```
R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

>   #load libraries
>   library(data.table)
data.table 1.14.2 using 4 threads (see ?getDTthreads).  Latest news: r-datatable.com
>   library(dplyr)

Attaching package: ‘dplyr’

The following objects are masked from ‘package:data.table’:

    between, first, last

The following objects are masked from ‘package:stats’:

    filter, lag

The following objects are masked from ‘package:base’:

    intersect, setdiff, setequal, union

>   library(rdrobust)
>   library(lmtest)
Loading required package: zoo
```

```
Attaching package: ‘zoo’

The following objects are masked from ‘package:base’:

    as.Date, as.Date.numeric

>   library(ggplot2)
Warning message:
In (function (kind = NULL, normal.kind = NULL, sample.kind = NULL)  :
  non-uniform 'Rounding' sampler used
>   library(AER)
Loading required package: car
Loading required package: carData

Attaching package: ‘car’

The following object is masked from ‘package:dplyr’:

    recode

Loading required package: sandwich
Loading required package: survival
>   library(gridExtra)

Attaching package: ‘gridExtra’

The following object is masked from ‘package:dplyr’:

    combine

>   library(sandwich)
>   library(dummies)
dummies-1.5.6 provided by Decision Patterns

>   library(stringr)
>   library(foreign)
>   library(arcdiagram)
>
>
################################################################################
#############################################################
>
>
```

```
> ##### Functions (Always RUN before proceeding to Tables and Plots)
################################################################################
>
>   .Weighting <- function(xx,kernel,zz){
+
+     if(kernel > 0){
+       if(kernel == 1){xx$wght <- zz - abs(xx$run)}
+       if(kernel == 2){xx$wght <- 0.75*(1 - ((xx$run/zz)**2))}
+       if(kernel == 3){xx$wght <- (1/((2*pi)**2))*exp(-0.5*((abs(xx$run/zz)**2)))}
+       xx$wght[xx$t == 1] <- xx$wght[xx$t == 1] / sum(xx$wght[xx$t == 1])
+       xx$wght[xx$t == 0] <- xx$wght[xx$t == 0] / sum(xx$wght[xx$t == 0])
+     }else{xx$wght <- 1/nrow(xx)}
+     return(xx)
+   }
>
>   .RDplot <- function(data,cut,band,bins,degree,x.title,y.title,title,con,sz,brk) {
+
+     if(degree == 1){mm <- (var ~ Score2 )}
+     if(degree == 2){mm <- (var ~ Score2 + I(Score2**2) )}
+     if(degree == 3){mm <- (var ~ Score2 + I(Score2**2) + I(Score2**3) )}
+
+     data <- subset(data, !is.na(var))
+
+     data$Score2 <- data$run - cut
+     data <- subset(data, abs(Score2) <= band)
+     vec <- seq(cut, band, length.out =  (bins+1))
+     vec_mid <- vec[-length(vec)] + diff(vec)/2
+     data$count <- 1 ;     i <- 0
+
+
+     for (iii in 0:1){
+       cc <- iii
+       if(iii == 0) cc <- -1
+       aux <- subset(data, t == iii)
+       aux$groups <- as.factor(cut(cc*aux$Score2,breaks=vec, include.lowest = TRUE))
+       R <- group_by(aux,groups) %>% summarise (var = mean(var), count = sum(count) )
#%>% complete(groups)
+       #R <- subset(R, !is.na(Var))
+       R$t <- iii
+       R$Score2 <- cc*vec_mid
+
+       mdl <- lm(mm , data = aux)
+       pred <- predict(mdl, R, interval = "confidence", level = con, na.action = na.pass)
+       PRE <- pred[,1] ; R$pre <- PRE
```

```
+       predSE <- predict(mdl, R, interval = "confidence", level =con, na.action = na.pass,
se=TRUE)$se.fit
+       t <- ((pred[1,1]-pred[1,2])/predSE[1])
+       V <- as.matrix(vcov(mdl) )
+
+       if(degree == 1){X <- data.frame(1,R$Score2)}
+       if(degree == 2){X <- data.frame(1,R$Score2,(R$Score2^2))}
+       if(degree == 3){X <- data.frame(1,R$Score2,(R$Score2^2),(R$Score2^3))}
+
+       X <- as.matrix(X) ; se_robust <- sqrt(diag(X %*% V %*% t(X)))
+       LOW <- pred[,1] - t*se_robust ; HIGH <- pred[,1] + t*se_robust
+       R$low <- LOW ;    R$high <- HIGH
+
+       if(iii == 0){R0 <- R ; plot_data <- R}else{R1 <- R ;plot_data <- rbind(plot_data,R)}
+     }
+
+     plot_data <- data.frame(plot_data)
+     plot_data <- filter(plot_data, !is.na(count))
+     plot_data$count2 <- plot_data$count/sum(plot_data$count)
+     if(sz==0)  plot_data$count2 <- 1
+     ymx <- max(plot_data$var)
+     ymn <- min(plot_data$var)
+
+     ggplot(data=plot_data, aes(x=Score2, y=var, ymax = ymx, ymin = ymn))   +   theme_bw() +
+       theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+       geom_point( data=R0, aes(x=Score2, y=var, size=count), shape=16,  color = "gray35",
bg="gray95") + #aes(size = count2/4)
+       geom_point( data=R1, aes(x=Score2, y=var,size=count), shape=16, color = "gray65",
bg="gray95") +
+       geom_vline(xintercept = 0, colour="gray25", linetype = "longdash") +
+       geom_line(data=R1, aes(x=Score2, y=pre),colour="gray65",  size=1.5, alpha=.75) +
+       geom_line(data=R0, aes(x=Score2, y=pre),colour="gray35",  size=1.5, alpha=.75) +
+       ggtitle(title) +
+       theme(plot.title = element_text(family = "A", size=12))+
+       xlab(x.title) + ylab(y.title) +
+       theme(axis.text = element_text(colour="black", size=12, family="A")) +
+       theme(axis.title = element_text(colour="black", size=12, family="A")) +
+       geom_ribbon(data=R0, aes(ymin=low,ymax=high),alpha=0.10) +
+       geom_ribbon(data=R1, aes(ymin=low,ymax=high),alpha=0.10) +
+       scale_x_continuous(breaks=brk) +
+       theme(legend.position="none")
+
+
+   }
```

```
>
>
################################################################################
#####################################################################
>
>
> ##### Setup (Always RUN before proceeding to Tables and Plots)
################################################################################
##
>
>   # set active R directory here with setwd() function
>   setwd("C:/Users/afrey4/Dropbox/Papers/Legsize/replication1")
>
>    # set fonts for plots
>   windowsFonts( A=windowsFont("Liberation Sans")) ; par(family="A") ; sz <- 11
>
>   # Set shortcuts for regression formulas
>   base_for <- "var ~  vari + run + trun + factor(cut) "
>   base_for2 <- "var ~  vari + run + trun  + factor(cut) + factor(sta)"
>   base_for2a <- "var ~  vari + run + I(run^2) + trun  + (trun^2) + factor(cut) + factor(sta)"
>   base_for2b <- "var ~  vari + run + I(run^2)+I(run^3) + trun  + (trun^2) +(trun^3)+ factor(cut) +
factor(sta)"
>   cov1 <- "+ men + urb  + p_budget +hepc+ poor1 + poor2 + desert + able + gdp + metr  + liter
"
>   cov2 <- "  + base.p + p11 + p12 + p13 + p15 + p40 + p45 + p55  "
>   iv <- "| .-vari  + t"
>
>   # Main data file
>   maindata <- fread("data_main.csv")
>   maindata$trun <- maindata$t * maindata$run
>
>   # Creates dummies for the mayor's party
>   mainparties <- c(11,12,13,15,40,45,55)
>   for (i in 1:length(mainparties)){
+     newvar <- paste0("p",mainparties[i])
+     maindata$newvar <- ifelse(maindata$m.party == mainparties[i],1,0)
+     names(maindata)[ncol(maindata)] <- newvar
+   }
>
>   # Find optimal bandwidth for each polynomial
>   covars1 <- dummy(maindata$cut, sep = ".")
Warning message:
In model.matrix.default(~x - 1, model.frame(~x - 1), contrasts = FALSE) :
  non-list contrasts argument ignored
```

```
>   covars <- covars1[,2:(ncol(covars1))]
>
>   b1 <- rdbwselect(maindata$seats, maindata$run, c = 0, p=1, q=2,  kernel = "tri", bwselect =
"cerrd" , covs=covars)$bws
[1] "Mass points detected in the running variable."
>   b2 <- rdbwselect(maindata$seats, maindata$run, c = 0, p=2, q=4,  kernel = "tri", bwselect =
"cerrd" , covs=covars)$bws
[1] "Mass points detected in the running variable."
>   b3 <- rdbwselect(maindata$seats, maindata$run, c = 0, p=3, q=5,  kernel = "tri", bwselect =
"cerrd" , covs=covars)$bws
[1] "Mass points detected in the running variable."
>
>
>
################################################################################
#########################################################################
>
>
> ##### Figure 1 (Plot A): Local political coalitions by party in Brazil
#############################################################################
>
>
>   coal <- fread("data_coal2012.csv")
>
>
>   plist <- c("PT","PSB","PDT","PMDB","PSDB","PSD","PP")
>   coalx <- subset( coal, party_a %in% plist & party_b %in% plist )
>   coalx$a <- as.character(coalx$party_a)
>   coalx$b <- as.character(coalx$party_b)
>
>   coalx$d <- coalx$b
>   coalx$d[coalx$a < coalx$b] <- coalx$a[coalx$a < coalx$b]
>   coalx$e <- coalx$b
>   coalx$e[coalx$a > coalx$b] <- coalx$a[coalx$a > coalx$b]
>
>   coalx <- data.table(coalx)
>   coalx$c <- 1
>   coaly <- subset(coalx, d == e)
>   coalx <- subset(coalx, d != e)
>
>   coaly <- group_by(coaly,d) %>% summarise(c=sum(c))
>   coalx <- group_by(coalx,d,e) %>% summarise(c=sum(c))
`summarise()` has grouped output by 'd'. You can override using the `.groups` argument.
>   coalx$w <- coalx$c /100
```

```
>   coaly$w <- coaly$c /1000
>   coalx <- data.frame(coalx)
>   lab <- as.matrix(coalx[,1:2])
>   arcplot(lab,col.arcs = "gray75" , lwd.arcs= coalx$w,
+         ordering=plist,
+         cex.nodes=coaly$w, col.nodes="gray10",bg.nodes="gray60",pch.nodes=21,
col.labels="gray10",las=1,
+         lwd.nodes = 1.5)
>
>
>
>
################################################################################
#####################################################################
>
>
> ##### Figure 1 (Plot B): Local political coalitions by party in Brazil
#############################################################################
>
>   band <- round(b1[1],2)
>   datax <- subset(maindata, abs(run) <= band & cut <= 300 & g.support == 1)
>
>   # Aggregate the data at different quantiles based on the variable "g.share"
>   seq <- 0:10/10 ; seq[1] <- seq[1] - 0.00001
>   seq2 <- 1:10/10 - 0.05
>   datax$g1 <- cut(datax$g.share,breaks=seq) ;  datax$g2 <- cut(datax$p.share,breaks=seq)
>   datax$c <- 1
>   plot_data1 <- group_by(datax,g1) %>% summarise(obs=sum(c) ) ; plot_data1$Coalition <-
"Gubernatorial"
>   plot_data2 <- group_by(datax,g2) %>% summarise(obs=sum(c) ) ; plot_data2$Coalition <-
"Presidential"
>   names(plot_data1)[1] <- "g" ; names(plot_data2)[1] <- "g"
>   plot_data <- rbind(plot_data1,plot_data2)
>   plot_data$value <- c(seq2,seq2)
>   plot_data$obs <- plot_data$obs*200/sum(plot_data$obs)
>
>   ggplot(data=plot_data, aes(x=value, y=obs, group=Coalition))   +   theme_bw() +
+     theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+     geom_bar(data=plot_data, aes(x=value, y=obs, group=Coalition, fill=Coalition),
stat="identity", position_dodge()) +
+     xlab("Share of Parties in a Coincident Coalition") + ylab("Percentage of Cases") +
+     theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+     theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+     theme(legend.title = element_text(colour="black", size=sz, family="A")) +
```

```
+     scale_x_continuous(breaks = c(0,.2,.4,.6,.8,1))+
+     scale_fill_manual(values=c("gray25","gray75"))+
+     theme(legend.position = c(0.75,0.75))
>
##### Figure 2: Number of council seats in Brazilian municipalities
################################################################################
>
>   plot_data <- subset(maindata, pop < 310)
>   d <- data.frame(density(plot_data$pop)$x,density(plot_data$pop)$y)
>   names(d) <- c("x","y")
>   d$y <- d$y*200+8.01 # adjust the scale
>
>   # Generate the plot
>   ggplot(data=plot_data, aes(x=pop,y=seats))   +   theme_bw() +
+     theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+     geom_point( data=plot_data, aes(x=pop, y=seats), shape=20, size=1.5, color = "gray35",
bg="gray35") +
+     geom_ribbon(data=d, aes(x=x,y=y,ymax=y,ymin=8),fill="gray60", alpha=.5) +
+     geom_vline(xintercept = c(15,30,50,80,120,160,300), colour="gray25", linetype = "solid",
size=0.5) +
+     xlab("Population (thousand)") + ylab("Number of council seats") +
+     theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+     theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+     scale_x_continuous(breaks = c(15,30,50,80,120,160,300))
>
>
################################################################################
#################################################################
>
>
> ##### Figure 3: Discontinuity in council size after 2012
################################################################################
########
>
>   band <- round(b1[1],2)
>   datax <- subset(maindata, abs(run) < band & cut <= 300  )
>   datax <- .Weighting(datax,1,band)
>
>   datax$var1 <- (datax$seats)
>   reg <- lm(var1 ~  factor(cut) ,data=datax)
>   datax$var <- summary(reg)$res
>
>   .RDplot(datax,cut=0,band=band,bins=18,degree=1,"Population (thousand,
normalized)","Seats (normalized)","",
```

```
+          con=0.95,sz=12,brk=c(-3,-2,-1,0,1,2,3))
>
>
################################################################################
#########################################################################
>
>
> ##### Table 1: Loss of electoral strength by the local incumbent party
#############################################################################
>
>
>   data_aux <- maindata
>   data_aux$vari <- data_aux$seats
>   band <- round(b1[1],2)
>
>   res <- data.frame(matrix(0,12,3))
>
>
>   for (i in 1:3){
+     for(j in 1:4){
+
+       # Set different empirical specifications (includes or excludes covariates)
+       if(i == 1)  formul <- as.formula(paste(base_for2,cov1,cov2,iv,sep=""))
+       if(i == 2)  formul <- as.formula(paste(base_for2,cov1,iv,sep=""))
+       if(i == 3)  formul <- as.formula(paste(base_for2,iv,sep=""))
+
+       # Set different variables as the outcome
+       if(j == 1) {data_aux$s <- data_aux$g.support*data_aux$p.support*data_aux$f.support ;
data_aux$var <- data_aux$indx*100 ; min_fcand <- 1}
+       if(j == 2) {data_aux$s <- data_aux$g.support ; data_aux$var <- data_aux$g.pc*100 ;
min_fcand <- 0}
+       if(j == 3) {data_aux$s <- data_aux$p.support ; data_aux$var <- data_aux$p.pc*100 ;
min_fcand <- 0}
+       if(j == 4) {data_aux$s <- data_aux$f.support ; data_aux$var <- data_aux$f.pc*100 ;
min_fcand <- 1}
+
+       datax <- subset(data_aux, abs(run) <= band & s == 1 & f.cand > min_fcand & cut <= 300
)
+       datax <- .Weighting(datax,1,band)
+
+       reg <- ivreg(formul  , weights=wght, data=datax)
+       r <- coeftest(reg,  vcov = vcovHC(reg), type="HC3")[2,]
+
+       tv <- abs(r[3])
```

```
+      stars <- "" ; if(tv >= 1.645){stars <- "†"} ; if(tv >= 1.96){stars <- "*"}
+      coeff <- round(r[1],3) ; sd <- round(r[2],3)
+
+      res[j*3-2,i] <- paste(format(coeff,nsmall=3),stars,sep="")
+      res[j*3-1,i] <- paste("(",format(sd,nsmall=3),")",sep="")
+
+
+      res[j*3,i] <- nrow(datax)
+
+    }
+  }
>
>
>  # For pasting in LatEx
>  for (i in 1:ncol(res)){
+    fake <- data.frame(matrix("&",nrow(res),1))
+    x <- cbind(fake,res[,i])
+    if(i == 1) {xx <- x}else{xx <- cbind(xx,x)}
+  }
>  last <-  c("\tabularnewline","\tabularnewline","\tabularnewline[0.25cm]")
>
>  xx[,ncol(xx)+1] <- c(last,last,last,last)
>  row.names(xx) <- c("Vote Share Index","(aggregates the elections below) ","Observations
",
+                   "Gubernatorial ","(2014)","Observations","Presidential ","(2014) ","Observations
",
+                   "Mayoral","(2016)","Observations   ") ;  names(xx) <- NULL
>  print(xx)

Vote Share Index                & -4.674* & -4.761* & -4.218*         \tabularnewline
(aggregates the elections below)  & (1.503) & (1.559) & (1.506)        \tabularnewline
Observations                    &    1114 &    1114 &    1114 \tabularnewline[0.25cm]
Gubernatorial                   & -4.936* & -4.904* & -4.573*         \tabularnewline
(2014)                          & (2.266) & (2.324) & (2.278)         \tabularnewline
Observations                    &    1305 &    1305 &    1305 \tabularnewline[0.25cm]
Presidential                    & -4.648* & -4.181† &  -3.918         \tabularnewline
(2014)                          & (2.005) & (2.492) & (2.465)         \tabularnewline
Observations                    &    1307 &    1307 &    1307 \tabularnewline[0.25cm]
Mayoral                         & -3.848* & -3.714* & -3.089†         \tabularnewline
(2016)                          & (1.877) & (1.892) & (1.836)         \tabularnewline
Observations                    &    1116 &    1116 &    1116 \tabularnewline[0.25cm]
>  rm(fake,x,data_aux,datax)
>
>
```

```
>
################################################################################
#####################################################################
>
>
> ##### Table 2: Effect of council size on the 2012 municipal election
################################################################################
>
>   data_aux <- maindata
>
>   data_aux$v.candi <- data_aux$v.cand/data_aux$seats
>   data_aux$vari <- data_aux$seats
>   data_aux$lsize <- log(data_aux$size)
>   data_aux$major.m <- data_aux$m.team/(data_aux$seats)
>   data_aux$major.g <- data_aux$mg.team/data_aux$m.team
>   data_aux$major.p <- data_aux$mp.team/data_aux$m.team
>
>   lista_var <-
c("v.parties","v.candi","lsize","m.parties.elec","m.team","major.m","g.share","p.share")
>
>   band <- round(b1[1],2)
>
>   res <- data.frame(matrix(0,length(lista_var)*2+1,3))
>
>   for (i in 1:3){
+     for(j in 1:length(lista_var)){
+
+
+       if(i == 1)  formul <- as.formula(paste(base_for2,cov1,cov2,iv,sep=""))
+       if(i == 2)  formul <- as.formula(paste(base_for2,cov1,iv,sep=""))
+       if(i == 3)  formul <- as.formula(paste(base_for2,iv,sep=""))
+
+       data_aux$var <- data_aux %>% select(lista_var[j])
+
+       data_aux$s <- data_aux$g.support*data_aux$p.support
+       datax <- subset(data_aux, abs(run) <= band & s == 1 & f.cand > 0 & cut <= 300 )
+       datax <- .Weighting(datax,1,band)
+
+       reg <- ivreg(formul  , weights=wght, data=datax)
+       r <- coeftest(reg,  vcov = vcovHC(reg), type="HC3")[2,]
+
+       tv <- abs(r[3])
+       stars <- "" ; if(tv >= 1.645){stars <- "???"} ; if(tv >= 1.96){stars <- "*"}
+       coeff <- round(r[1],3) ; sd <- round(r[2],3)
```

```
+
+      res[j*2-1,i] <- paste(format(coeff,nsmall=3),stars,sep="")
+      res[j*2,i] <- paste("(",format(sd,nsmall=3),")",sep="")
+
+    }
+  }
>   res[17,] <- nrow(datax)
>
>
>   #### For pasting in latex
>   for (i in 1:ncol(res)){
+     fake <- data.frame(matrix("&",nrow(res),1))
+     x <- cbind(fake,res[,i])
+     if(i == 1) {xx <- x}else{xx <- cbind(xx,x)}
+   }
>   last <-  c("\tabularnewline","\tabularnewline[0.25cm]")
>
>   xx[,ncol(xx)+1] <- c(last,last,last,last,last,last,last,last,"\tabularnewline[0.25cm]")
>   row.names(xx) <- c("Total parties running","(number of parties) ","Total candidates
running","(per seat) ","Total parties in winning coalition","(log) ",
+                      "Coalition parties elected","(number of parties)    ", "Coalition councilors
elected","(number of councilors)",
+                      "Coalition share elected","(share of total seats)",
+                      "Gubernatorial alignment ","(share of coincident coalition)","Presidential
alignment","(share of coincident coalition) ",
+                      "Observations") ;  names(xx) <- NULL
>   print(xx[1:6,]) ; print(xx[7:12,]) ; print(xx[13:17,])

Total parties running              &   0.233 &   0.232 &   0.208         \tabularnewline
(number of parties)                & (0.386) & (0.384) & (0.386) \tabularnewline[0.25cm]
Total candidates running           &  -0.143 &  -0.139 &  -0.157         \tabularnewline
(per seat)                         & (0.240) & (0.241) & (0.254) \tabularnewline[0.25cm]
Total parties in winning coalition &   0.060 &   0.059 &   0.068         \tabularnewline
(log)                              & (0.055) & (0.055) & (0.055) \tabularnewline[0.25cm]

Coalition parties elected    &  0.433* &  0.430* &  0.471*         \tabularnewline
(number of parties)          & (0.159) & (0.159) & (0.159) \tabularnewline[0.25cm]
Coalition councilors elected &  0.473* &  0.478* &  0.520*         \tabularnewline
(number of councilors)       & (0.208) & (0.207) & (0.208) \tabularnewline[0.25cm]
Coalition share elected      &  -0.002 &  -0.002 &   0.002         \tabularnewline
(share of total seats)       & (0.020) & (0.020) & (0.020) \tabularnewline[0.25cm]

Gubernatorial alignment          &  -0.027 &  -0.025 &  -0.024         \tabularnewline
(share of coincident coalition)  & (0.026) & (0.026) & (0.025) \tabularnewline[0.25cm]
```

```
Presidential alignment           &  -0.003 &  -0.003 &  -0.004         \tabularnewline
(share of coincident coalition)  & (0.019) & (0.022) & (0.022) \tabularnewline[0.25cm]
Observations                     &    1305 &    1305 &    1305 \tabularnewline[0.25cm]
>   rm(fake,x,data_aux,datax)
>
>
>
################################################################################
#########################################################################
>
>
> ##### Figure 4: Heterogeneity in reverse coattails: the role of unaligned electoral incentives
################################################
>
>   data_aux <- maindata
>
>   data_aux$vari <- data_aux$seats
>
>   # set formulas for the interacted effects
>   base_for_i <- "var ~  vari + ivari  + i + run + trun + irun + itrun + factor(cut) + factor(sta)"
>   iv_i <- "| .-vari -ivari  + t + it"
>   formul <- as.formula(paste(base_for_i,cov1,"+ log(size)",cov2,iv_i,sep=""))
>
>   band <- round(b1[1],2)
>
>   res <- data.frame(matrix(0,6,2))
>
>   for (i in 1:2){
+
+     # Set different variables as the outcome
+     if(i == 1) {data_aux$s <- data_aux$g.support ; data_aux$var <- data_aux$g.pc*100 ;
data_aux$share <- data_aux$g.share}
+     if(i == 2) {data_aux$s <- data_aux$p.support ; data_aux$var <- data_aux$p.pc*100 ;
data_aux$share <- data_aux$p.share}
+
+     datax <- subset(data_aux, abs(run) <= band & s == 1 & f.cand > 0 & cut <= 300  )
+     datax <- .Weighting(datax,1,band)
+
+     # Create dummies to split the sample into High and Low shares (see the text)
+     datax$j <- summary(lm(datax$share ~ factor(datax$sta)+factor(datax$cut) ))$res #use
variation within state / assignment window
+     med <- round(median(datax$j[!is.na(datax$j)]),5)
+     datax$i <- ifelse(datax$j >= med,1,0)
+
```

```
+     # Create the interaction terms
+     datax$ivari <- datax$i*datax$vari
+     datax$it <- datax$i*datax$t
+     datax$itrun <- datax$i*datax$t*datax$run
+     datax$irun <- datax$i*datax$run
+
+     reg <- ivreg(formul  , weights=wght, data=datax)
+     results <- coeftest(reg,  vcov = vcovHC(reg, type="HC3"))[2:3,]
+     v <- vcovHC(reg, type="HC3")
+
+     res[i*3-2,] <- results[1,1:2]
+     res[i*3,] <- results[2,1:2]
+     res[i*3-1,1] <- results[1,1]  + results[2,1]
+     res[i*3-1,2] <- (v[2,2]  + v[3,3]  + 2*v[2,3] )^.5
+
+   }
>
>   res$Coalition <- "Gubernatorial"
>   res$Coalition[4:6] <- "Presidential"
>
>   res$high <- res$X1+1.96*res$X2 ; res$low <- res$X1-1.96*res$X2
>   nam <- c("Low share","High share","Difference")
>   res$names <- c(1,2,3,1,2,3)
>
>   ggplot(res, aes(x=names, y=X1, ymin=low, ymax=high, group=Coalition)) + theme_bw() +
+     theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+     geom_pointrange(aes(color=Coalition, group=Coalition), position = position_dodge(width =
1/2), shape=16) +
+     scale_color_manual(values=c("gray25","gray75"), name="2014 Election") +
+     scale_y_continuous(breaks=c(-30,-15,0,15,30)) +
+     scale_x_continuous(breaks=c(1,2,3), labels=nam)+
+     geom_hline(yintercept=0, linetype=2)+
+     theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+     theme(axis.title.x = element_blank()) +
+     theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+     theme(legend.text = element_text(colour="black", size=sz, family="A")) +
+     theme(legend.title = element_text(colour="black", size=sz, family="A")) +
+     ylab("Vote percentage")+
+     theme(legend.position = c(0.75,0.15))
>
>
>
```

```
>
################################################################################
#####################################################################
>
##### Figure 5: Change in winning mayoral coalitions between 2012 and 2016
#####################################################################
>
>   data_aux <- maindata
>
>   data_aux$vari <- data_aux$seats
>
>   band <- round(b1[1],2)
>   data_aux$m.base <- (data_aux$m.team )/data_aux$seats       # number of council members
in base
>   data_aux$f.base <- (data_aux$f.team )/data_aux$seats       # number of council members in
future coalition
>   data_aux$m.base.run <- (data_aux$mr.team)/data_aux$seats   # number of council
members in base, rerunning
>   data_aux$f.base.run <- (data_aux$fr.team)/data_aux$seats   # number of council members
in future coalition, rerunning
>
>
>   vec <- c("m.base","f.base","m.base.run","f.base.run","f.share","extreme.1","extreme.2")
>   res <- data.frame(matrix(0,length(vec),2))
>
>   for (i in 1:nrow(res)){
+
+     # set variable to be used
+     num <- match(vec[i],names(data_aux))
+     data_aux$var <- data_aux[[num]]
+
+     # set formula
+     formul2 <- as.formula(paste(base_for2,cov1,"+log(size) ",cov2,iv,sep=""))
+
+     datax <- subset(data_aux, abs(run) <= band & cut <= 300 & f.cand > 1 & f.support >= 1 &
!is.na(var) & abs(var) != Inf  )
+     datax <- .Weighting(datax,1,band)
+
+     reg <- ivreg(formul2  , weights=wght, data=datax)
+     r <- coeftest(reg,  vcov = vcovHC(reg, type="HC3"))[2,]
+
+     res[i,1] <- r[1]
+     res[i,2] <- r[2]
+
```

```
+  }
>  res$high <- res$X1+1.96*res$X2 ; res$low <- res$X1-1.96*res$X2
>
>  names <- c("Share of \n coalition \n members \n in 2012 \n (post-election)",
+           "Share of \n coalition \n members \n in 2016 \n (pre-election)",
+           "Share of \n coalition \n members \n in 2012 \n (post-election) \n (reelec. only)",
+           "Share of \n coalition \n members \n in 2016 \n (pre-election) \n (reelec. only)",
+           "Common partners: \n 2016 vs. 2012 \n (parties)",
+           "Ideological \n Incoherence \n 2016 / 2012 (I)",
+           "Ideological \n Incoherence \n 2016 / 2012 (II)")
>
>  res$names <- 1:7
>  res <- res[order(-res$names),]
>  res$names <- as.factor(res$names)
>  ggplot(res, aes(x=names, y=X1, ymin=low, ymax=high)) + theme_bw() +
+   theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+   geom_pointrange(shape=16) +
+   scale_x_discrete(labels=names)+
+   geom_hline(yintercept=0, linetype=2)+
+   theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+   theme(axis.title.x = element_blank()) +
+   theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+   ylab("Effect")
>
>
>
>
>
>
>
################################################################################
#####################################################################
>
>
> ##### Figure 6: Policy outcomes in 2013-2016
################################################################################
####################
>
>  data_aux <- maindata
>  data_aux$vari <- data_aux$seats
>
>  band <- round(b1[1],2)
>
>  vec <- c("sch1","sch2","sch3","hea1","hea2","tran","lf.budget","p1","p2","p3","p4","lf.budget")
```

```
>
>   res <- data.frame(matrix(0,length(vec),2))
>   i <- 1
>   for (i in 1:nrow(res)){
+
+     num <- match(vec[i],names(data_aux))
+     data_aux$var <- data_aux[[num]]
+
+     formul2 <- as.formula(paste(base_for2,cov1,cov2,iv,sep=""))
+
+     datax <- subset(data_aux, abs(run) <= band & cut <= 300 &  !is.na(var)    )
+     datax <- .Weighting(datax,1,band)
+
+     if(i >= 12) {
+       datax$var <- abs(summary(lm(var ~ t ,  data=datax) )$res )
+     }
+
+     reg <- ivreg(formul2  , weights=wght, data=datax)
+     r <- coeftest(reg,  vcov = vcovHC(reg, type="HC3"))[2,]
+
+     res[i,1] <- r[1]
+     res[i,2] <- r[2]
+
+   }
>
>   res$coeff <- res$X1/res$X2
>   res$high <- res$coeff+1.96 ; res$low <- res$coeff-1.96
>
>   nam <- c("Preschool \n enrollment","Elementary \n enrollment","Middle School \n
enrollment",
+            "Infant \n mortality","Health \n visits","Federal \n discretionary transfers",
+            "Total budget","Budget in \n Health","Budget in \n Education","Budget in \n
Urbanization","Budget in \n Security",
+            "Spending Volatility")
>
>   res$names <- 1:length(vec)
>   ggplot(res, aes(x=names, y=coeff, ymin=low, ymax=high)) + theme_bw() +
+     theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+     geom_pointrange(shape=16) +
+     scale_x_continuous(breaks=1:length(vec), labels=nam)+
+     geom_hline(yintercept=0, linetype=2)+
+     coord_flip() +
+     theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+     theme(axis.title.y = element_blank()) +
```

```
+     theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+     ylab("Normalized Effect")
>
>
>
################################################################################
#####################################################################
>
>
> ##### Figure 7: Effects on the profile of elected politicians
################################################################################
####
>
>
>   data_aux <- maindata
>   data_aux$vari <- data_aux$seats
>
>   band <- round(b1[1],2)
>
>   vec <-  c("mean.edu2","mean.age","mean.sex","mean.ide","mean.recent")
>
>   res <- data.frame(matrix(0,length(vec),2))
>
>   for (i in 1:nrow(res)){
+
+     num <- match(vec[i],names(data_aux))
+     data_aux$var <- data_aux[[num]]
+
+     formul2 <- as.formula(paste(base_for2,cov1,cov2,iv,sep=""))
+
+     datax <- subset(data_aux, abs(run) <= band & cut <= 300 & f.cand > 1 & f.support == 1 &
!is.na(var) & abs(var) != Inf  )
+     datax <- .Weighting(datax,1,band)
+
+     reg <- ivreg(formul2  , weights=wght, data=datax)
+     r <- coeftest(reg,  vcov = vcovHC(reg, type="HC3"))[2,]
+
+     res[i,1] <- r[1]
+     res[i,2] <- r[2]
+
+
+   }
>
>   res$coeff <- res$X1/res$X2
```

```
>   res$high <- res$coeff+1.96 ; res$low <- res$coeff-1.96
>
>   res$names <- c("Education","Age","Gender","Left Ideology","Newcomer to Politics")
>   plot1 <- ggplot(res, aes(x=names, y=coeff, ymin=low, ymax=high)) + theme_bw() +
+     theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+     geom_pointrange(shape=16) +
+     ggtitle("Council's profile")+
+     geom_hline(yintercept=0, linetype=2)+
+     coord_flip() +
+     theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+     theme(axis.title.y = element_blank()) +
+     theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+     theme(plot.title = element_text(colour="black", size=sz, family="A")) +
+     ylab("Normalized Effect")
>
>
>   vec <- c("edu","age","sex","job","lame_duck")
>
>   res <- data.frame(matrix(0,length(vec),2))
>
>   for (i in 1:nrow(res)){
+
+     num <- match(vec[i],names(data_aux))
+     data_aux$var <- data_aux[[num]]
+
+     formul2 <- as.formula(paste(base_for2,cov1,cov2,iv,sep=""))
+
+     datax <- subset(data_aux, abs(run) <= band & cut <= 300 & f.cand > 1 & f.support == 1 &
!is.na(var) & abs(var) != Inf  )
+     datax <- .Weighting(datax,1,band)
+
+     reg <- ivreg(formul2  , weights=wght, data=datax)
+     r <- coeftest(reg,  vcov = vcovHC(reg, type="HC3"))[2,]
+
+     res[i,1] <- r[1]
+     res[i,2] <- r[2]
+
+
+   }
>   res$coeff <- res$X1/res$X2
>   res$high <- res$coeff+1.96 ; res$low <- res$coeff-1.96
>
>   res$names <- c("Education","Age","Gender","Public Servant","Reelected Incumbent")
>   plot2 <- ggplot(res, aes(x=names, y=coeff, ymin=low, ymax=high)) + theme_bw() +
```

```
+     theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+     geom_pointrange(shape=16) +
+     ggtitle("Mayor's profile")+
+     geom_hline(yintercept=0, linetype=2)+
+     coord_flip() +
+     theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+     theme(axis.title.y = element_blank()) +
+     theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+     theme(plot.title = element_text(colour="black", size=sz, family="A")) +
+     ylab("Normalized Effect")
>
>
>
>   grid.arrange( plot1 , plot2 ,nrow=1)
>
>
>
################################################################################
#########################################################################
>
 ##### Figure 8: Electoral and policy preferences of voters
################################################################################
#####
>
>   ### This Figure is only replicated with data from the LAPOP survey,
>   # which needs to be downloaded from their website:
>
>   # http://datasets.americasbarometer.org/database/index.php?freeUser=true
>   # file name: 54861031Brazil LAPOP AmericasBarometer 2012 Rev1_W.DTA
>   # Please rename the file: Brazil_LAPOP_2012.dta
>
>
>   ### LAPOP'S STATEMENT ON THE DISTRIBUTION OF DATASETS
>   #Replication requests should refer to the LAPOP websites where the data can be
downloaded.
>   #Authors should provide syntax/code that allows for replication of their analyses.
>   #Authors are not permitted to transfer LAPOP-obtained data sets to third parties,
>   #even journals, as that would violate the terms of the "click license."
>
>
>   ### with the LAPOP file above...
>   data_aux <-
read.dta(file="C:/Users/afrey4/Dropbox/Data/Barometer/Brazil_LAPOP_2012.dta")
>   data_aux <- data_aux[,c("municipio","soc2a","soc2b")]
```

```
>   data_aux$ibge <- as.numeric(substring(data_aux$municipio,3,8))
>
>   # Get data from the 2010 Presidential election
>   elec <- fread("data_elec2010.csv")
>
>   datax <- subset(data_aux, !is.na(soc2a) &!is.na(soc2b)  )
>
>   # Merge datasets
>   datax <- merge(datax, elec, by="ibge")
>
>   datax$HEC <- ifelse(datax$soc2a %in% c("Saúde") | datax$soc2b %in% c("Saúde") ,1,0)
>   datax$c <- 1
>
>   # Aggregate the data at the level of municipality
>   gg <- group_by(datax, municipio) %>% summarise(hec=mean(HEC), pt=mean(pc) ,
tot=sum(c))
>   gg <- subset(gg, tot>=10)
>   gg1 <- gg[,c("municipio","pt")] ; names(gg1)[2] <- "y"
>   gg1$g <- "Votes for the presidential incumbent (PT)"
>   gg2 <- gg[,c("municipio","hec")] ;names(gg2)[2] <- "y"
>   gg2$g <- "Report policy preference for Health"
>   plot_data <- rbind(gg1,gg2)
>
>
>   ggplot(plot_data, aes(x=y, group=g)) + theme_bw() +
+     theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())+
+     geom_density( aes(group=g, fill=g) , size=1, alpha=0.6)+
+     scale_fill_manual(values=c("gray10","gray90"), name="") +
+     theme(axis.text = element_text(colour="black", size=sz, family="A")) +
+     theme(legend.text = element_text(colour="black", size=sz, family="A")) +
+     theme(legend.title  = element_text(colour="black", size=sz, family="A")) +
+     theme(axis.title = element_text(colour="black", size=sz, family="A")) +
+     xlab("Share of respondents in the municipality that:") + ylab("Data density") +
+     theme(legend.position="bottom")
>
>
>
##############################################################################
#########################################################################
>
```