Online Appendix for:

"congressbr: An R Package for Analyzing Data from Brazil's Chamber of Deputies and Federal Senate"*

Robert Myles McDonnell   Guilherme Jardim Duarte   Danilo Freire

28 August 2018

1   Introduction

In this appendix, we present R code to reproduce the numeral analyses included in our article "congressbr: An R Package for Analyzing Data from Brazil's Chamber of Deputies and Federal Senate," published in the *Latin American Research Review*. Please contact the authors should you require any assistance with the code.

2   Installation

All required packages are available for download on the Comprehensive R Archive Network (CRAN). Please run the commands below to install them:

***McDonnell**: Data scientist, First Data Corporation. Email address: robertmylesmcdonnell@gmail.com. **Duarte**: Data editor, JOTA. Email address: gjardimduarte@gmail.com. **Freire**: Postdoctoral Research Associate, Political Theory Project, Brown University. Email address: danilofreire@brown.edu. Corresponding author. The current stable version is available on CRAN at https://cran.r-project.org/web/packages/congressbr. You can install the latest build of congressbr from its GitHub code page at https://github.com/RobertMyles/congressbr. Comments and suggestions are most welcome and can be submitted either by contacting the authors or by opening an issue at https://github.com/RobertMyles/congressbr/issue

```r
install.packages('tidyverse')

install.packages('congressbr')

install.packages('MCMCpack')

install.packages('lubridate')
```

3 Number of Parties in Congress by Year

With congressbr, one can obtain the number of parties in the Chamber of Deputies or in the Senate with only a few lines of code. Firstly, we load the necessary packages:

```r
library(congressbr)

library(dplyr)

library(ggplot2)

library(lubridate)
```

Next, we will use a pre-processed dataframe included in congressbr with nominal votes in Brazil's Chamber of Deputies.

```r
data('cham_nominal_votes')

head(cham_nominal_votes)
```
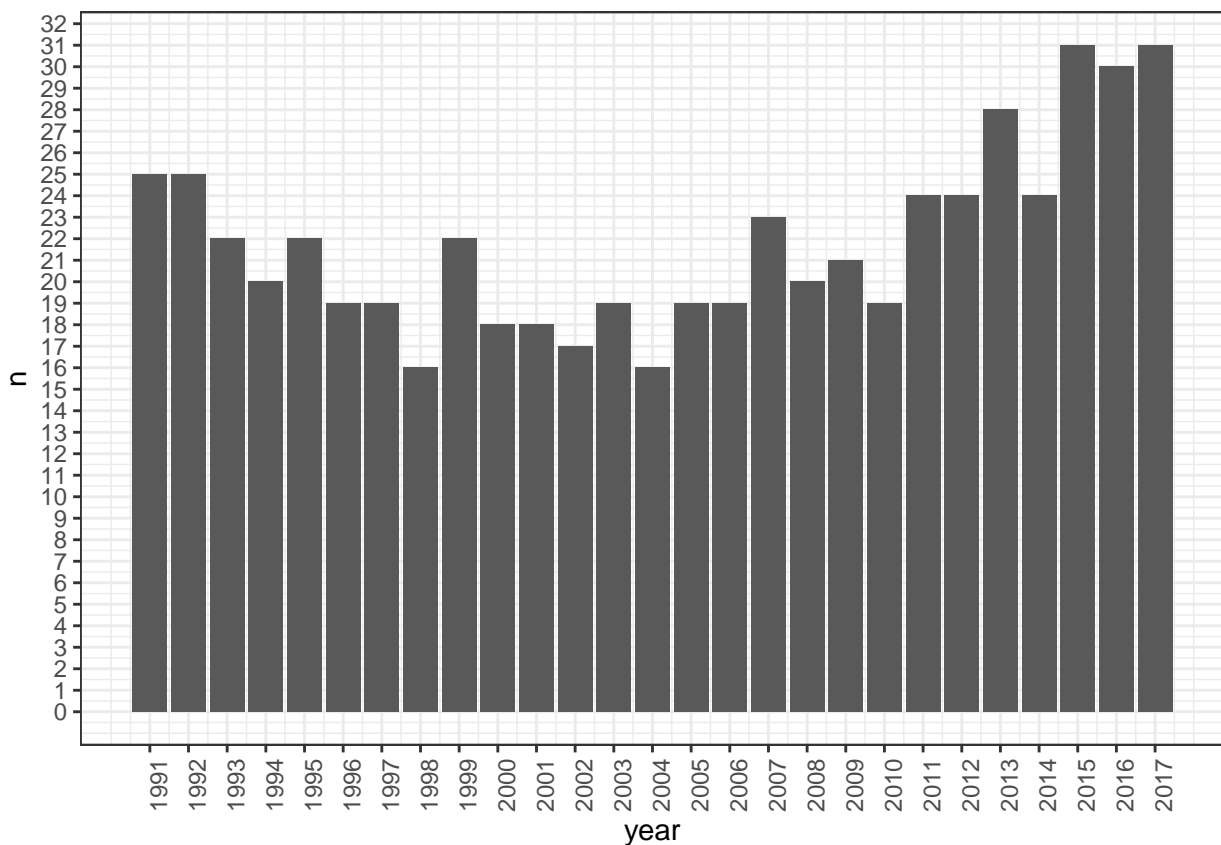
```
## # A tibble: 6 x 9

##   vote_date           rollcall_id legislator_id legislator_name

##   <dttm>              <chr>       <chr>         <chr>

## 1 1991-12-17 00:00:00 PL-4580-19~ 73979         Alceste Almeida

## 2 1991-12-17 00:00:00 PL-4580-19~ 73981         Francisco Rodr~

## 3 1991-12-17 00:00:00 PL-4580-19~ 133881        Joao Fagundes

## 4 1991-12-17 00:00:00 PL-4580-19~ 133880        Marcelo Luz

## 5 1991-12-17 00:00:00 PL-4580-19~ 160608        Teresa Juca

## 6 1991-12-17 00:00:00 PL-4580-19~ 73807         Ruben Bento

## # ... with 5 more variables: legislator_vote <dbl>,

## #   legislator_party <chr>, legislator_state <chr>, GOV_orientation <dbl>,
```

```
## #   bill <chr>
```

Now we clean these data and create a graph with the number of parties by year. Note that we use the `dplyr` and `ggplot2` packages as well.

```
cham_nominal_votes %>%

  mutate(year = year(vote_date)) %>%

  distinct(year, legislator_party) %>%

  count(year) %>%

  ggplot(aes(x = year, y = n)) +

  geom_col() +

  theme_bw() +

  scale_x_continuous(breaks=1991:2017) +

  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +

  scale_y_continuous(breaks = 0:33)
```

4   Rice Index

It is also straightforward to calculate common metrics for parliamentary behavior with congressbr. In this example, we show how to estimate the Rice Index for party cohesion.

For convenience, we will use cham_nominal_votes example dataset again.

```
data("cham_nominal_votes")
head(cham_nominal_votes)

## # A tibble: 6 x 9
##   vote_date           rollcall_id legislator_id legislator_name
##   <dttm>              <chr>       <chr>         <chr>
## 1 1991-12-17 00:00:00 PL-4580-19~ 73979         Alceste Almeida
## 2 1991-12-17 00:00:00 PL-4580-19~ 73981         Francisco Rodr~
## 3 1991-12-17 00:00:00 PL-4580-19~ 133881        Joao Fagundes
## 4 1991-12-17 00:00:00 PL-4580-19~ 133880        Marcelo Luz
## 5 1991-12-17 00:00:00 PL-4580-19~ 160608        Teresa Juca
## 6 1991-12-17 00:00:00 PL-4580-19~ 73807         Ruben Bento
## # ... with 5 more variables: legislator_vote <dbl>,
## #   legislator_party <chr>, legislator_state <chr>, GOV_orientation <dbl>,
## #   bill <chr>
```

We then create a quick routine to calculate the Rice Index. To do so, we have to program a function that returns a vector with the votes (coded '1' for aye, and '0' for nay) on a certain roll call, then calculates the absolute value of two times the number of 'aye' votes minus the length of the vector divided by the length of the vector.

```
rice <- function(votes){

  votes <- votes[!is.na(votes)]

  denominator <- length(votes)

  numerator <- abs(2*sum(votes) - denominator)

  numerator/denominator

}
```

Next, we apply our custom function to calculate results for each roll in a certain year and for each party, once again making use of some functions from the `dplyr` package. We filter out missing votes (NA) and get the year of each rollcall before estimating the index.
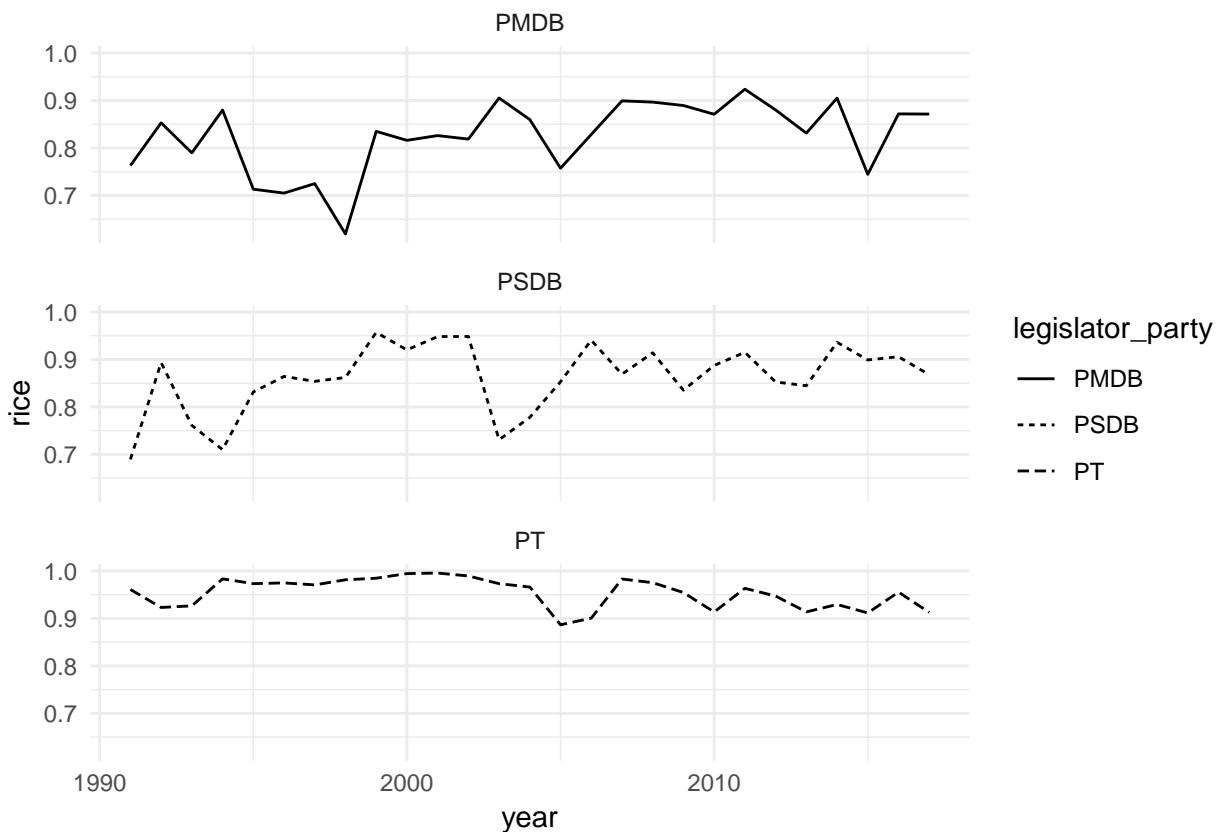
```
cham_rice_index <- cham_nominal_votes %>%

  filter(!is.na(legislator_vote)) %>%

  mutate(year = year(vote_date)) %>%

  group_by(rollcall_id, year, legislator_party) %>%

  summarise(rice = rice(legislator_vote))
```

The code below estimates the average Rice Index by party and year

```
cham_rice_index <- cham_rice_index %>%

  group_by(year, legislator_party) %>%

  summarise(rice = mean(rice))
```

Having done this, we can plot the results for three selected parties: PT, PMDB and PSDB.

```
ggplot(cham_rice_index %>%

        filter(legislator_party %in% c("PT", "PMDB", "PSDB")),

    aes(x = year, y = rice, linetype = legislator_party)) +

    geom_line() +

    facet_wrap(~legislator_party, nrow = 3) + theme_minimal()
```

5 Ideal Point Analysis

In our last example, we will perform an ideal point analysis. The MCMCpack package is required for this type of model, so we load it along with congressbr and dplyr:

```r
library(MCMCpack)

library(congressbr)

library(tidyverse)
```

Here we will use the pre-processed data of senate votes included in congressbr.

```r
data("sen_nominal_votes")
```

Some simple data wrangling is necessary to get the data ready for analysis.

```r
sen <- sen_nominal_votes %>%

  filter(vote_date >= "2013-02-01", !is.na(senator_id)) %>%

  mutate(legislature = ifelse(vote_date >= "2016-05-15", "Temer", "Dilma"),

    sen_id = paste(senator_name, legislature, sep = ":"),
```

```
      senator_vote = as.numeric(senator_vote))
```

Now we will put the same data inside a "roll call matrix" object before running the routine, using the `vote_to_rollcall()` function available in `congressbr`.

```
votes <- vote_to_rollcall(votes = sen$senator_vote,
                          legislators = sen$sen_id,
                          bills = sen$bill,
                          ideal = FALSE)
```

Next, we will employ the `MCMCirt1d()` function from `MCMCpack` and draw a sufficient sample.

```
mcmc_votes <- MCMCirt1d(votes, burnin = 2500, mcmc = 50000, thin = 10,
                        verbose = 0, seed = 1234, drop.constant.items = T,
                        theta.constraints =
                            list(`Jose Agripino:Dilma` = "+",
                                 `Vanessa Grazziotin:Dilma` = "-"))
```

We then extract values for the mean and standard deviation for the estimated parameters (θ):

```
mc <- summary(mcmc_votes)


# extract ideal points ("theta"):

thetas <- as_tibble(mc$statistics[, 1:2]); colnames(thetas) <- c("mean", "SD")

id <- unlist(dimnames(votes)[1])
```

Before plotting the results, we transform those values into a dataframe.

```
legis <- tibble(mean = thetas$mean, id = id) %>%

  separate(id, into = c("name", "legislature"), sep = ":", drop = F) %>%

  filter(name %in% c("Renan Calheiros", "Aecio Neves", "Lindbergh Farias",

                     "Vanessa Grazziotin", "Tasso Jereissati", "Romario",

                     "Magno Malta"))
```

Finally, a few ggplot2 commands and we can generate the plot (this code is somewhat lengthy due to the unusual nature of the plot).

```
legis %>%

  {ggplot(.) +

      geom_line(aes(x = legislature, y = mean, group = name), color = "black", size = 1.5) +

    geom_point(aes(x = legislature, y = mean, group = name), color = "black", size = 5,

              shape = 21, , fill = "white") +

    theme_minimal(base_size = 11) + ylab("Ideal Point") +

    scale_color_brewer(palette = "Dark2") + xlab("President") +

    geom_text(data = subset(., legislature == "Dilma" & name == "Renan Calheiros"),

              aes(x = legislature, y = mean, group = name, label = name),

              size = 4, hjust = 1.1) +

    geom_text(data = subset(., legislature == "Dilma" & name == "Aecio Neves"),

              aes(x = legislature, y = mean, group = name, label = name),

              size = 4, hjust = 1.1) +

    geom_text(data = subset(., legislature == "Dilma" & name == "Lindbergh Farias"),

              aes(x = legislature, y = mean, group = name, label = name),

              size = 4, hjust = 1.1) +

    geom_text(data = subset(., legislature == "Dilma" & name == "Vanessa Grazziotin"),

              aes(x = legislature, y = mean, group = name, label = name),

              size = 4, hjust = 1.1) +

    geom_text(data = subset(., legislature == "Dilma" & name == "Tasso Jereissati"),

              aes(x = legislature, y = mean, group = name, label = name),

              size = 4, hjust = 1.1) +

    geom_text(data = subset(., legislature == "Dilma" & name == "Romario"),

              aes(x = legislature, y = mean, group = name, label = name),

              size = 4, hjust = 1.1) +

    geom_text(data = subset(., legislature == "Dilma" & name == "Magno Malta"),

              aes(x = legislature, y = mean, group = name, label = name),

              size = 4, hjust = 1.1, vjust = -.35) +

    theme(legend.position = "none",

          panel.grid.minor.y = element_blank(),
```

```
        panel.grid.major.x = element_blank(),

        plot.title = element_text(size = 16)

    ) +

    labs(title = "Changes in ideal points, Dilma & Temer Presidencies",

         subtitle = "Selected senators") +

    annotate("text", label = sprintf("\u2190 Government"), y = -0.2, x = "Temer",

             angle = 90) +

    annotate("text", label = sprintf("Opposition \u2192"), y = 1.15, x = "Temer",

             angle = 90)

}
```


Changes in ideal points, Dilma & Temer Presidencies
Selected senators