Local Suffrage Increases Citizenship Acquisition: Evidence from the European Union Replication Log: Spain - Trajectory Balancing & Event Study

Hannah M. Alarian

March 16, 2024

```
#Use file Spain.dta
#Variable information can be found in Codebook tab Spain

#####Libraries-----------------------------------------------------------------
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
library(foreach)
```

```
##
## Attaching package: 'foreach'
##
## The following objects are masked from 'package:purrr':
##
##     accumulate, when
```

```
library(doParallel)
```

```
## Loading required package: iterators
## Loading required package: parallel
```

```
require(parallel)
library(panelView)
```

```
## ## See bit.ly/panelview4r for more info.
## ## Report bugs -> yiqingxu@stanford.edu.
```

```
library(devtools)
```

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

```
library(readstata13)
library(dplyr)
library(fixest)
library(modelsummary)
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'
##
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

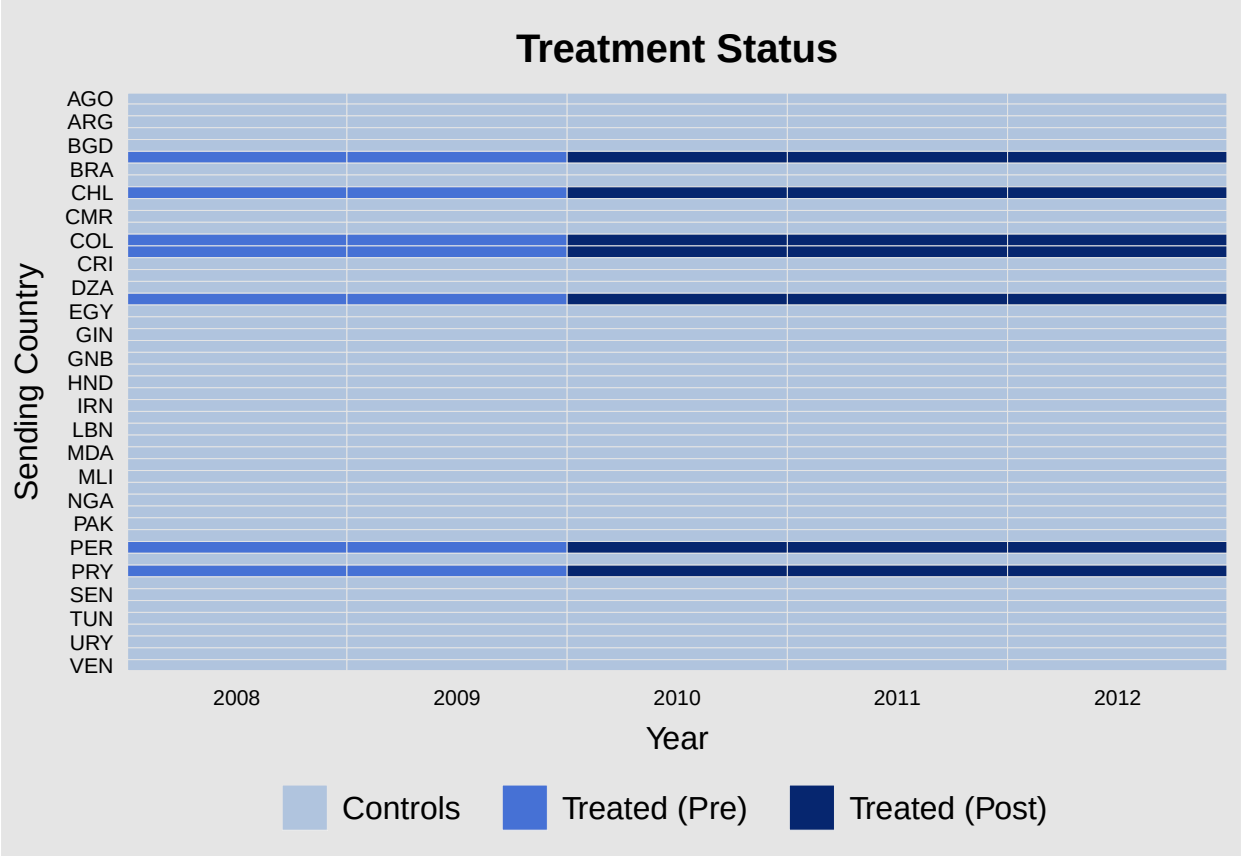
```
library(kbal) #install via github #install_github("chadhazlett/KBAL")
library(tjbal) #install via github #install_github("xuyiqing/tjbal")
```

```
## ## Syntax has changed since v.0.4.0.
##
## ## See http://bit.ly/tjbal4r for more info.
## ## Comments and suggestions -> yiqingxu@stanford.edu.
```

```
#Data Preparation-----------------------------------------------------------------
#Set working directory to location of data. Call in data from working directory
#setwd()
spain <- read.dta13("~/Dropbox/Research/Under Review/Local Suffrage/JOP/Data/Spain/Replication/Spain.dta
head(spain)
```

```
##   scou scountrylabel rcou rcountrylabel year citacq scode        sgdp sunemploy
## 1  AFG   Afghanistan  ESP         Spain 2009      9   AFG 42879504384       8.2
## 2  AFG   Afghanistan  ESP         Spain 2010      4   AFG 46495657984       8.1
## 3  AFG   Afghanistan  ESP         Spain 2011      8   AFG 49338253312       8.2
## 4  AFG   Afghanistan  ESP         Spain 2012      8   AFG 56460103680       8.0
## 5  AFG   Afghanistan  ESP         Spain 2013     14   AFG 58662371328       8.4
## 6  AFG   Afghanistan  ESP         Spain 2014      6   AFG 60240695296       8.6
##   sdc sdc_C spolity spolity2 comlang_ethno colony permres  Ll1sgdp l1sunemp
## 1   1     1     -66       NA             0      0      NA       NA       NA
## 2   1     1     -66       NA             0      0      NA 24.48166      8.2
## 3   1     1     -66       NA             0      0      NA 24.56262      8.1
## 4   1     1     -66       NA             0      0      NA 24.62197      8.2
## 5   1     1     -66       NA             0      0      NA 24.75680      8.0
## 6   1     1      -1       -1             0      0      NA 24.79506      8.4
##   NAmerica Africa MENA Oceania Asia Europe SCAmerica municipal SpainVote
## 1        0      0    0       0    1      0         0         0         0
## 2        0      0    0       0    1      0         0         0         0
## 3        0      0    0       0    1      0         0         0         0
## 4        0      0    0       0    1      0         0         0         0
## 5        0      0    0       0    1      0         0         0         0
## 6        0      0    0       0    1      0         0         0         0
##   time2010 DiD lpermres IHS_citacq
## 1        0   0       NA   2.893444
## 2        1   0       NA   2.094712
## 3        1   0       NA   2.776472
## 4        1   0       NA   2.776472
## 5        1   0       NA   3.333477
## 6        1   0       NA   2.491780
```

```r
#Removing Missing Cases and Unbalanced
spain <- spain[spain$scou != "BIH" & spain$scou != "BLR" & spain$scou != "ISR" &
                 spain$scou != "SYR" & spain$scou != "IRQ" & spain$scou != "CUB" &
                 spain$scou != "TUR" & spain$scou!= "GEO", ]

#Re-Coding Treatment variable
spain$treat <- 0
spain$treat[spain$municipal == 1 & spain$year >= 2010] <- 1
spain$treat[spain$scou =="CPV" & spain$year >= 2010] <- 1 #to run treatment at same time

#Classifying for Event Study
spain$treated <- 0
spain$treated[spain$scou == "ECU" | spain$scou == "BOL" | spain$scou == "PER"
            | spain$scou == "COL" | spain$scou ==  "CHL" | spain$scou =="PRY"
            | spain$scou == "CPV"] <- 1
spainT <- spain[spain$treated == 1,]

#Relabeling variables for graphs
spain <- spain %>%
  rename("Acquisitions" = "citacq",
    Residents = lpermres,
    GDPpc = sgdp,
    Unemployment = sunemploy,
    Democracy = spolity2,
    Colony = colony,
    Language = comlang_ethno,
    Perm.Res= permres
  )

#Subsetting for balance prior to 2012 & up to 2015 & including 2007
spain12 <- spain[spain$year >= 2008 & spain$year <= 2012, ]
spain12 <- spain12[spain12$scou %in% names(which(table(spain12$scou) > 4)), ]

spain15 <- spain[spain$year >= 2008, ]
spain15 <- spain15[spain15$scou %in% names(which(table(spain15$scou) > 7)), ]

spain7 <- spain[spain$year >= 2007, ]
spain7 <- spain7[spain7$scou %in% names(which(table(spain7$scou) > 8)), ]

spainT <- spainT[spainT$year >= 2008 & spainT$year <= 2012, ]

#Looking at data structure
panelview(Acquisitions ~ treat, data = spain12, show.id = c(1:49),
          index = c("scou", "year"), xlab = "Year", ylab = "Sending Country",
          axis.lab.gap = c(0,1), pre.post = TRUE)
```




```
with(spainT, table(year, treat))
```

```
##       treat
## year   0 1
##   2008 7 0
##   2009 7 0
##   2010 0 7
##   2011 0 7
##   2012 0 7
```

```
#Replicating Analysis-----------------------------------------------------------
##Models##
#Average Treatment Effect on the Treated by Year (Mean Balanced)
out.mbal <- tjbal(Y= "Acquisitions", D = "treat",
                  X = c("Colony", "Language","Unemployment"),
                  data = spain12,
                  X.avg.time = list(c(2008), c(2008), c(2008:2009)),
                  estimator = "mean",
                  index = c("scou", "year"), parallel = F,
                  demean = T, vce = "boot", nsims = 1000, seed = 1013, conf.lvl = .9)
```

```
## Seek balance on:
## Acquisitions.dm2008, Acquisitions.dm2009, Colony, Language, Unemployment
##
## Optimization:
## bias.ratio = 0.0000; num.dims = 4 (mbal)
##
```

```
## Balance Table
##                       mean.tr mean.co.pre mean.co.pst    sd.tr sd.co.pre
## Acquisitions.dm2008 -17.8571     49.5357    -17.8571 465.2988  165.4094
## Acquisitions.dm2009  17.8571    -49.5357     17.8571 465.2988  165.4094
## Colony                0.8571      0.2857      0.8571   0.3780    0.4572
## Language              0.8571      0.2857      0.8571   0.3780    0.4572
## Unemployment          7.5571      8.5619      7.5571   3.0492    4.6109
##                     sd.co.pst diff.pre diff.pst
## Acquisitions.dm2008   76.4561  -0.1448        0
## Acquisitions.dm2009   76.4561   0.1448        0
## Colony                 0.3499   1.5119        0
## Language               0.3499   1.5119        0
## Unemployment           4.0309  -0.3295        0
##
## Bootstrapping...
## ..................
```

```
#Trajectory Balancing - Maximizing Covariate Inclusion
out.mbal1r <- tjbal(Y= "Acquisitions", D = "treat",
                    X = c("Colony", "Language","Unemployment", "Democracy", "GDPpc"),
                    data = spain12,
                    X.avg.time = list(c(2008), c(2008), c(2008:2009), c(2008:2009), c(2008:2009)),
                    estimator = "mean",
                    index = c("scou", "year"), parallel = F,
                    demean = T, vce = "boot", nsims = 1000, seed = 1013, conf.lvl = .9)
```

```
## Seek balance on:
## Acquisitions.dm2008, Acquisitions.dm2009, Colony, Language, Unemployment, Democracy, GDPpc
##
## Optimization:
## bias.ratio = 0.0000; num.dims = 6 (mbal)
##
## Balance Table
##                            mean.tr   mean.co.pre   mean.co.pst        sd.tr
## Acquisitions.dm2008 -1.785710e+01  4.953570e+01 -1.785710e+01 4.652988e+02
## Acquisitions.dm2009  1.785710e+01 -4.953570e+01  1.785710e+01 4.652988e+02
## Colony               8.571000e-01  2.857000e-01  8.571000e-01 3.780000e-01
## Language             8.571000e-01  2.857000e-01  8.571000e-01 3.780000e-01
## Unemployment         7.557100e+00  8.561900e+00  7.557100e+00 3.049200e+00
## Democracy            8.071400e+00  3.476200e+00  8.071400e+00 1.789500e+00
## GDPpc                1.842273e+11  1.131949e+12  1.842274e+11 1.750498e+11
##                        sd.co.pre    sd.co.pst diff.pre diff.pst
## Acquisitions.dm2008 1.654094e+02 9.124100e+01  -0.1448        0
## Acquisitions.dm2009 1.654094e+02 9.124100e+01   0.1448        0
## Colony              4.572000e-01 3.499000e-01   1.5119        0
## Language            4.572000e-01 3.499000e-01   1.5119        0
## Unemployment        4.610900e+00 3.642700e+00  -0.3295        0
## Democracy           5.768900e+00 3.164000e+00   2.5679        0
## GDPpc               2.905114e+12 6.381477e+11  -5.4140        0
##
## Bootstrapping...
## ..................
```

```
#Trajectory Balancing with All Available Time Points (2007 - 2015)
out.mbal7 <- tjbal(Y= "Acquisitions", D = "treat",
```

```
                  X = c("Colony", "Language","Unemployment", "Democracy", "GDPpc"),
                  data = spain7,
                  X.avg.time = list(c(2008), c(2008), c(2007:2009), c(2007:2009), c(2007:2009)),
                  estimator = "mean",
                  index = c("scou", "year"), parallel = F,
                  demean = T, vce = "boot", nsims = 1000, seed = 1013, conf.lvl = .9)
```

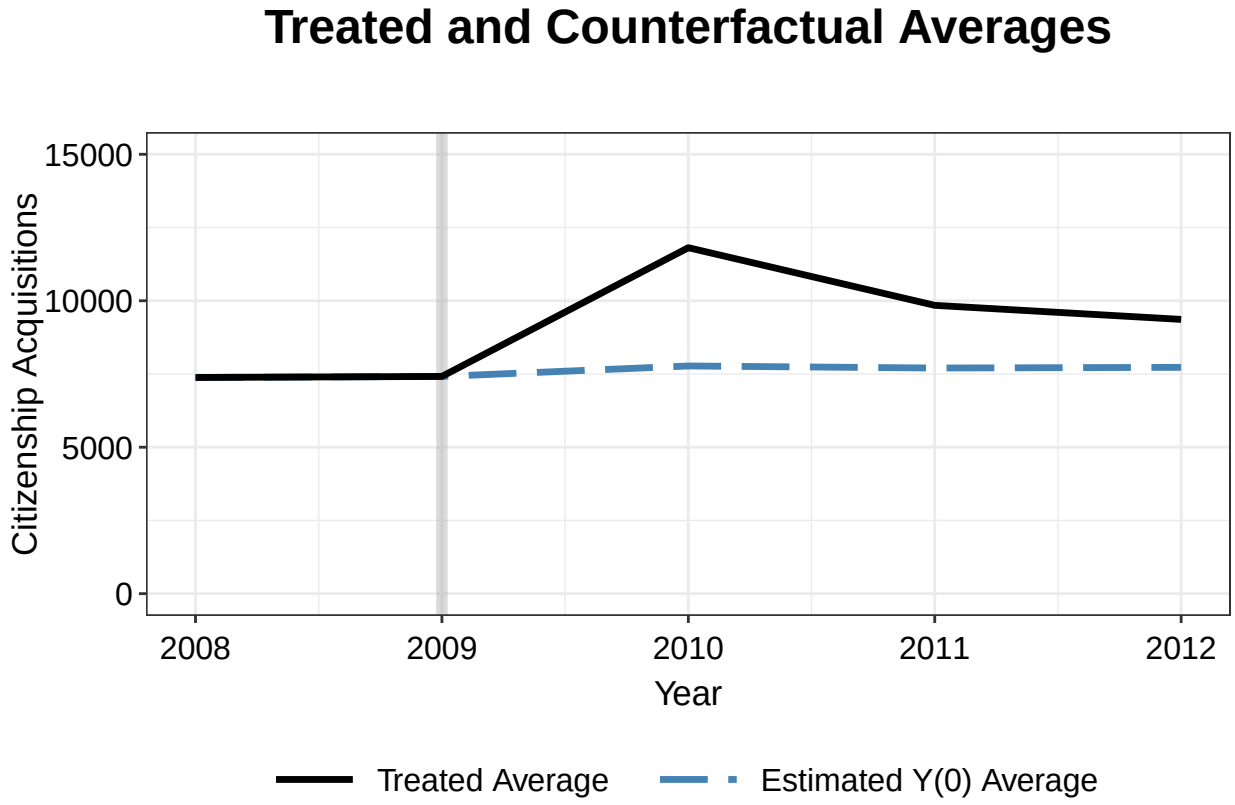
```
## Seek balance on:
## Acquisitions.dm2007, Acquisitions.dm2008, Acquisitions.dm2009, Colony, Language, Unemployment, Democ
##
## Optimization:
## bias.ratio = 0.8000; num.dims = 1 (mbal)
##
## Balance Table
##                            mean.tr   mean.co.pre  mean.co.pst        sd.tr
## Acquisitions.dm2007 -7.976190e+02 -2.102380e+01 -6.59028e+01 1.021532e+03
## Acquisitions.dm2008  3.809524e+02  6.004760e+01  4.57481e+01 5.989176e+02
## Acquisitions.dm2009  4.166667e+02 -3.902380e+01  2.01547e+01 7.720542e+02
## Colony               8.571000e-01  2.857000e-01  9.95800e-01 3.780000e-01
## Language             8.571000e-01  2.857000e-01  9.95200e-01 3.780000e-01
## Unemployment         7.552400e+00  8.531700e+00  6.11610e+00 2.830200e+00
## Democracy            8.095200e+00  3.500000e+00  8.61070e+00 1.781700e+00
## GDPpc                1.810744e+11  1.116156e+12  3.17936e+12 1.725793e+11
##                          sd.co.pre    sd.co.pst diff.pre diff.pst
## Acquisitions.dm2007 8.220550e+01 1.111972e+02  -0.7602  -0.7163
## Acquisitions.dm2008 1.583654e+02 7.741270e+01   0.5358   0.5597
## Acquisitions.dm2009 1.817135e+02 1.289996e+02   0.5902   0.5136
## Colony              4.572000e-01 6.500000e-02   1.5119  -0.3668
## Language            4.572000e-01 6.910000e-02   1.5119  -0.3653
## Unemployment        4.639100e+00 1.555200e+00  -0.3460   0.5075
## Democracy           5.721800e+00 1.993600e+00   2.5791  -0.2893
## GDPpc               2.872835e+12 5.837799e+12  -5.4183 -17.3734
##
## Bootstrapping...
## ..................
```

```
#Permanent Residents Trajectory Balancing (2008 - 2015)
out.mbal15r2 <- tjbal(Y= "Acquisitions", D = "treat",
                      X = c("Perm.Res"),
                      data = spain15,
                      X.avg.time = list(c(2008:2009)),
                      estimator = "mean",
                      index = c("scou", "year"), parallel = F,
                      demean = T, vce = "boot", nsims = 1000, seed = 1013, conf.lvl = .9)
```

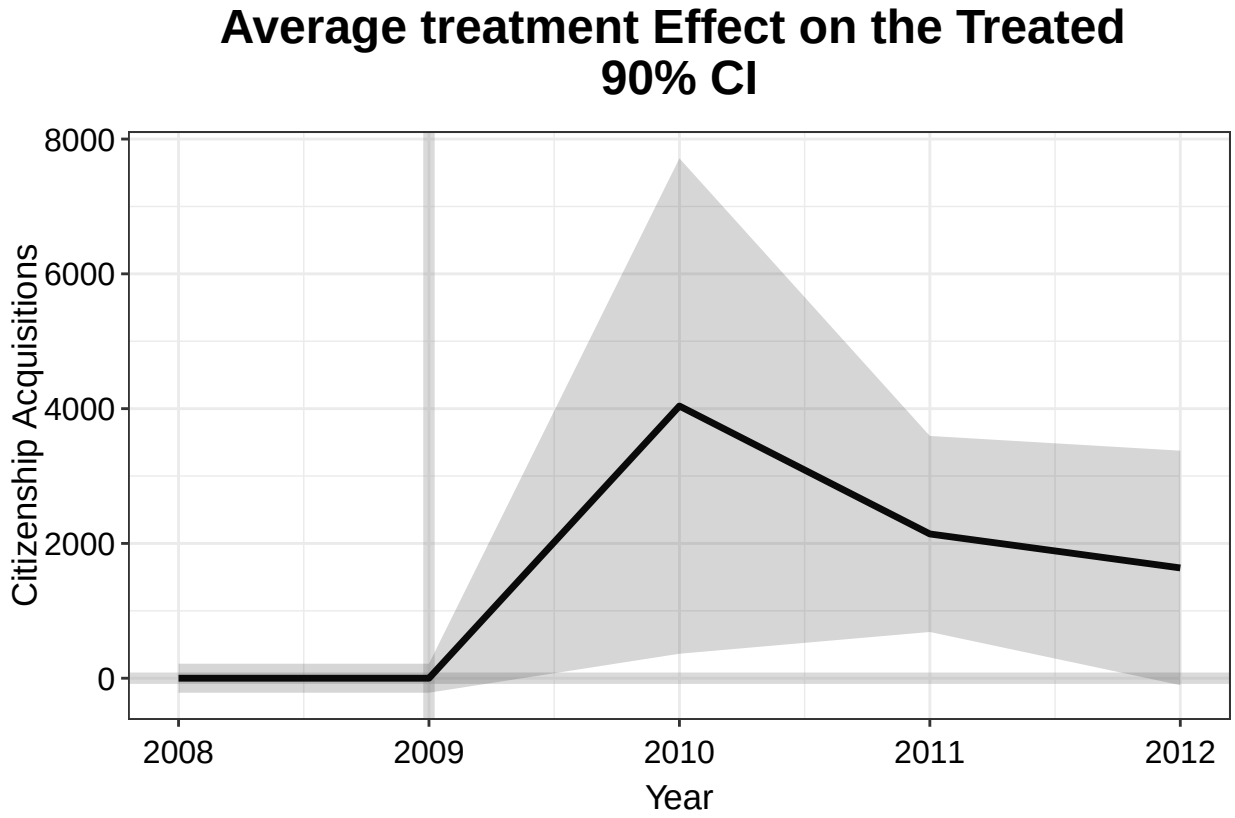
```
## Seek balance on:
## Acquisitions.dm2008, Acquisitions.dm2009, Perm.Res
##
## Optimization:
## bias.ratio = 0.8469; num.dims = 1 (mbal)
##
## Balance Table
##                          mean.tr mean.co.pre mean.co.pst       sd.tr   sd.co.pre
## Acquisitions.dm2008    -17.8571     49.5357    102.5681    465.2988    165.4094
```

```
## Acquisitions.dm2009     17.8571    -49.5357   -102.5681    465.2988    165.4094
## Perm.Res            144658.2143  40172.2143  78247.3151 158247.1123 113560.8188
##                       sd.co.pst diff.pre diff.pst
## Acquisitions.dm2008    259.9889  -0.1448  -0.2588
## Acquisitions.dm2009    259.9889   0.1448   0.2588
## Perm.Res            189188.1608   0.6603   0.4197
##
## Bootstrapping...
## ..................
```

```r
#Event Study
set.seed(1013) #setting seed for replication
M <- feols(citacq ~ i(year, ref = 2009), data = spainT,
           cluster = 'scode')
names(M$coefficients) <- c("2009", "2008", "2010", "2011", "2012")

#Figure 3-------------------------------------------------------------
#Panel a
plot(out.mbal, type = "ct", ylim = c(0, 15000),
     main = "Treated and Counterfactual Averages\n",
     ylab = "Citizenship Acquisitions", xlab = "Year", count = F)
```



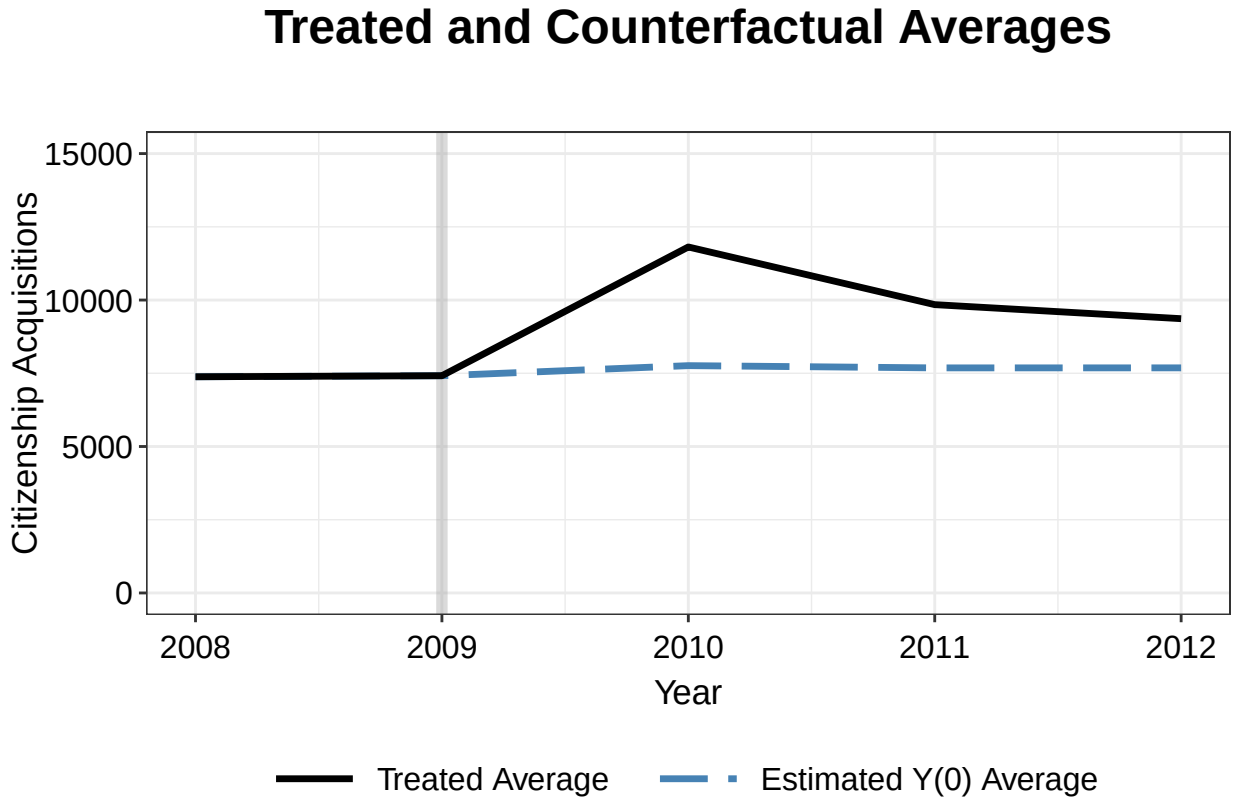

```r
#Panel b
plot(out.mbal, ylab = "Citizenship Acquisitions", xlab = "Year",
     main = "Average treatment Effect on the Treated \n90% CI", count = F)
```



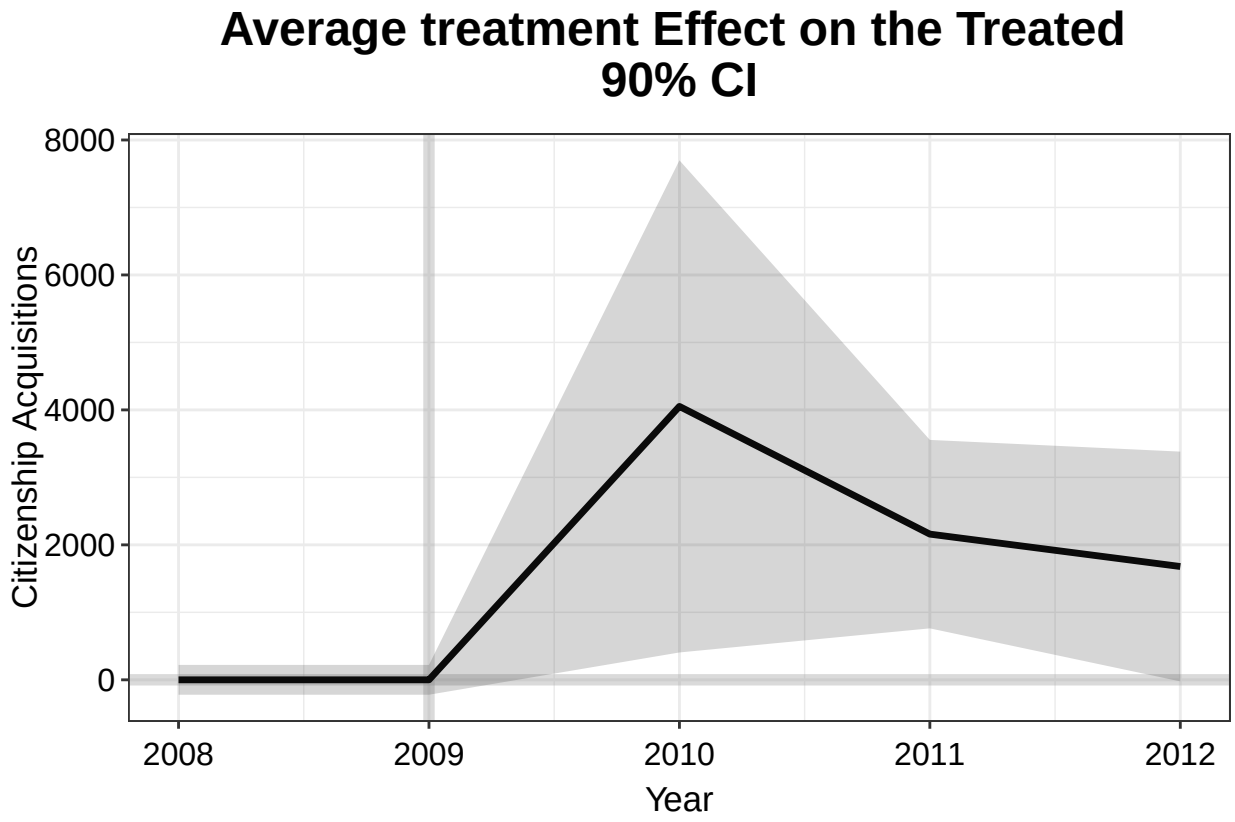

```
#Appendix B.1.2-------------------------------------------------------------
print(out.mbal)
```

```
## Call:
## tjbal.default(data = spain12, Y = "Acquisitions", D = "treat",
##     X = c("Colony", "Language", "Unemployment"), X.avg.time = list(c(2008),
##         c(2008), c(2008:2009)), index = c("scou", "year"), demean = T,
##     estimator = "mean", vce = "boot", conf.lvl = 0.9, nsims = 1000,
##     parallel = F, seed = 1013)
##
##    ~ by Period (including Pre-treatment Periods):
##       ATT   S.E. z-score CI.lower CI.upper p.value n.Treated
## 2008    0  130.8   0.000   -215.1    215.1  1.0000         7
## 2009    0  130.8   0.000   -215.1    215.1  1.0000         7
## 2010 4039 2234.5   1.808    363.4   7714.1  0.0707         7
## 2011 2141  884.3   2.421    686.1   3595.0  0.0155         7
## 2012 1637 1057.7   1.547   -103.0   3376.5  0.1218         7
##
## Average Treatment Effect on the Treated:
##       ATT S.E. z-score CI.lower CI.upper p.value
## [1,] 2605 1010    2.58    944.5     4266  0.0099
```

```
#Appendix C.1.2-------------------------------------------------------------
#Panel a
plot(out.mbal1r, type = "ct", ylim = c(0, 15000), main = "Treated and Counterfactual Averages\n",
     ylab = "Citizenship Acquisitions", xlab = "Year", count = F)
```



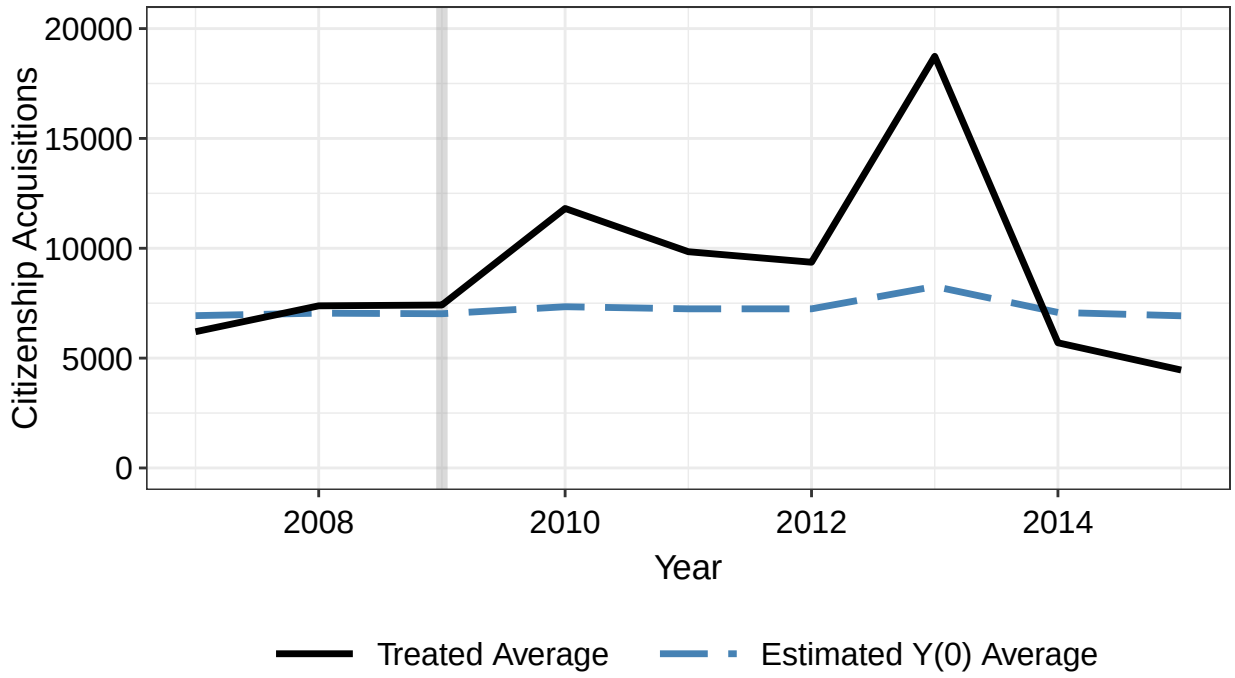

```
#Panel b
plot(out.mbal1r, ylab = "Citizenship Acquisitions", xlab = "Year",
     main = "Average treatment Effect on the Treated \n90% CI", count = F)
```



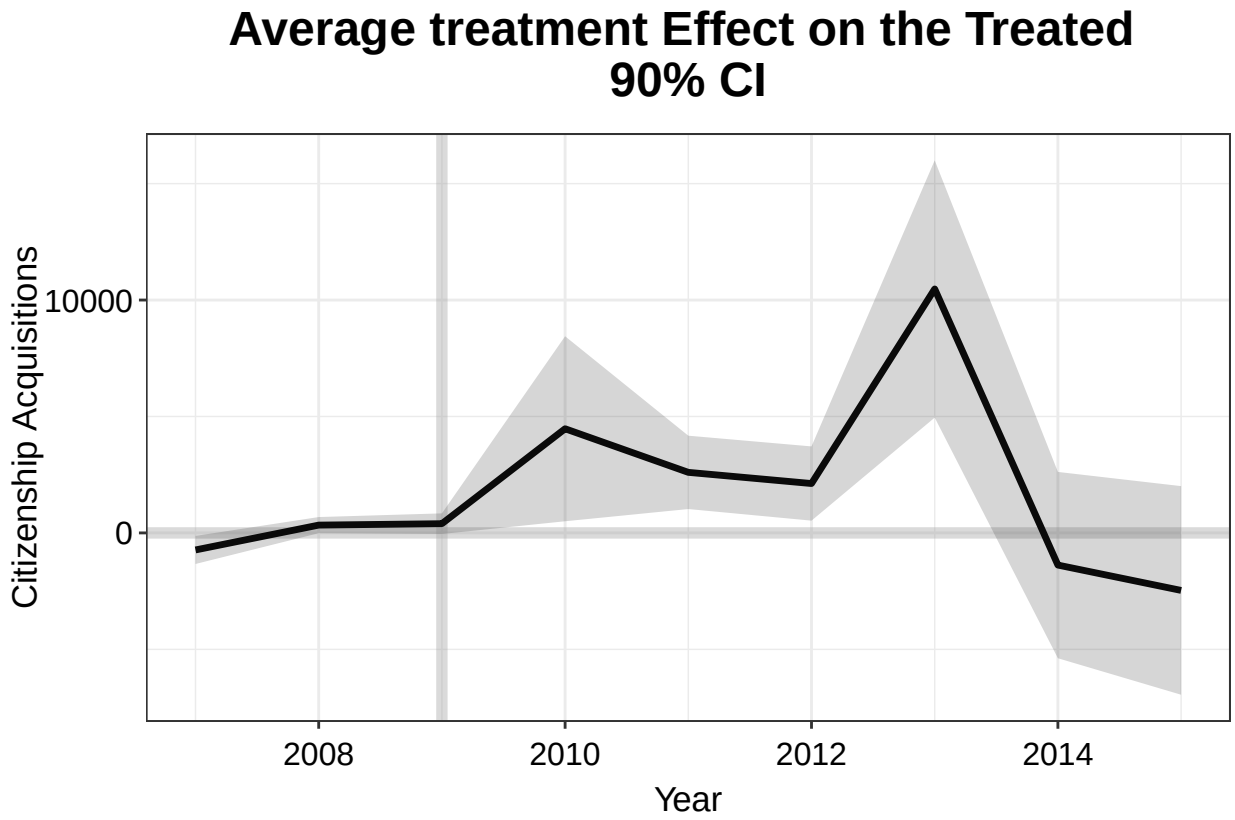

```
#Table
print(out.mbal1r)
```

```
## Call:
## tjbal.default(data = spain12, Y = "Acquisitions", D = "treat",
##     X = c("Colony", "Language", "Unemployment", "Democracy",
##         "GDPpc"), X.avg.time = list(c(2008), c(2008), c(2008:2009),
##         c(2008:2009), c(2008:2009)), index = c("scou", "year"),
##     demean = T, estimator = "mean", vce = "boot", conf.lvl = 0.9,
##     nsims = 1000, parallel = F, seed = 1013)
##
##    ~ by Period (including Pre-treatment Periods):
##       ATT   S.E. z-score CI.lower CI.upper p.value n.Treated
## 2008    0  134.3   0.000   -220.9    220.9  1.0000         7
## 2009    0  134.3   0.000   -220.9    220.9  1.0000         7
## 2010 4052 2217.0   1.828    405.3   7698.6  0.0676         7
## 2011 2159  849.0   2.543    762.6   3555.5  0.0110         7
## 2012 1679 1035.2   1.622    -23.5   3382.0  0.1048         7
##
## Average Treatment Effect on the Treated:
##       ATT  S.E. z-score CI.lower CI.upper p.value
## [1,] 2630 977.3   2.691     1022     4238  0.0071
```

```
#Appendix C.1.3-------------------------------------------------------------
#Panel a
plot(out.mbal7, type = "ct", ylim = c(0, 20000), main = "Treated and Counterfactual Averages\n",
     ylab = "Citizenship Acquisitions", xlab = "Year", count = F)
```




```
#Panel b
plot(out.mbal7, ylab = "Citizenship Acquisitions", xlab = "Year",
     main = "Average treatment Effect on the Treated \n90% CI", count = F)
```

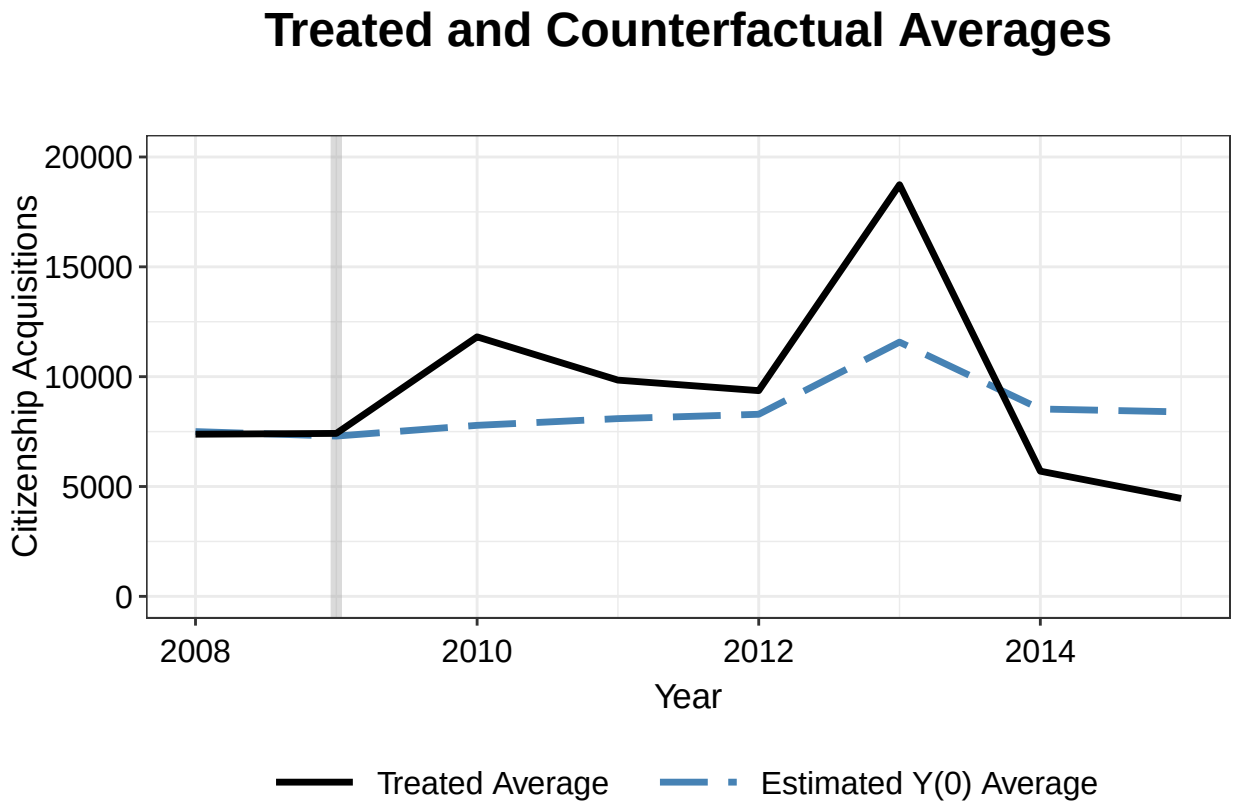



```
#Table
print(out.mbal7)
```

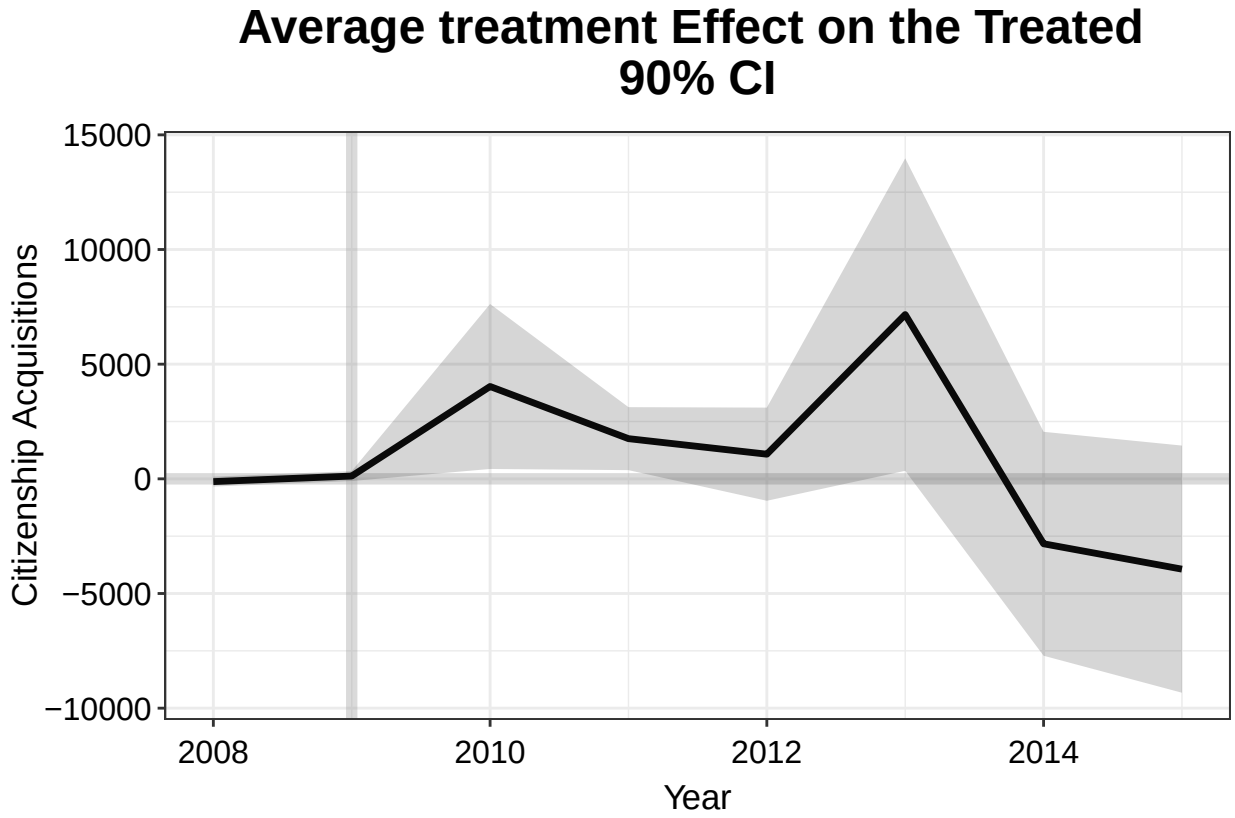
```
## Call:
## tjbal.default(data = spain7, Y = "Acquisitions", D = "treat",
##     X = c("Colony", "Language", "Unemployment", "Democracy",
##         "GDPpc"), X.avg.time = list(c(2008), c(2008), c(2007:2009),
##         c(2007:2009), c(2007:2009)), index = c("scou", "year"),
##     demean = T, estimator = "mean", vce = "boot", conf.lvl = 0.9,
##     nsims = 1000, parallel = F, seed = 1013)
##
##    ~ by Period (including Pre-treatment Periods):
##          ATT   S.E. z-score CI.lower CI.upper p.value n.Treated
## 2007  -731.7  366.1 -1.9986 -1333.91   -129.5  0.0456         7
## 2008   335.2  210.6  1.5919   -11.16    681.6  0.1114         7
## 2009   396.5  269.0  1.4741   -45.94    839.0  0.1405         7
## 2010  4472.9 2414.8  1.8523   500.86   8444.9  0.0640         7
## 2011  2600.2  957.9  2.7144  1024.55   4175.8  0.0066         7
## 2012  2120.5  969.1  2.1881   526.46   3714.5  0.0287         7
## 2013 10475.8 3358.3  3.1194  4951.95  15999.7  0.0018         7
## 2014 -1381.3 2432.2 -0.5679 -5381.99   2619.4  0.5701         7
## 2015 -2468.9 2723.2 -0.9066 -6948.19   2010.3  0.3646         7
##
## Average Treatment Effect on the Treated:
##       ATT S.E. z-score CI.lower CI.upper p.value
## [1,] 2637 1052   2.506    906.3     4367  0.0122
```

```
#Appendix C.1.4----------------------------------------------------------------
#Panel a
```

```
plot(out.mbal15r2, type = "ct", ylim = c(0, 20000), main = "Treated and Counterfactual Averages\n",
     ylab = "Citizenship Acquisitions", xlab = "Year", count = F)
```



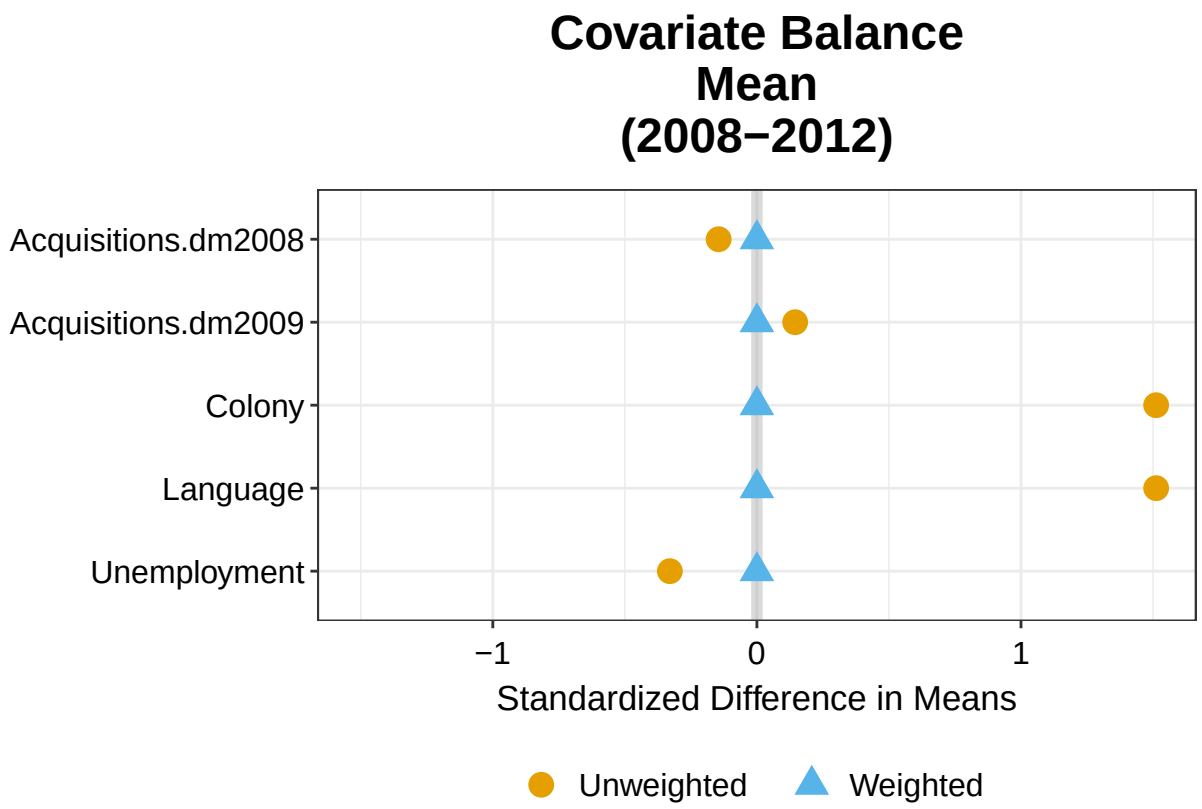

```
#Panel b
plot(out.mbal15r2, ylab = "Citizenship Acquisitions", xlab = "Year",
     main = "Average treatment Effect on the Treated \n90% CI", count = F)
```



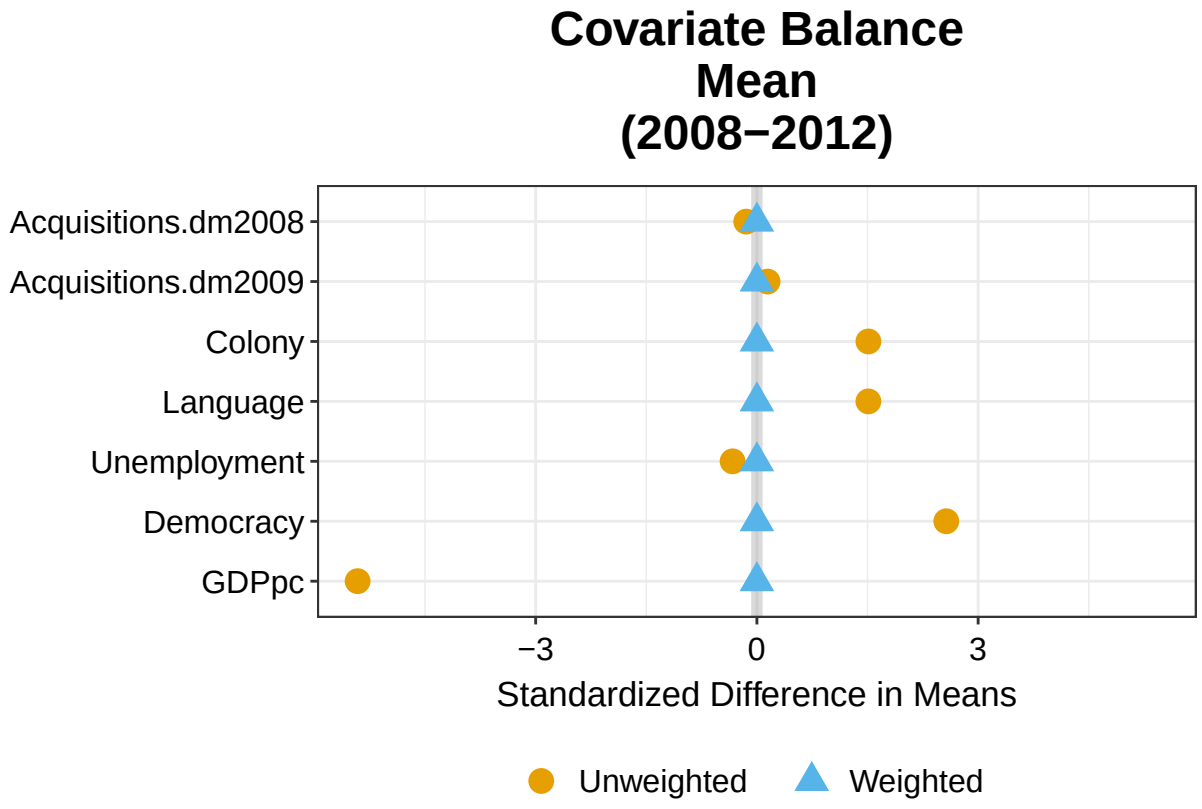

```
#Table
print(out.mbal15r2)
```

```
## Call:
## tjbal.default(data = spain15, Y = "Acquisitions", D = "treat",
##     X = c("Perm.Res"), X.avg.time = list(c(2008:2009)), index = c("scou",
##         "year"), demean = T, estimator = "mean", vce = "boot",
##     conf.lvl = 0.9, nsims = 1000, parallel = F, seed = 1013)
##
##    ~ by Period (including Pre-treatment Periods):
##          ATT   S.E. z-score CI.lower CI.upper p.value n.Treated
## 2008  -120.4  131.2 -0.9177  -336.28    95.43  0.3588         7
## 2009   120.4  131.2  0.9177   -95.43   336.28  0.3588         7
## 2010  4026.6 2186.7  1.8414   429.79  7623.39  0.0656         7
## 2011  1752.3  834.9  2.0989   379.05  3125.57  0.0358         7
## 2012  1074.7 1235.0  0.8702  -956.70  3106.10  0.3842         7
## 2013  7163.1 4142.2  1.7293   349.78 13976.41  0.0838         7
## 2014 -2831.5 2968.5 -0.9539 -7714.25  2051.17  0.3401         7
## 2015 -3938.6 3276.0 -1.2023 -9327.12  1449.90  0.2293         7
##
## Average Treatment Effect on the Treated:
##       ATT S.E. z-score CI.lower CI.upper p.value
## [1,] 1208 1539  0.7846    -1324     3740  0.4327
```

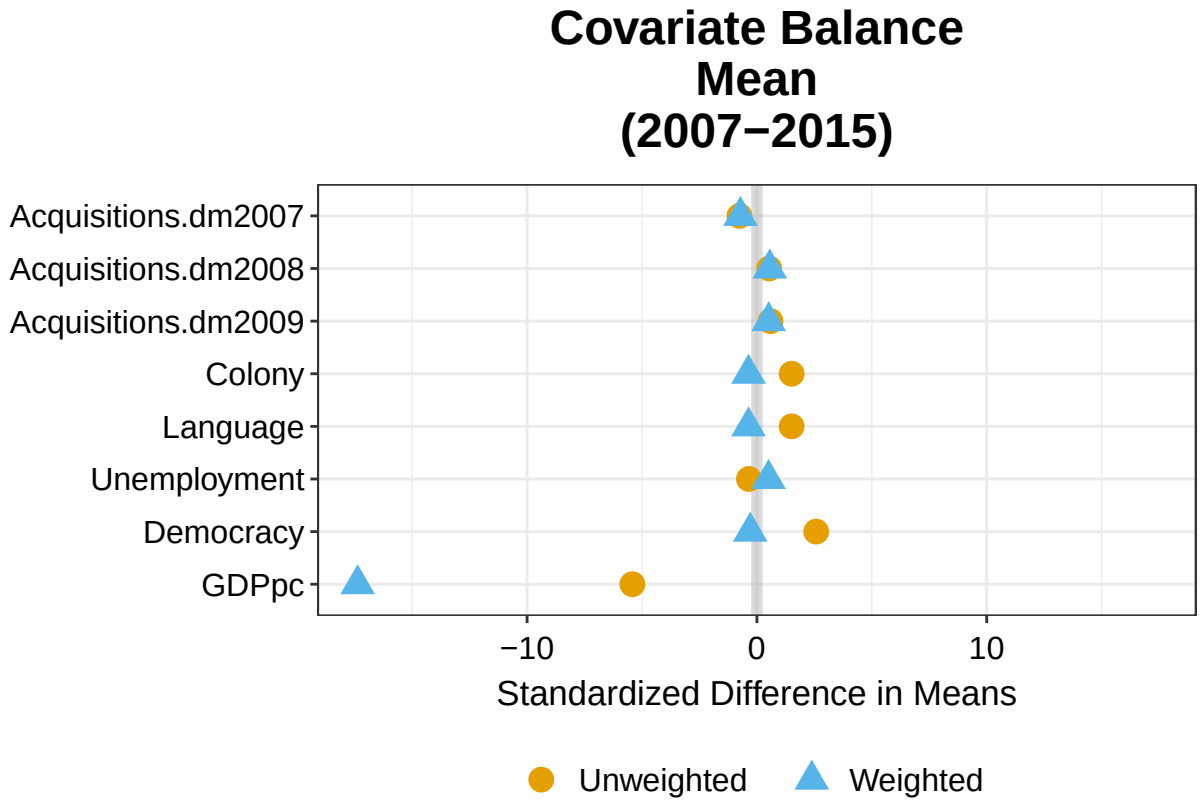
```
#Appendix C.1.5-----------------------------------------------------------------
#Standardized Difference in Weighted and Unweighted Coefficients
#Top Right panel
plot(out.mbal, type = "balance", main = "Covariate Balance\nMean\n(2008-2012)")
```



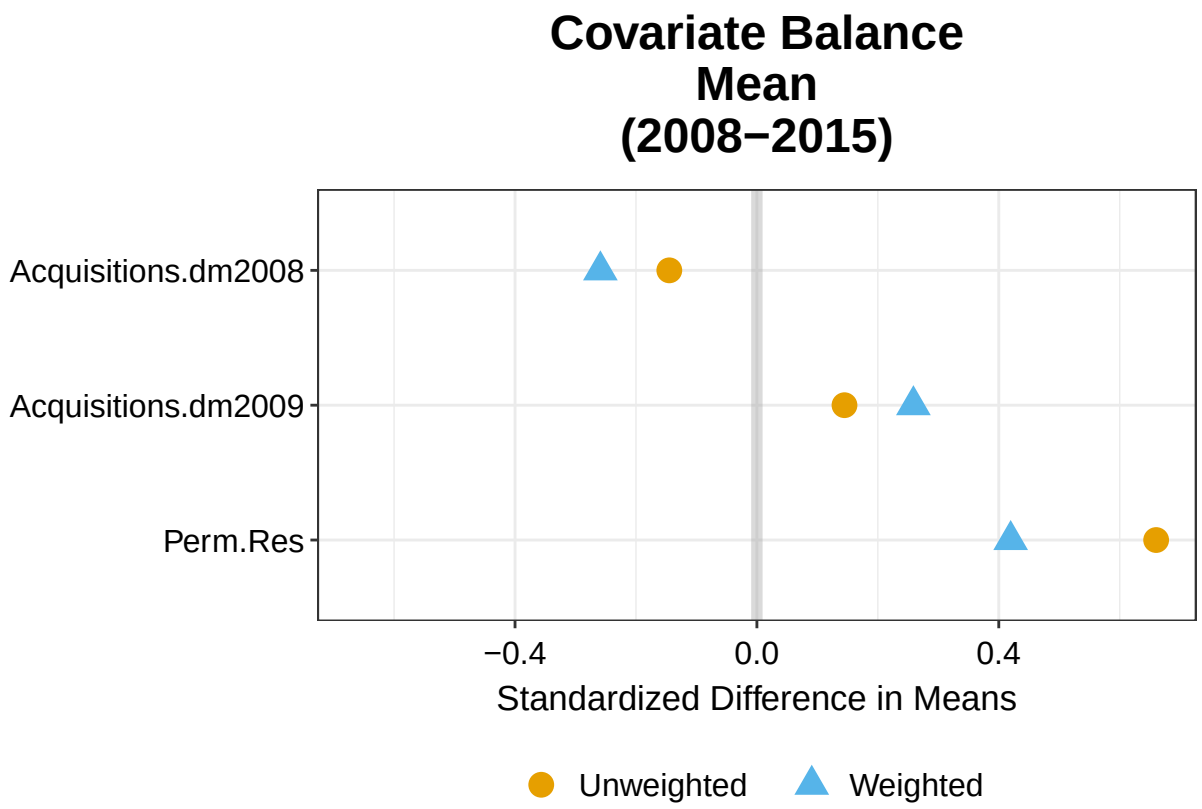

```
#Top Left panel
plot(out.mbal1r, type = "balance", main = "Covariate Balance\nMean\n(2008-2012)")
```



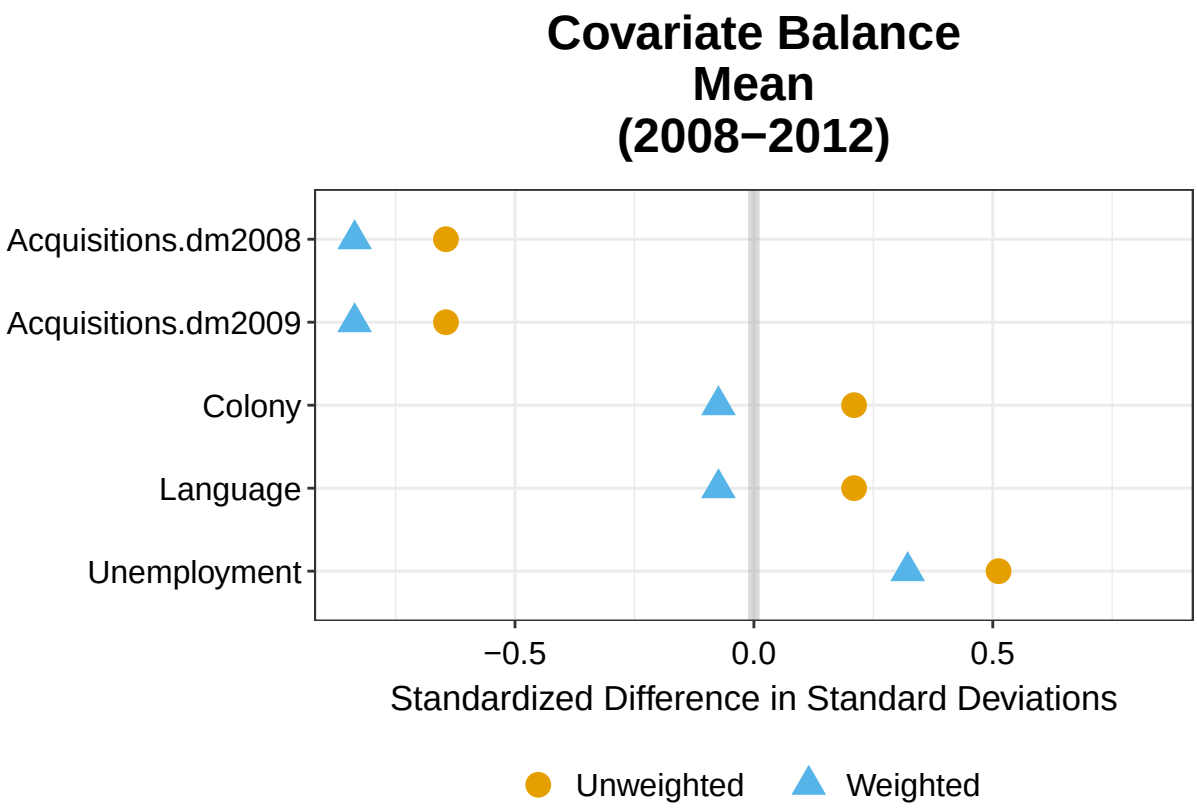
```r
#Bottom Right
plot(out.mbal7, type = "balance", main = "Covariate Balance\nMean\n(2007-2015)")
```




```r
#Bottom Left
plot(out.mbal15r2, type = "balance", main = "Covariate Balance\nMean\n(2008-2015)")
```

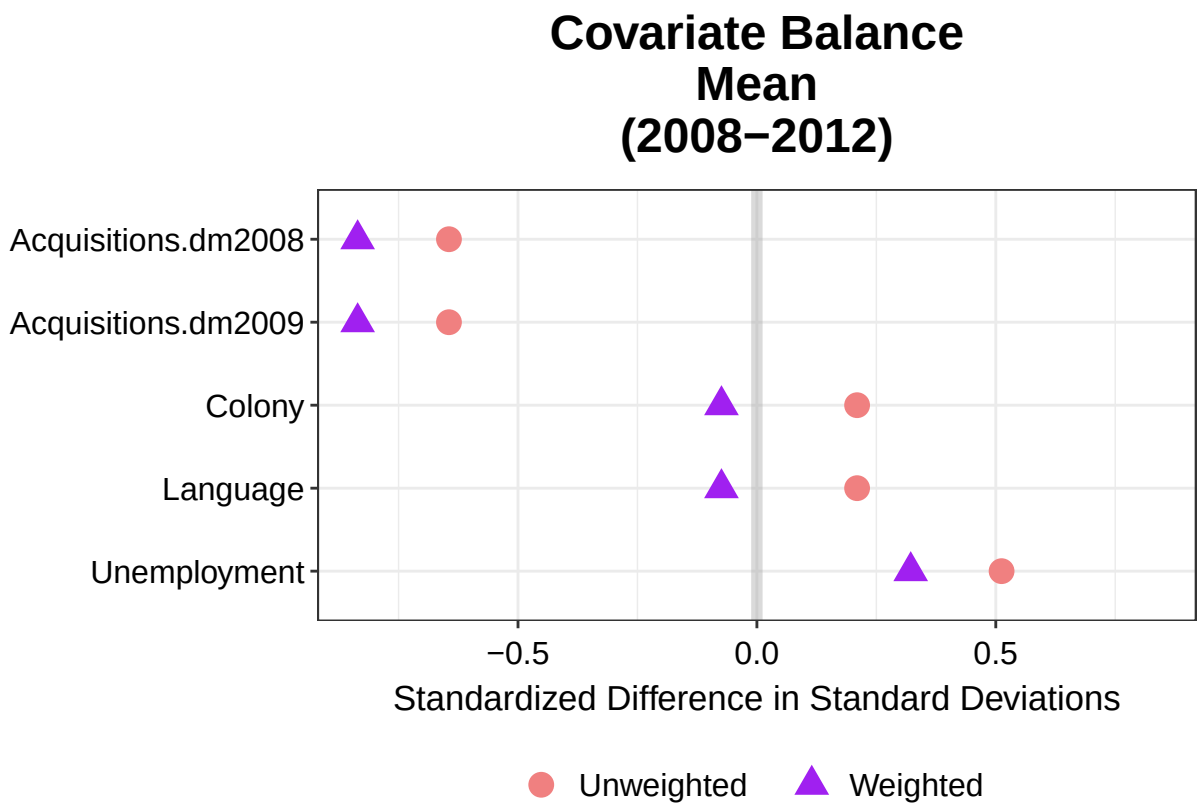



```
#Standardized Difference in Weighted and Unweighted Standard Deviations
#Top Right panel
sdm <- plot(out.mbal, type = "balance", stat = "sd",
            main = "Covariate Balance\nMean\n(2008-2012)")
```

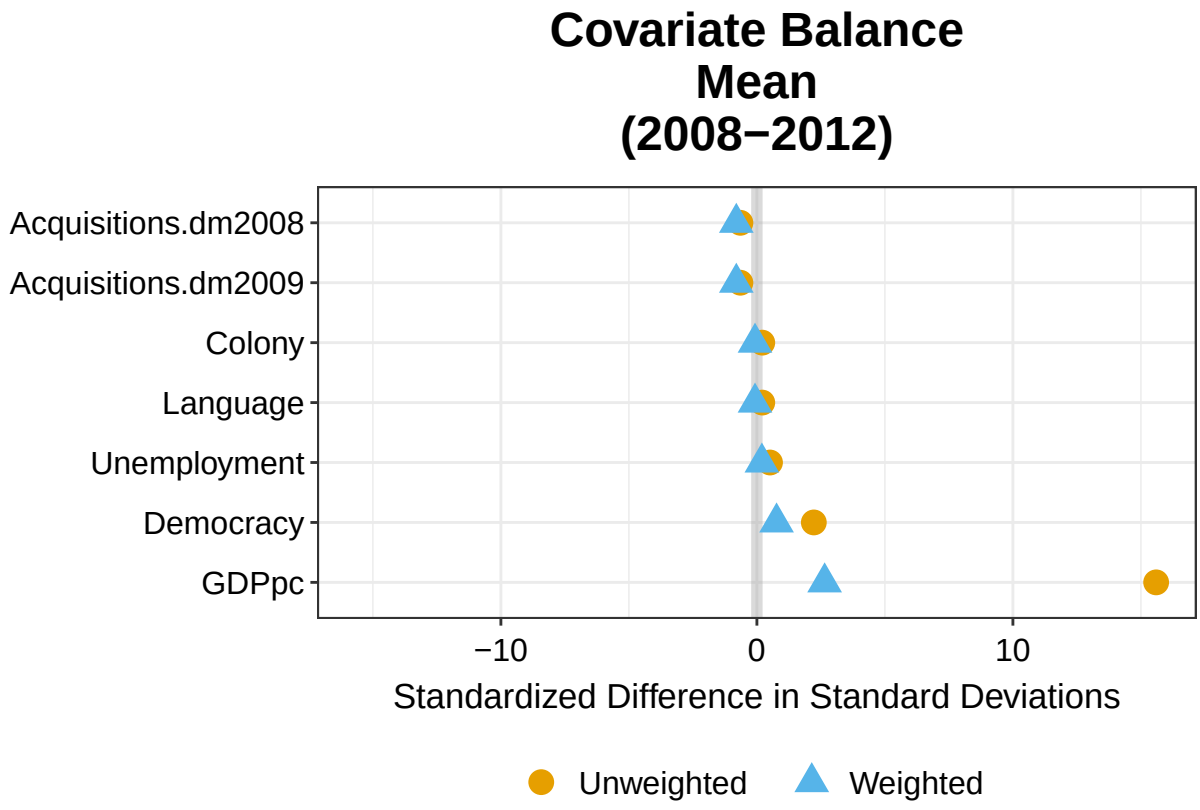



```
sdm + scale_color_manual(values = c("lightcoral", "purple"))
```

```
## Scale for colour is already present.
## Adding another scale for colour, which will replace the existing scale.
```

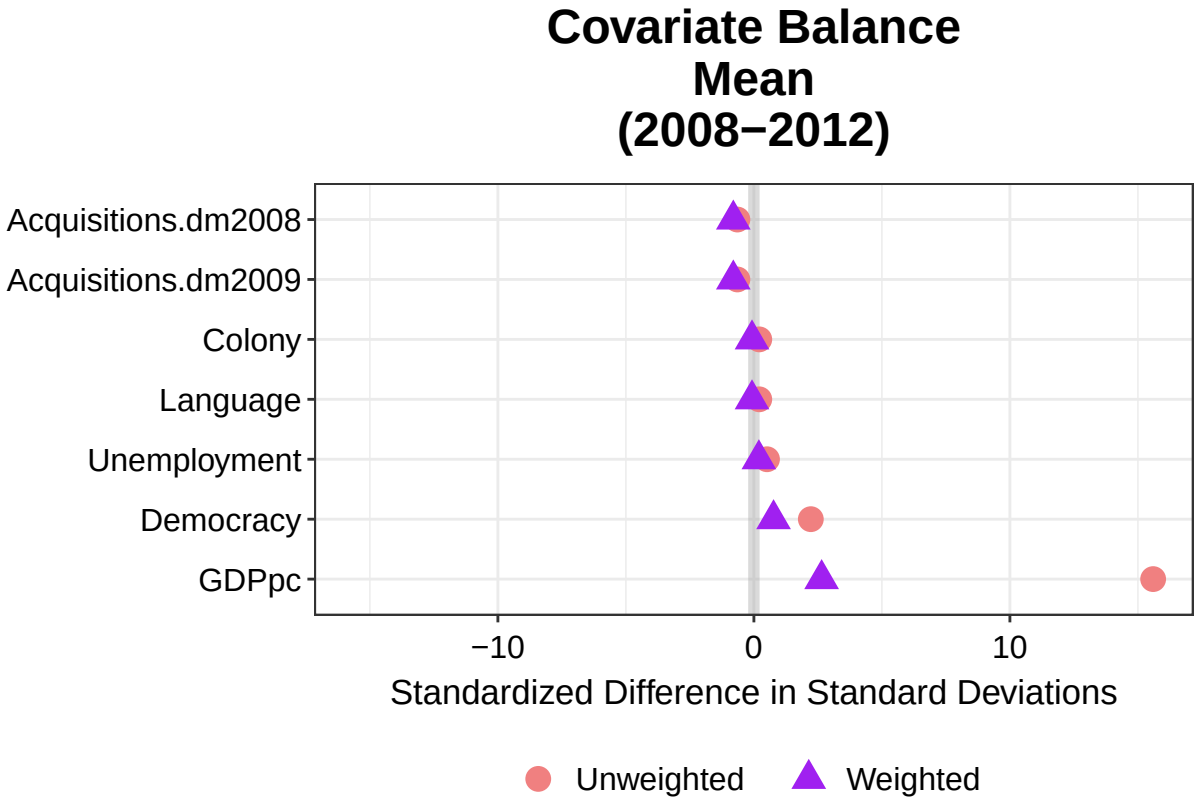



```
#Top Left panel
sdmMax <- plot(out.mbal1r, type = "balance", stat = "sd",
               main = "Covariate Balance\nMean\n(2008-2012)")
```

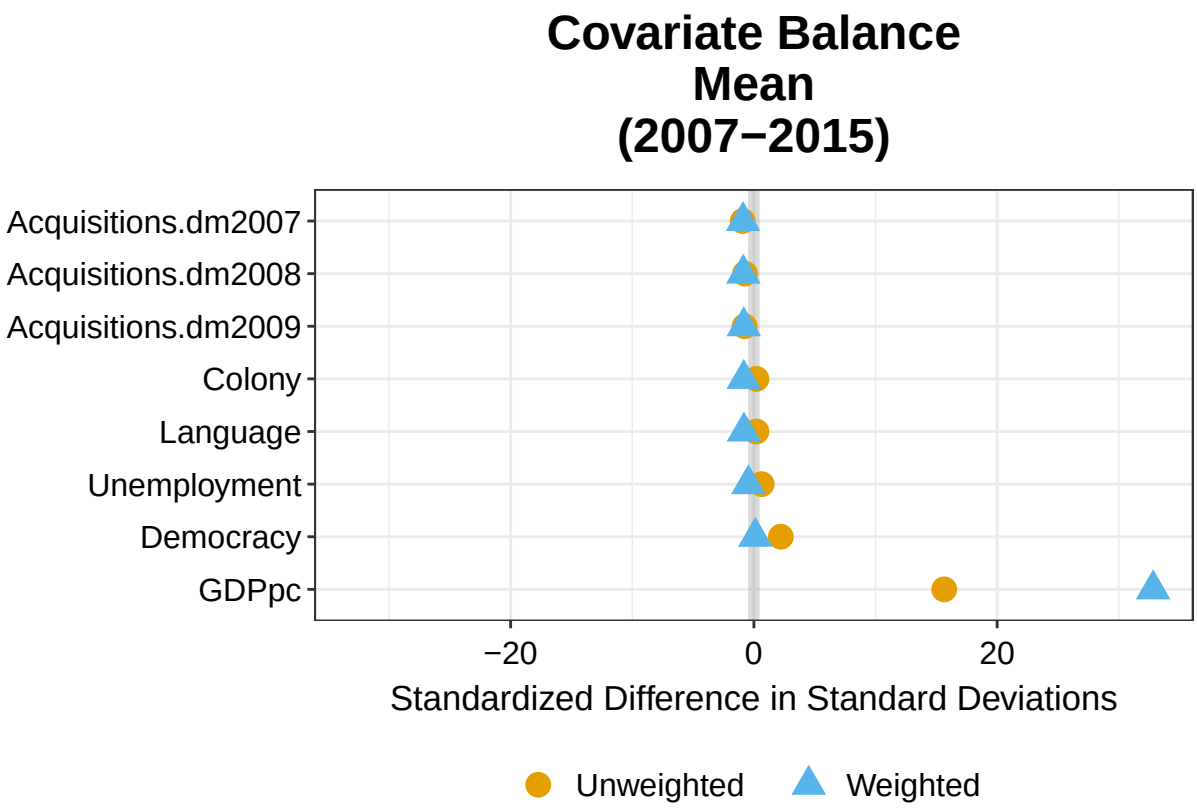


```
sdmMax + scale_color_manual(values = c("lightcoral", "purple"))
```

```
## Scale for colour is already present.
## Adding another scale for colour, which will replace the existing scale.
```

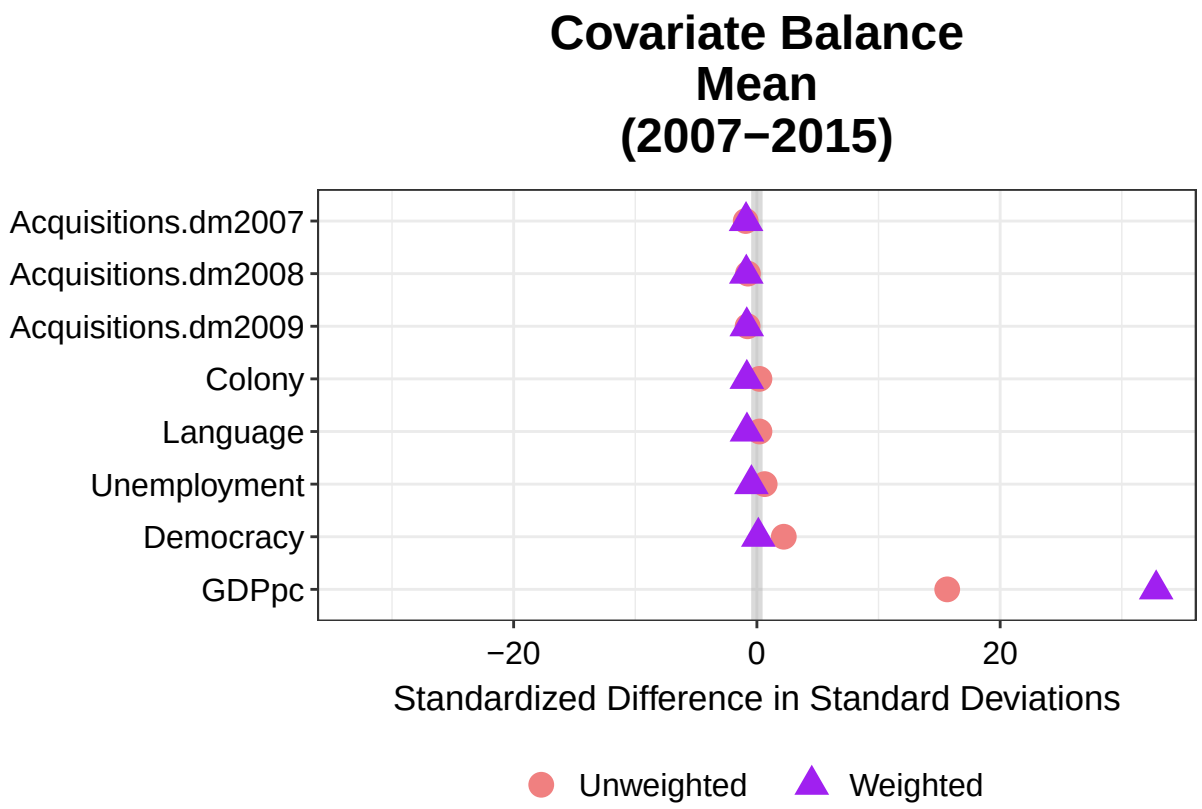



```
#Bottom Right
sdk <- plot(out.mbal7, type = "balance", stat = "sd",
            main = "Covariate Balance\nMean\n(2007-2015)")
```

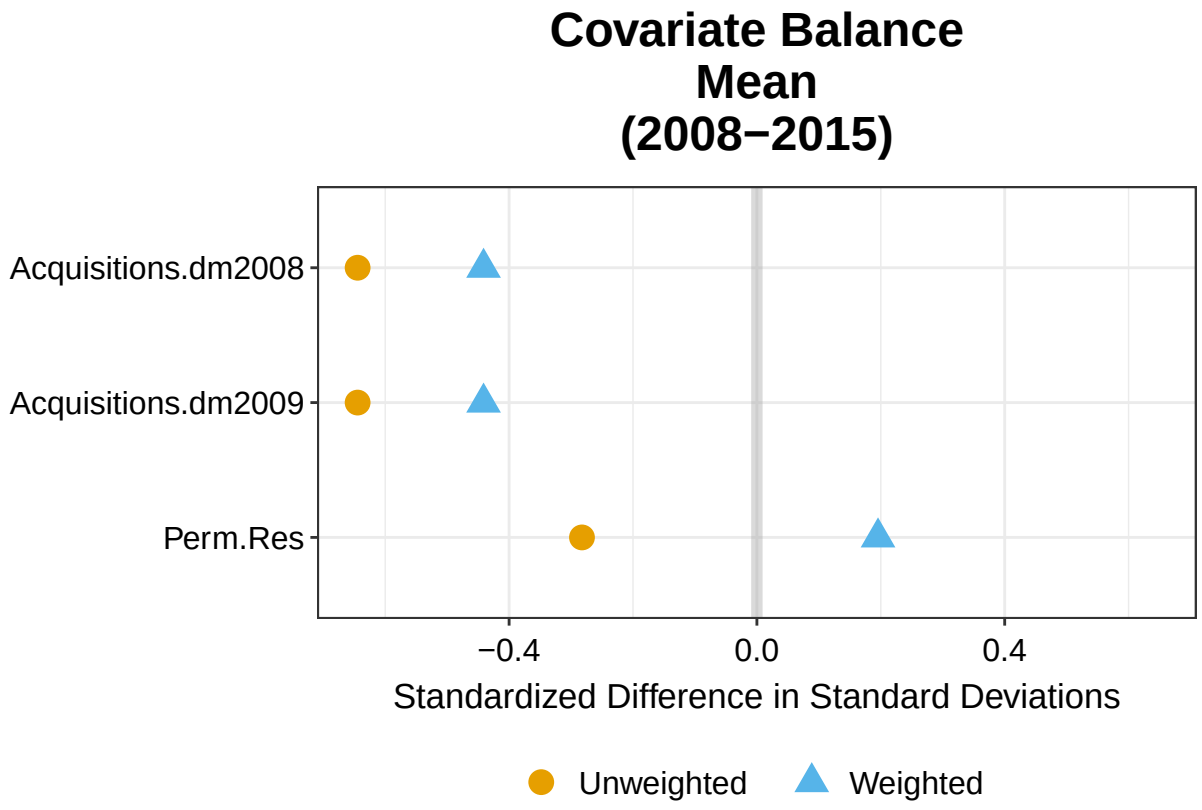



```
sdk + scale_color_manual(values = c("lightcoral", "purple"))
```

```
## Scale for colour is already present.
## Adding another scale for colour, which will replace the existing scale.
```

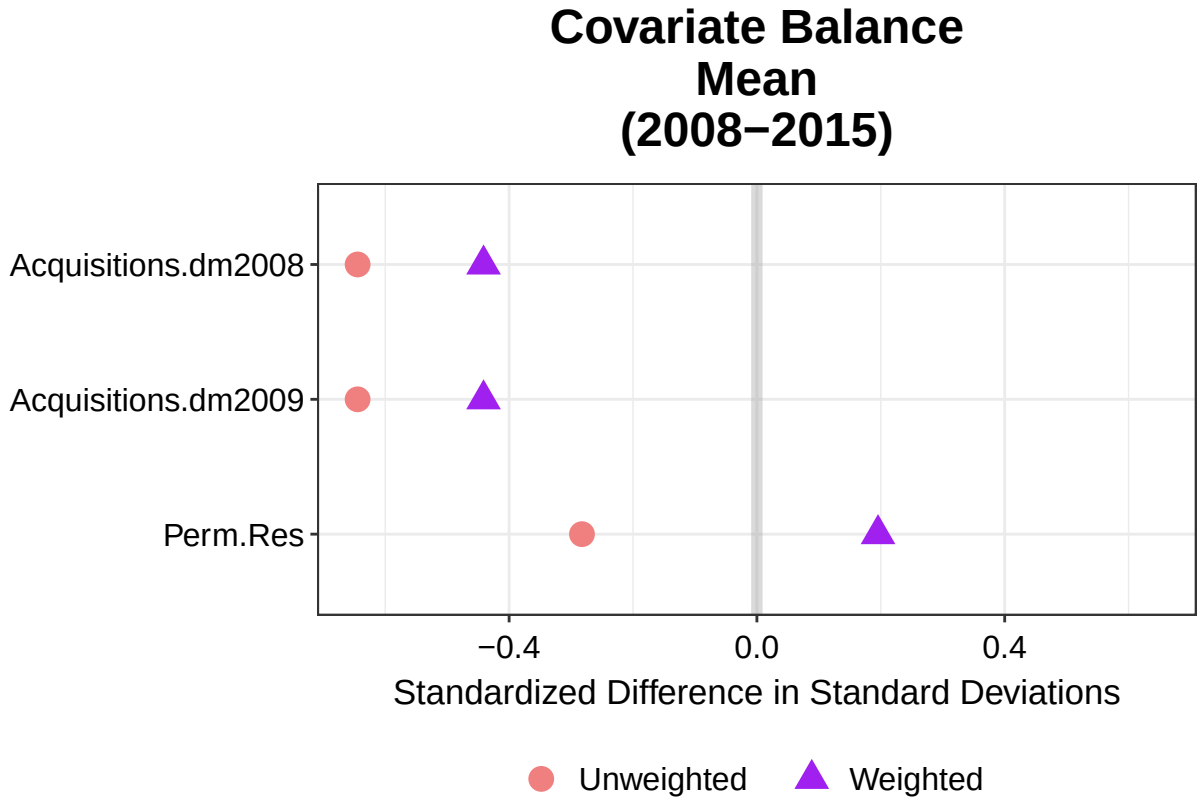



```
#Bottom Left
sdres <- plot(out.mbal15r2, type = "balance", stat = "sd",
              main = "Covariate Balance\nMean\n(2008-2015)")
```



```
sdres + scale_color_manual(values = c("lightcoral", "purple"))
```

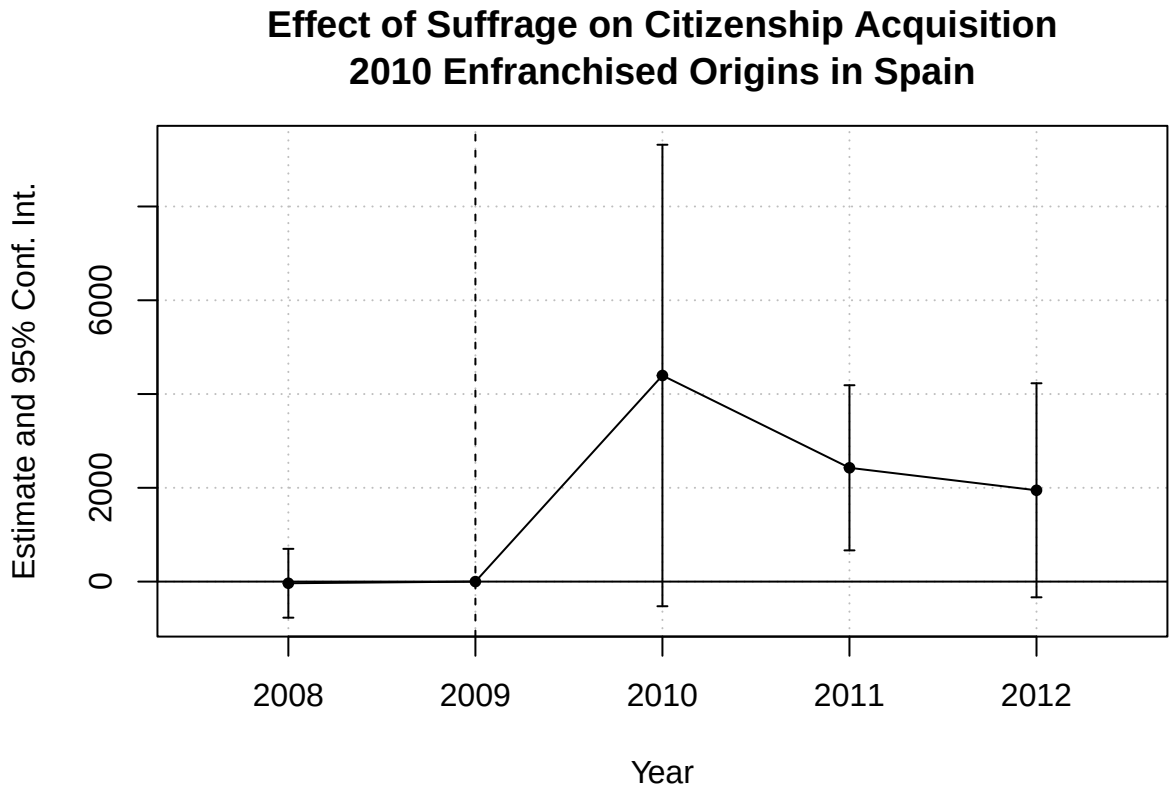
```
## Scale for colour is already present.
## Adding another scale for colour, which will replace the existing scale.
```




```
#Appendix C.1.6-------------------------------------------------------------
coefplot(M, drop = '(Intercept)',
         pt.join = TRUE, ref = c("2009" = 3), ref.line = TRUE,
         main = "Effect of Suffrage on Citizenship Acquisition\n2010 Enfranchised Origins in Spain",
         xlab = "Year",
         dict = c("year::2008"="2008", "year::2010"="2010",
                  "year::2011"="2011", "year::2012"="2012"))
```

	(1)
2008	−35.714
	(374.448)
2009	7416.429
	(4014.598)
2010	4395.429
	(2510.947)
2011	2426.857*
	(898.501)
2012	1947.286
	(1164.477)
Num.Obs.	35
AIC	759.7
RMSE	10 841.52
Std.Errors	by: scode

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001




```
modelsummary(M,
             stars = T,
             coef_map = c("2008","2009","2010","2011","2012"),
             gof_omit = 'DF|Deviance|R2|Log.Lik.|BIC|Std. Errors')
```