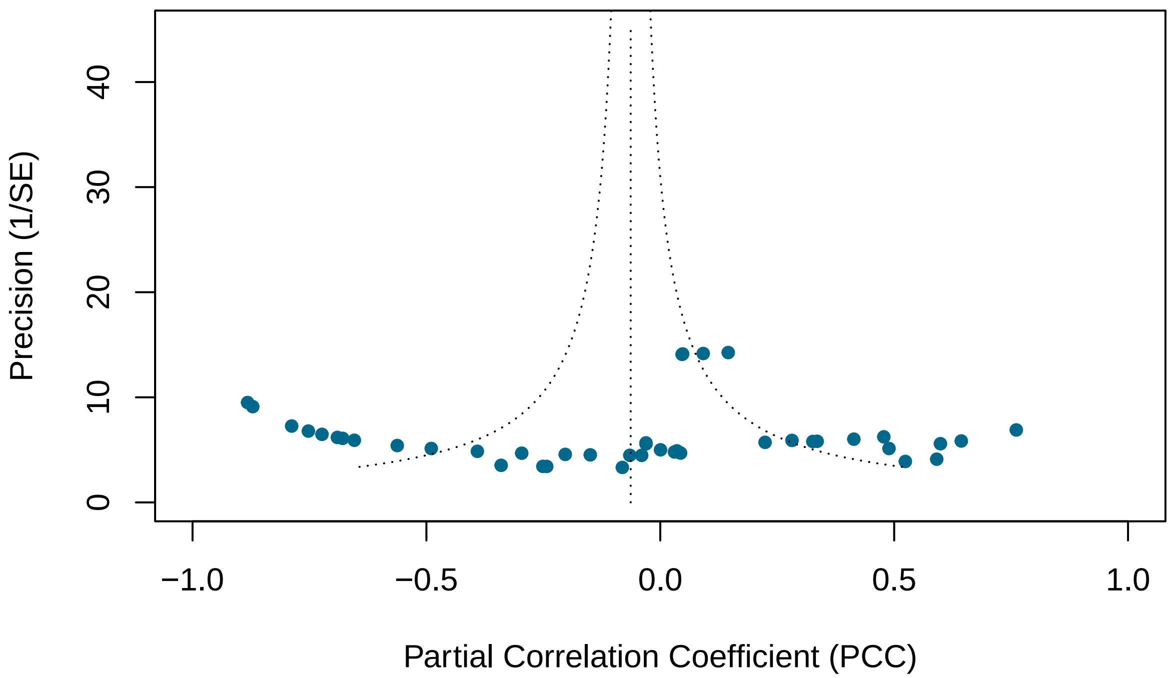
FAT - unpublished papers

```
FDII_not_published <- subset(FDII, if_published == 0)

m.gen.not.published <- metagen(TE = FDII_not_published$PCC,
                           seTE = FDII_not_published$PCC_se,
                           studlab = FDII_not_published$citation,
                           data = FDII_not_published,
                           sm = "SMD",
                           comb.fixed = FALSE,
                           comb.random = TRUE,
                           method.tau = "REML",
                           hakn = TRUE,
                           title = "FDI Impacts on Income Inequality (single
                           ↪  country)")

eggers.test(m.gen.not.published)
```

```
## Eggers' test of the intercept
```

```
## ==============================
##
##  intercept       95% CI      t          p
##     -0.536 -1.45 - 0.38 -1.152 0.2551092
##
## Eggers' test does not indicate the presence of funnel plot asymmetry.
```

```
PCC_Funnel_Graph_not_published <- metafor::funnel(m.gen.not.published,
                                          xlim = c(-1, 1),
                                          ylim = rev(c(45, 0)),
                                          studlab = FALSE,
                                          xlab = "Partial Correlation Coefficient
                                          ↪  (PCC)",
                                          ylab = "Precision (1/SE)",
                                          yaxis = "invse",
                                          col = "deepskyblue4",
                                          bg = "deepskyblue4",
                                          cex = 0.8)
```



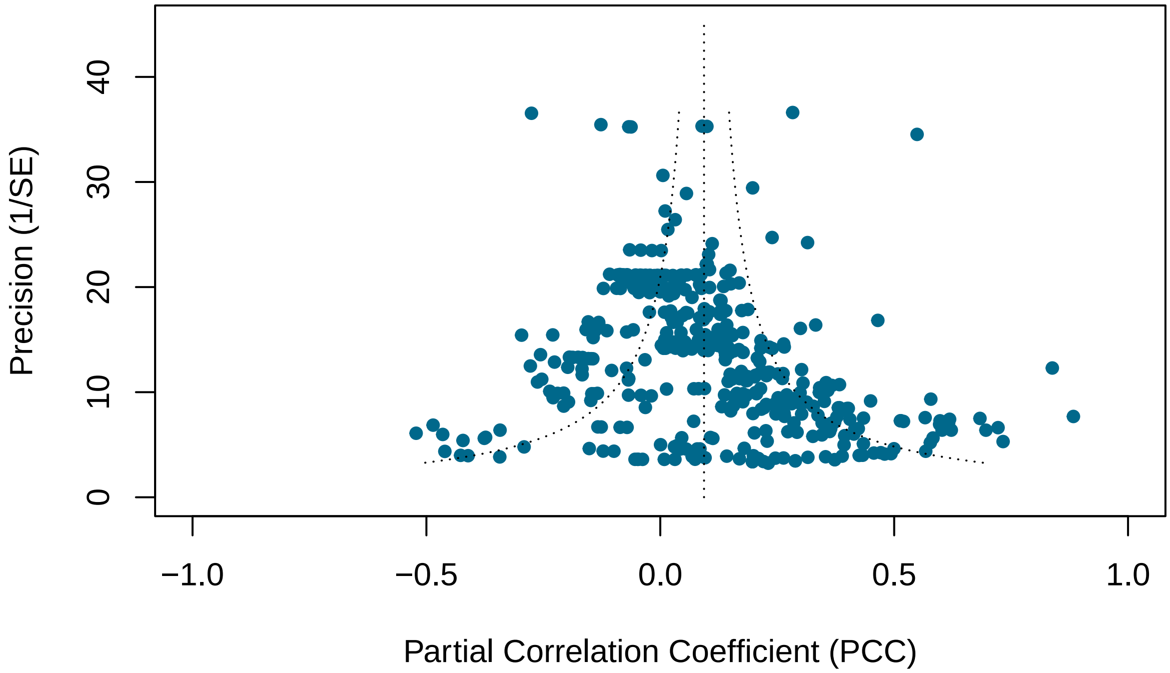
FAT - core tests (from those in the permutest with p < 0.05)

```
FDII_core_tests <- subset(FDII, GDP == 1 |
                GDPpc == 1 |
                GNI == 1 |
                export_incentives == 1 |
                GDPgr == 1 |
                gov_quality == 1 |
                unemployment == 1 |
                `GDPpc^2` == 1 |
                gov_exp == 1 |
                education_middle_school_enrollment == 1 |
                popgr == 1 |
                financial_development == 1 |
                labor_productivity == 1 |
```

```r
                Country_size == 1 |
                import == 1 |
                `GDP^2` == 1)

m.gen.core.tests <- metagen(TE = FDII_core_tests$PCC,
                       seTE = FDII_core_tests$PCC_se,
                       studlab = FDII_core_tests$citation,
                       data = FDII_core_tests,
                       sm = "SMD",
                       comb.fixed = FALSE,
                       comb.random = TRUE,
                       method.tau = "REML",
                       hakn = TRUE,
                       title = "FDI Impacts on Income Inequality (core tests)")

eggers.test(m.gen.core.tests)
```

```
## Eggers' test of the intercept
## =============================
##
##  intercept      95% CI     t                p
##      1.371 0.91 - 1.83 5.806 0.00000001224257
##
## Eggers' test indicates the presence of funnel plot asymmetry.
```

```r
PCC_Funnel_Graph_core_tests <- metafor::funnel(m.gen.core.tests,
                                           xlim = c(-1, 1),
                                           ylim = rev(c(45, 0)),
                                           studlab = FALSE,
                                           xlab = "Partial Correlation Coefficient
                                           ↪  (PCC)",
                                           ylab = "Precision (1/SE)",
                                           yaxis = "invse",
                                           col = "deepskyblue4",
                                           bg = "deepskyblue4",
                                           cex = 0.8)
```



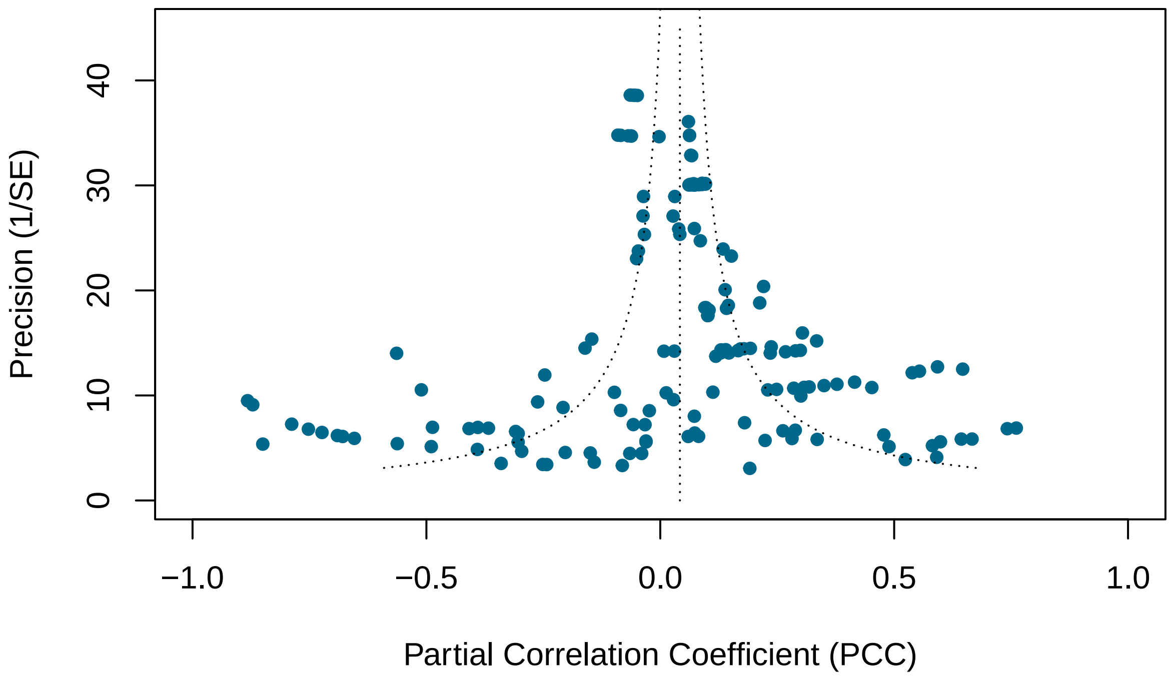
FAT - no core tests

```
FDII_no_core_tests <- subset(FDII, !(GDP == 1 | GDPpc == 1 | GNI == 1 |
↪  export_incentives == 1 | GDPgr == 1 | gov_quality == 1 | unemployment == 1 |
↪  `GDPpc^2` == 1 | gov_exp == 1 | education_middle_school_enrollment == 1 | popgr
↪  == 1 | financial_development == 1 | labor_productivity == 1 | Country_size == 1
↪  | import == 1 | `GDP^2` == 1))

m.gen.no.core.tests <- metagen(TE = FDII_no_core_tests$PCC,
                            seTE = FDII_no_core_tests$PCC_se,
                            studlab = FDII_no_core_tests$citation,
                            data = FDII_no_core_tests,
                            sm = "SMD",
                            comb.fixed = FALSE,
                            comb.random = TRUE,
                            method.tau = "REML",
                            hakn = TRUE,
                            title = "FDI Impacts on Income Inequality (core tests)")

eggers.test(m.gen.no.core.tests)
```

```
## Eggers' test of the intercept
## =============================
##
##  intercept      95% CI      t            p
##      0.694 0.17 - 1.22 2.612 0.009815691
##
## Eggers' test indicates the presence of funnel plot asymmetry.
```

```
PCC_Funnel_Graph_no_core_tests <- metafor::funnel(m.gen.no.core.tests,
                                                   xlim = c(-1, 1),
                                                   ylim = rev(c(45, 0)),
                                                   studlab = FALSE,
                                                   xlab = "Partial Correlation
                                                   ↪  Coefficient (PCC)",
```

```
                                     ylab = "Precision (1/SE)",
                                     yaxis = "invse",
                                     col = "deepskyblue4",
                                     bg = "deepskyblue4",
                                     cex = 0.8)
```

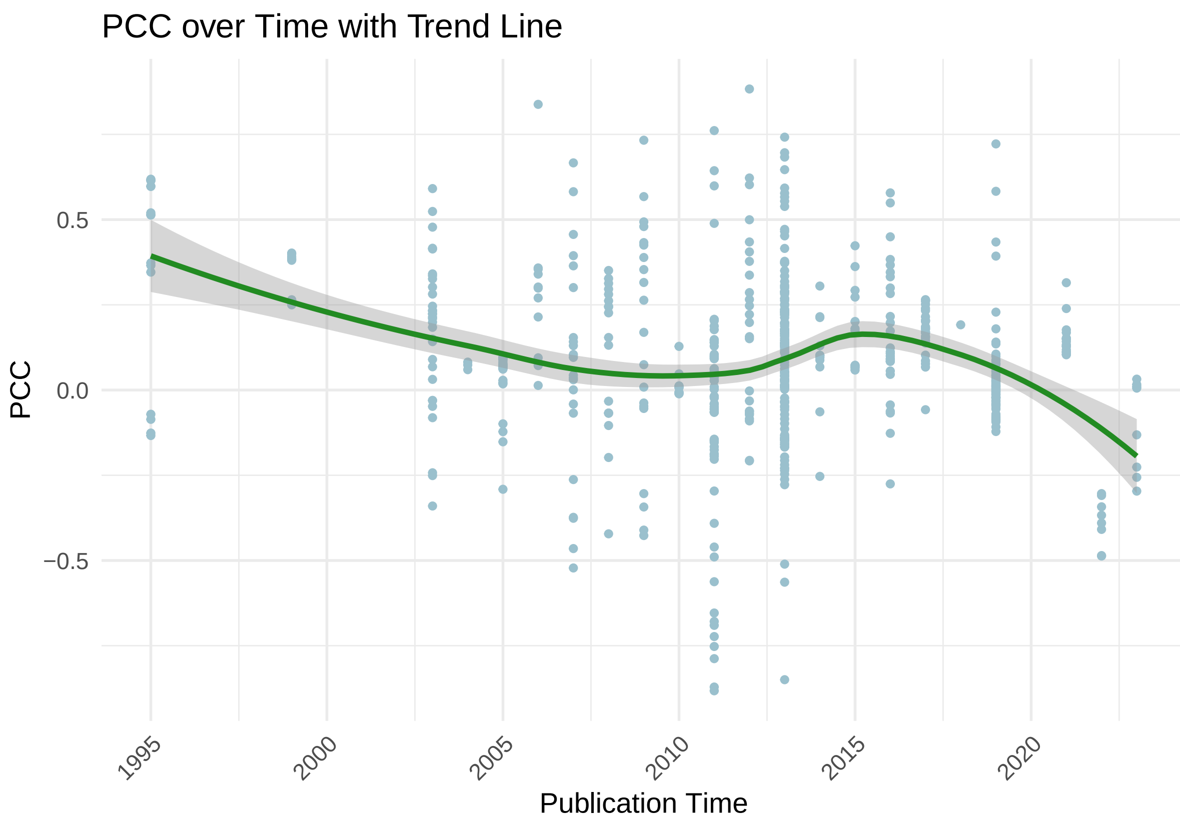


Tests over time

PCC over time (to publication)

```
ggplot(FDII, aes(x = publication_time, y = PCC)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "loess", col = "forestgreen") +
  labs(title = "PCC over Time with Trend Line", x = "Publication Time", y = "PCC") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

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



This is super interesting, because it could be showing changes in the underlying economic phenomenon

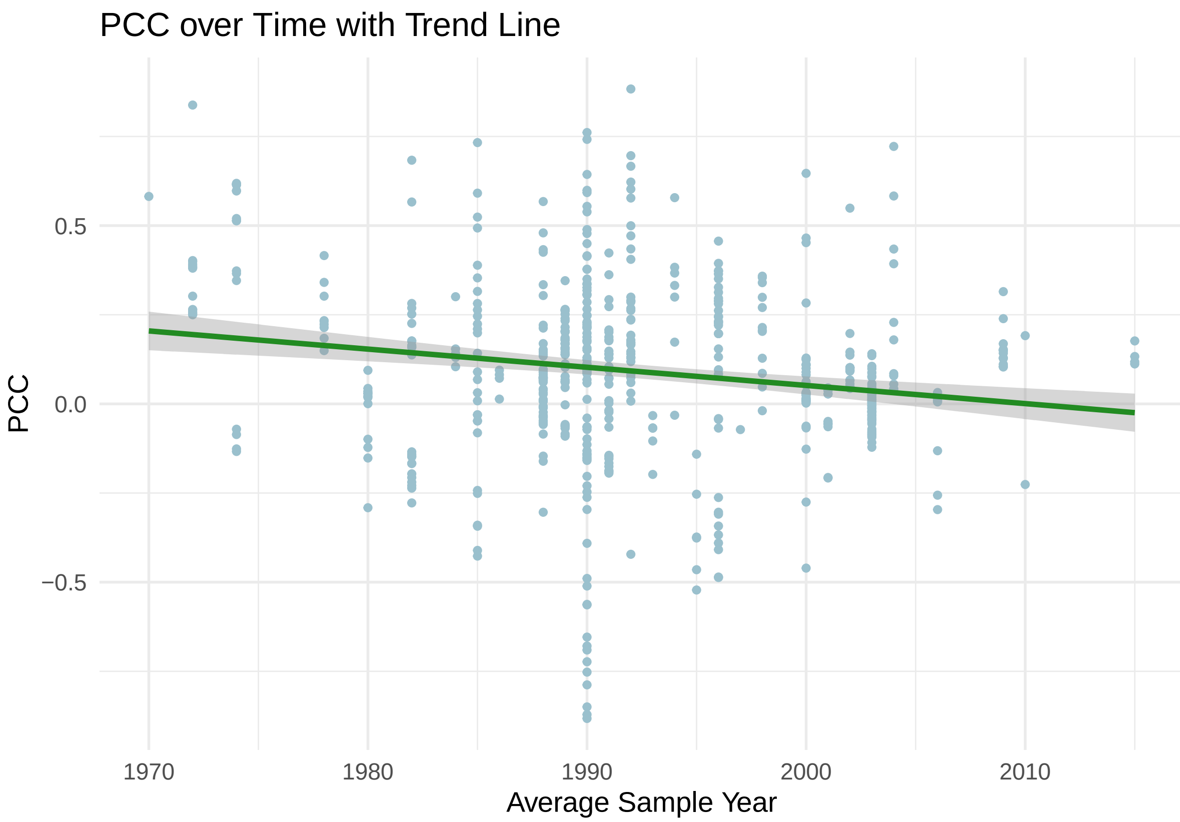
```
summary(lm(FDII$PCC ~ FDII$publication_time))
```

```
##
## Call:
## lm(formula = FDII$PCC ~ FDII$publication_time)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98255 -0.10791  0.00262  0.13298  0.79183
##
## Coefficients:
##                        Estimate Std. Error t value    Pr(>|t|)
## (Intercept)           17.905310   3.322600   5.389 0.000000101 ***
## FDII$publication_time -0.008854   0.001651  -5.362 0.000000117 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2424 on 614 degrees of freedom
## Multiple R-squared:  0.04473,    Adjusted R-squared:  0.04318
## F-statistic: 28.75 on 1 and 614 DF,  p-value: 0.0000001166
```

PCC trend over time (average year of the sample)

```
ggplot(FDII, aes(x = avg_sample_year, y = PCC)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "lm", col = "forestgreen") +
  labs(title = "PCC over Time with Trend Line", x = "Average Sample Year", y =
  ↪  "PCC") +
  theme_minimal()
```

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



We do, as expected

```r
summary(lm(FDII$PCC ~ FDII$avg_sample_year))
```

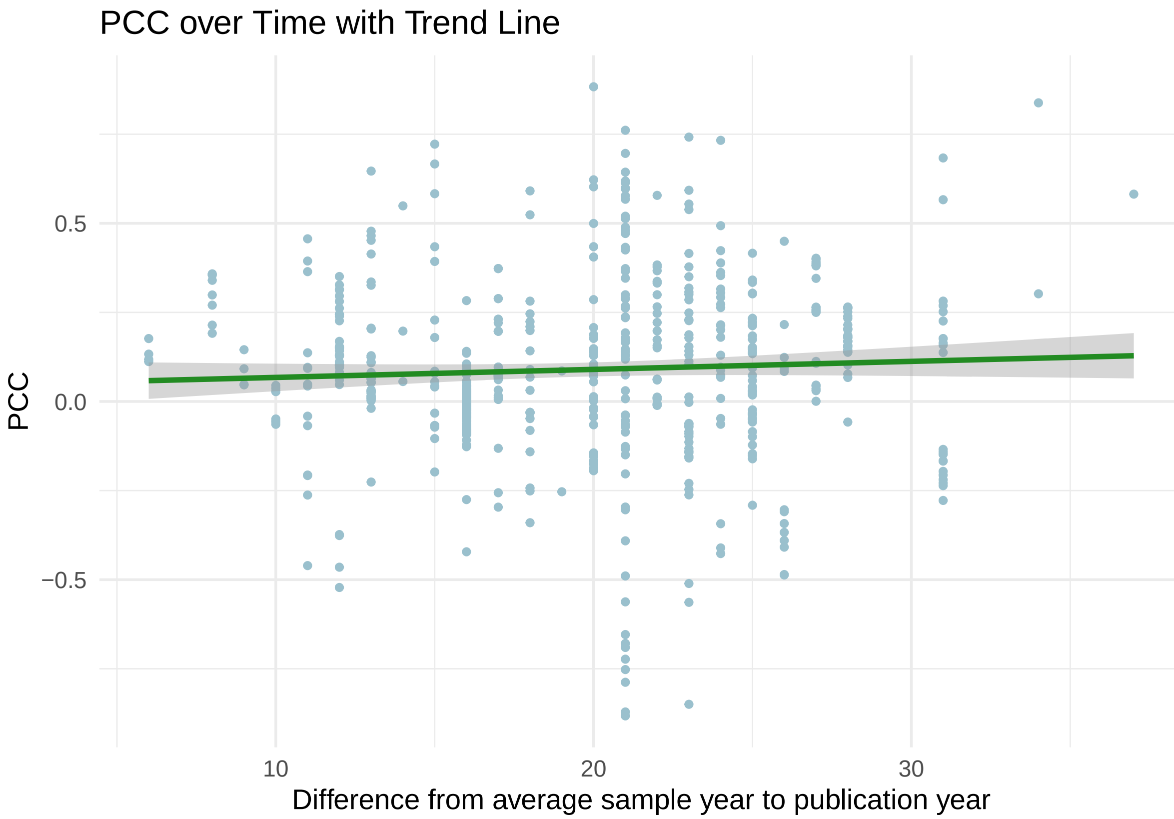
```
##
## Call:
## lm(formula = FDII$PCC ~ FDII$avg_sample_year)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98498 -0.11107  0.00022  0.12535  0.79073
##
## Coefficients:
##                       Estimate Std. Error t value   Pr(>|t|)
## (Intercept)          10.241536   2.265050   4.522 0.00000737 ***
## FDII$avg_sample_year -0.005095   0.001137  -4.482 0.00000881 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2441 on 614 degrees of freedom
## Multiple R-squared:  0.03168,    Adjusted R-squared:  0.03011
## F-statistic: 20.09 on 1 and 614 DF,  p-value: 0.000008813
```

Difference between publication and sample year trend

```r
FDII$difference <- FDII$publication_time - FDII$avg_sample_year

ggplot(FDII, aes(x = difference, y = PCC)) +
  geom_point(size = 1, color = "lightblue3") +
```

```r
  geom_smooth(method = "lm", col = "forestgreen") +
  labs(title = "PCC over Time with Trend Line", x = "Difference from average sample
  ↪  year to publication year", y = "PCC") +
  theme_minimal()
```

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




```r
summary(lm(FDII$PCC ~ FDII$difference))
```

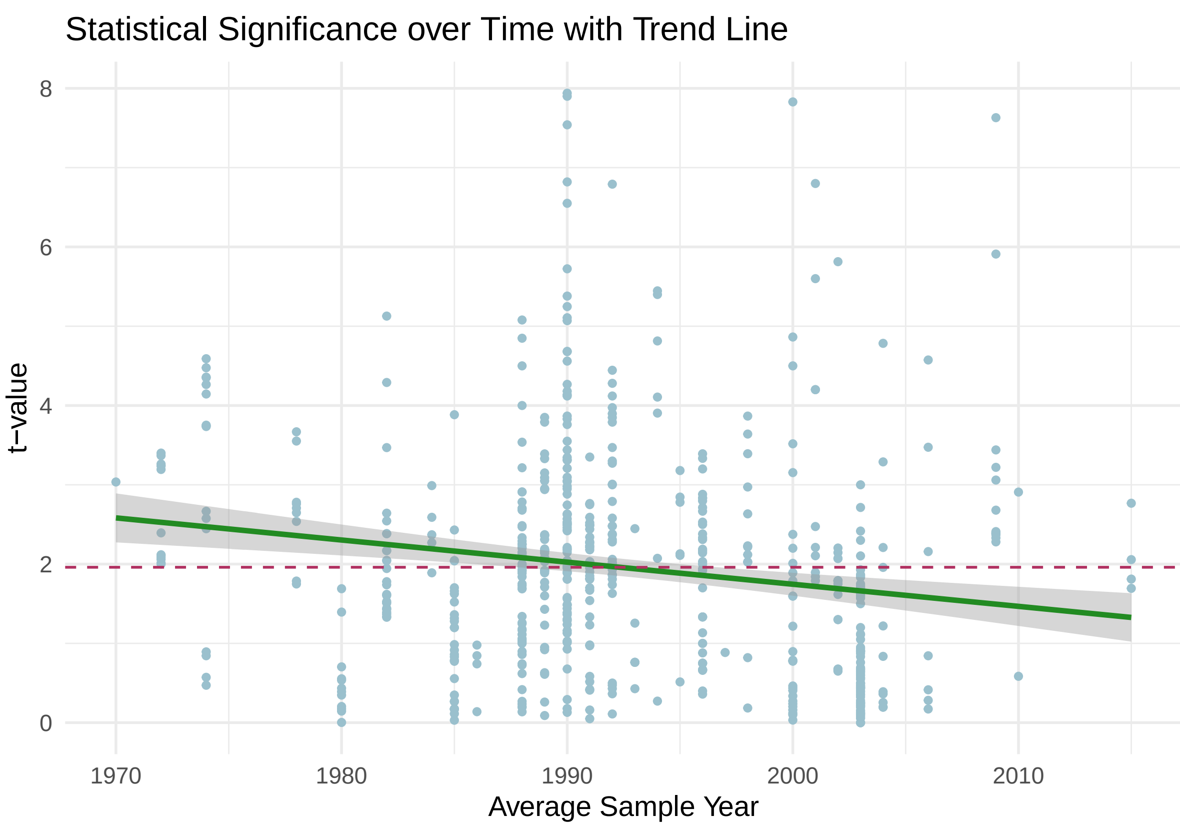
```
##
## Call:
## lm(formula = FDII$PCC ~ FDII$difference)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.97458 -0.11401 -0.00833  0.12937  0.79320
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.044973   0.036121   1.245    0.214
## FDII$difference 0.002254   0.001772   1.272    0.204
##
## Residual standard error: 0.2477 on 614 degrees of freedom
## Multiple R-squared:  0.002629,   Adjusted R-squared:  0.001004
## F-statistic: 1.618 on 1 and 614 DF,  p-value: 0.2038
```

I wanted to see whether articles that are published closer to their average sample year have greater impacts or not. I could test later whether these are heterogeneous or not. They are not statistically significant, suggesting that the recency of the article bears little-to-no weight in the likelihood of prediction. Stanley talks in the book about how papers are often quicker to publish is there are statistically-significant results.

Average sample year trend (statistical significance)

```
FDII$mod_t_value <- ifelse(sqrt(FDII$t_value_calculated^2) < 8,
                           sqrt(FDII$t_value_calculated^2),
                           NA)


ggplot(FDII, aes(x = avg_sample_year, y = mod_t_value)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "lm", col = "forestgreen") +
  geom_hline(yintercept = 1.96, color = "maroon", linetype = "dashed") +
  labs(title = "Statistical Significance over Time with Trend Line",
       x = "Average Sample Year",
       y = "t-value") +
  theme_minimal()
```

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

## Warning: Removed 6 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 6 rows containing missing values or values outside the scale
## range (`geom_point()`).
```




```
summary(lm(FDII$mod_t_value ~ FDII$avg_sample_year))
```

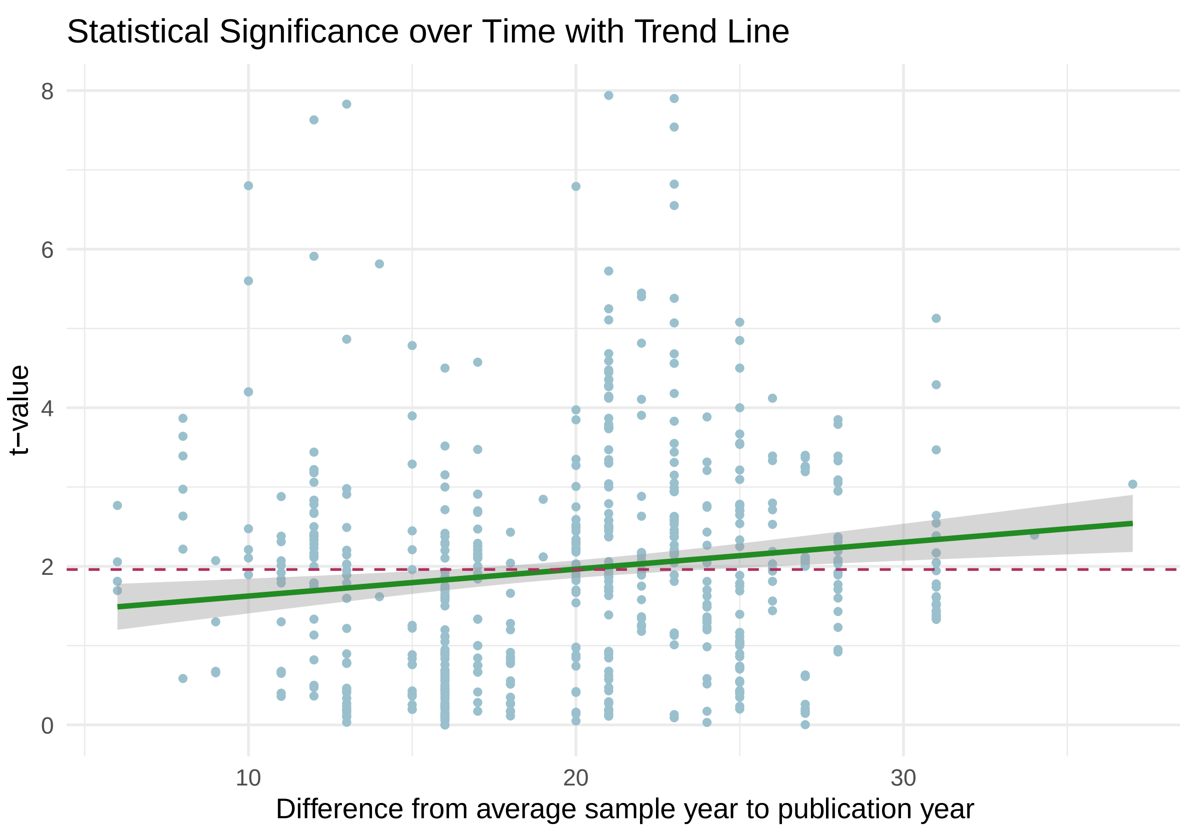
```
##
## Call:
## lm(formula = FDII$mod_t_value ~ FDII$avg_sample_year)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -2.3008 -1.0349 -0.0765  0.6100  6.1353
##
## Coefficients:
##                      Estimate Std. Error t value  Pr(>|t|)
## (Intercept)          57.54155   12.93309   4.449 0.0000103 ***
## FDII$avg_sample_year -0.02790    0.00649  -4.298 0.0000200 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.383 on 608 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.02949,    Adjusted R-squared:  0.0279
## F-statistic: 18.48 on 1 and 608 DF,  p-value: 0.00002004
```

Difference between publication and sample year trend (statistical significance)

```r
ggplot(FDII, aes(x = difference, y = mod_t_value)) +
  geom_point(size = 1, color = "lightblue3") +
  geom_smooth(method = "lm", col = "forestgreen") +
  geom_hline(yintercept = 1.96, color = "maroon", linetype = "dashed") +
  labs(title = "Statistical Significance over Time with Trend Line",
       x = "Difference from average sample year to publication year",
       y = "t-value") +
  theme_minimal()
```

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

## Warning: Removed 6 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 6 rows containing missing values or values outside the scale
## range (`geom_point()`).
```




```r
summary(lm(FDII$mod_t_value ~ FDII$difference))
```

```
##
## Call:
## lm(formula = FDII$mod_t_value ~ FDII$difference)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.1990 -1.0700 -0.1144  0.6286  6.1021
##
## Coefficients:
##                 Estimate Std. Error t value       Pr(>|t|)
## (Intercept)      1.28561    0.20449   6.287 0.000000000618 ***
## FDII$difference  0.03394    0.01003   3.384        0.00076 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.391 on 608 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.01849,    Adjusted R-squared:  0.01687
## F-statistic: 11.45 on 1 and 608 DF,  p-value: 0.0007601
```

Machine Learning Setup

```r
# Use World Bank API
library(WDI)

# Import CPI data
```

```r
## Corruption Perceptions Index data
CPI_Corruption <- read_excel("CPI_2012-2022.xlsx")

## WGI Corruption Indicators data
WGI_Corruption <- read_excel("WGI_Corruption_Indicators.xlsx")
## Adjust WGI data from a scale of -2.5-2.5 to 0-100
WGI_adjustment <- WGI_Corruption[, 5:ncol(WGI_Corruption)]
WGI_adjustment <- (WGI_adjustment + 2.5) * 20
WGI_adjustment <- round(WGI_adjustment, 3)
WGI_Corruption[, 5:ncol(WGI_Corruption)] <- WGI_adjustment
```

```r
dat = WDI(indicator='CC.EST', country =
↪  c('CHN','FRA','GUY','USA','CAN','SWE','NPL'), start=2012, end=2022)

dat$CC.EST <- (dat$CC.EST + 2.5)*20
```

```r
ggplot(dat, aes(year, CC.EST, color=country)) + geom_line() +
    xlab('Year') + ylab('Corruption Prediction') + ylim(0,100)
```