Replication of Figures, Tables and Numbers. Forecasting Elections in Multi-Party Systems: A Bayesian Approach Combining Polls and Fundamentals

Marcel Neunhoeffer, Lukas F. Stoetzer, Thomas Gschwend,
Simon Munzert & Sebastian Sternberg

6 June 2018

Introduction

This file replicates all the figures, tables and numbers in "Forecasting Elections in Multi-Party Systems: A Bayesian Approach Combining Polls and Fundamentals" (Stoetzer et al. 2018). As re-running all the models can be time consuming, this file takes the MCMC draws in the dataverse and reproduces the Figures and Numbers in the main text as well as all the Figures, Tables and Numbers in the Supplementary Materials. If you run the setup chunk all the necessary packages to replicate the results will be installed and loaded.

If you want to re-run our models you find the R files in `code/R`. You first want to re-run the pre-training of the structural model by running `code/R/01_xx_structural_pre_train_stan.R`, adjust xx accordingly for the country (ger or nz) you want to run the model for. Then you can run the combined model by running `code/R/02_xx_structural_pre_train_stan.R`. You will then have to set the working directory to `code/R`.

The Stan code for the models can be found in `code/model_code`. All the data we used to estimate our models is stored in `data/ger` for Germany and `data/nz` for New Zealand.

R Environment

This code was last tested and run on 6 June 2018.

platform	x86_64-apple-darwin15.6.0
arch	x86_64
os	darwin15.6.0
system	x86_64, darwin15.6.0
status	
major	3
minor	4.3
year	2017
month	11
day	30
svn rev	73796
language	R
version.string	R version 3.4.3 (2017-11-30)
nickname	Kite-Eating Tree
RStudio	1.1.383

Dataverse Structure

Dataverse

- code

- model_code - Contains the Stan model files
- R - Contains the R code necessary for Replication
- data
  - ger - Contains all input data for Germany (Polls and Structural Predictors)
    * Polls
    * Structural
    * structural_inits
  - nz - Contains all input data for New Zealand (Polls and Structural Predictors)
    * Polls
    * Structural
    * structural_inits
- output
  - ger - Contains the output of the models for Germany
    * draws -forecasts
    * plots
  - nz - Contains the output of the models for New Zealand
    * draws
    * forecasts
    * plots
- Replication.Rmd - This file, replicates everything in the article

Questions

If you have any further questions or encounter issues while replicating the results please let us know via email to either (or both) Marcel Neunhoeffer (marcel.neunhoeffer@gess.uni-mannheim.de) or Lukas F. Stoetzer (lukas.stoetzer@uzh.ch).

Replication - Article

The following code chunks will replicate the figures, numbers and tables from the main text and the Supplementary Materials.

Final Forecast in Text (p. 5)

```r
# Final Forecast in Text

df <- readRDS(file="output/ger/draws/combined_model/draws_forcast_levels_2017_2.RDS")

forecast <- df$forecast

colnames(forecast) <- df$party_names

adjustOrder <- match(c("cdu", "spd", "lin", "gru", "fdp", "afd", "oth"), df$party_names)


forecast <- forecast[, adjustOrder]

# Mean Forecast
round(apply(forecast, 2, mean)*100, 1)
```

```
##  cdu  spd  lin  gru  fdp  afd  oth
```

```
## 36.2 22.1  9.2  7.8  9.2  9.8  5.7
```
```
# 5/6 Credible Intervals
round(t(apply(forecast, 2, quantile, c(1/12, 11/12)))*100, 1)
```
```
##
## parameters 8.333333% 91.66667%
##        cdu      30.7      41.5
##        spd      19.8      24.2
##        lin       7.2      11.4
##        gru       5.9       9.6
##        fdp       7.2      11.4
##        afd       7.4      12.5
##        oth       3.8       7.8
```
```
# RMSE in Text

# 2017 Election results in the same order

election_res <- c(32.9, 20.5, 9.2, 8.9, 10.7, 12.6, 5)


round(sqrt(mean((apply(forecast, 2, mean)*100 - election_res)^2)), 2)
```
```
## [1] 1.88
```

Probabilities in Text

```
# Probabilities in text

# Probability of 7 (counting CDU and CSU as two parties) parties entering parliament
sum(apply(forecast[,1:6] >= 0.05, 1, all)) / nrow(forecast) * 100
```
```
## [1] 95.2
```
```
# Probability for AfD to become third strongest party
sum(t(apply(-forecast[,1:6], 1, rank))[,6] == 3) / nrow(forecast) * 100
```
```
## [1] 43
```
```
# Probability for Left to become third strongest party
sum(t(apply(-forecast[,1:6], 1, rank))[,3] == 3) / nrow(forecast) * 100
```
```
## [1] 25.73333
```
```
# Probability for Greens to become third strongest party
sum(t(apply(-forecast[,1:6], 1, rank))[,4] == 3) / nrow(forecast) * 100
```
```
## [1] 6
```
```
# Probability for FDP to become third strongest party
sum(t(apply(-forecast[,1:6], 1, rank))[,5] == 3) / nrow(forecast) * 100
```
```
## [1] 24.7
```
```
# Coalition Probabilities

coa_df <- forecast
```

```r
# Set share of parties below 5-percent-threshold to 0

coa_df[coa_df[,1:6] < 0.05] <- 0


# Majority for Grand Coalition?
# (When coalition vote share is greater than half of the sum of the vote shares
# of parties anbove the 5% thershold.)

gc <- (coa_df[,1] + coa_df[,2]) > (apply(coa_df[,1:6], 1, sum) / 2)

sum(gc) / nrow(forecast) * 100
```

```
## [1] 99.66667
```

```r
# Majority for Jamaica Coalition?

jam <- (coa_df[,1] + coa_df[,4] + coa_df[,5]) > (apply(coa_df[,1:6], 1, sum) / 2)

sum(jam) / nrow(forecast) * 100
```

```
## [1] 95.8
```

```r
# Majority for Black-Yellow?

by <- (coa_df[,1] + coa_df[,5]) > (apply(coa_df[,1:6], 1, sum) / 2)

sum(by) / nrow(forecast) * 100
```

```
## [1] 27.4
```

```r
# Majority for Black-Green?

bg <- (coa_df[,1] + coa_df[,4]) > (apply(coa_df[,1:6], 1, sum) / 2)

sum(bg) / nrow(forecast) * 100
```

```
## [1] 15.53333
```

```r
# Majority for Red-Red-Green?

rrg <- (coa_df[,2] + coa_df[,3] + coa_df[,4]) > (apply(coa_df[,1:6], 1, sum) / 2)

sum(rrg) / nrow(forecast) * 100
```

```
## [1] 0.8
```

Figure 1
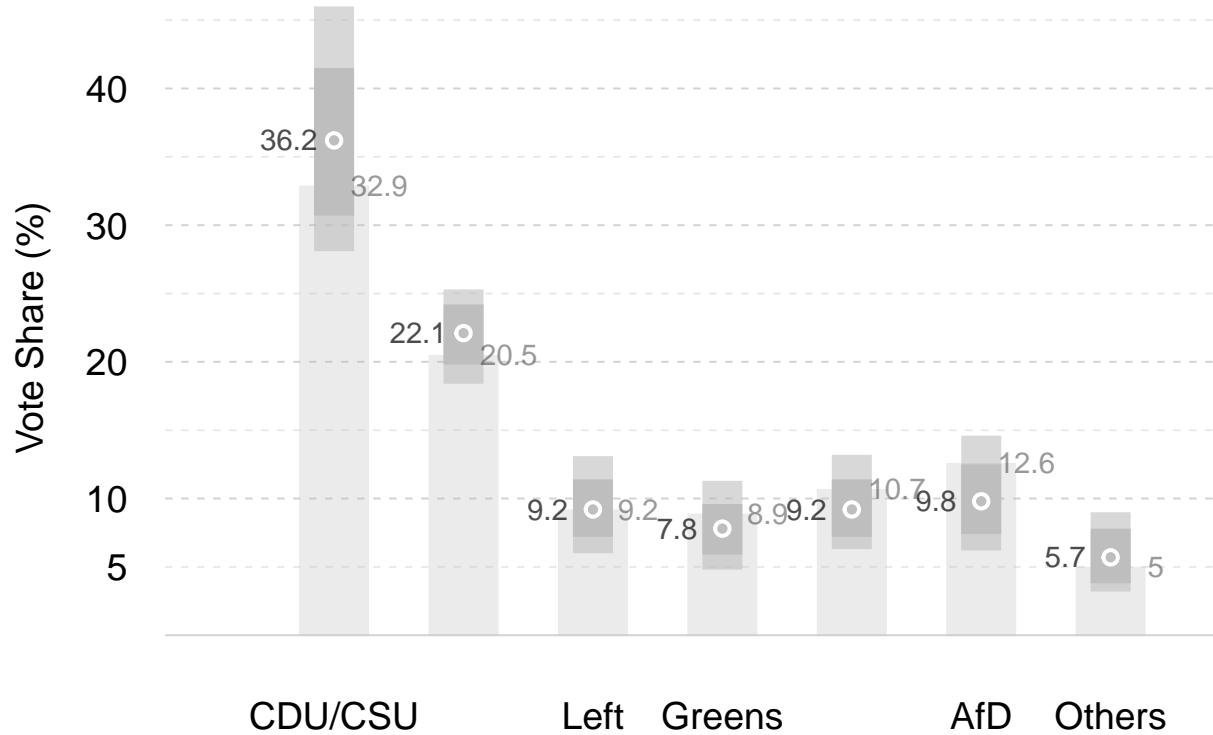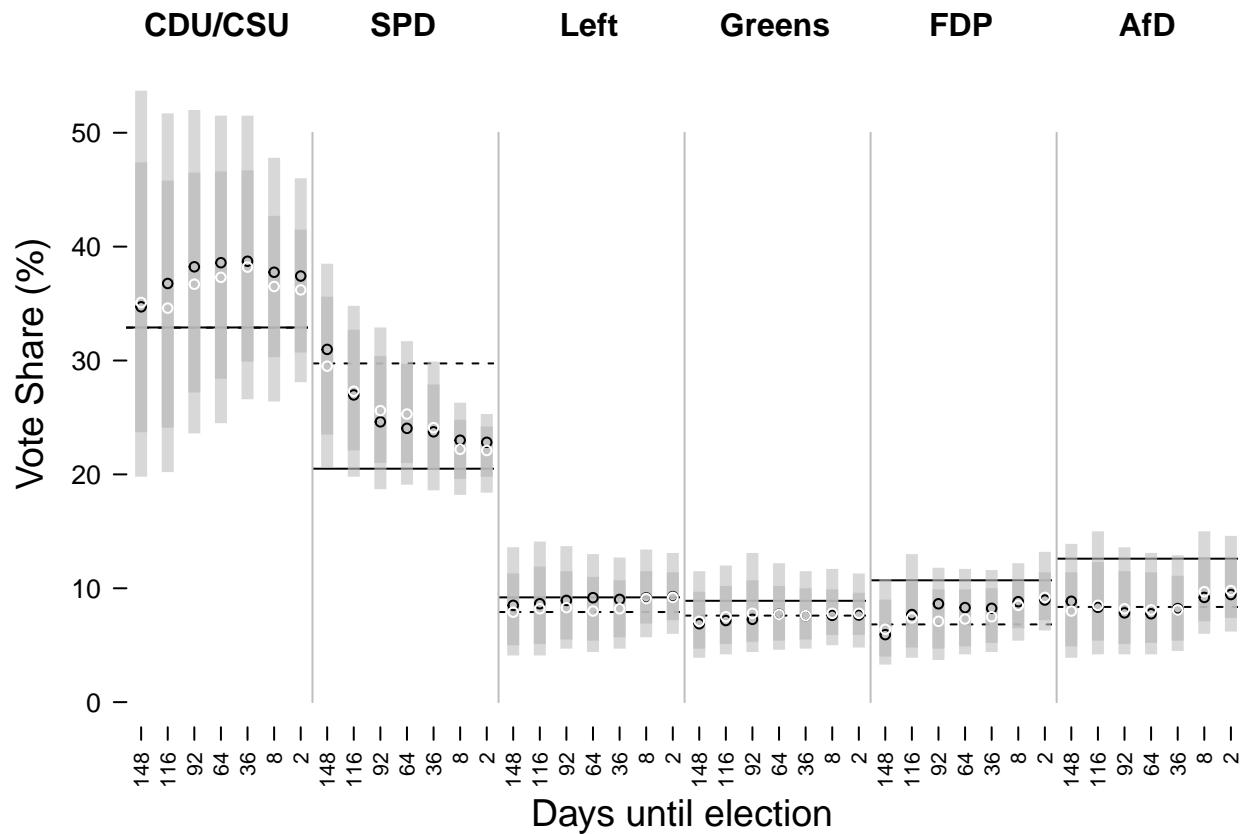
```r
source("code/R/Figure1.R")
```

Figure 2
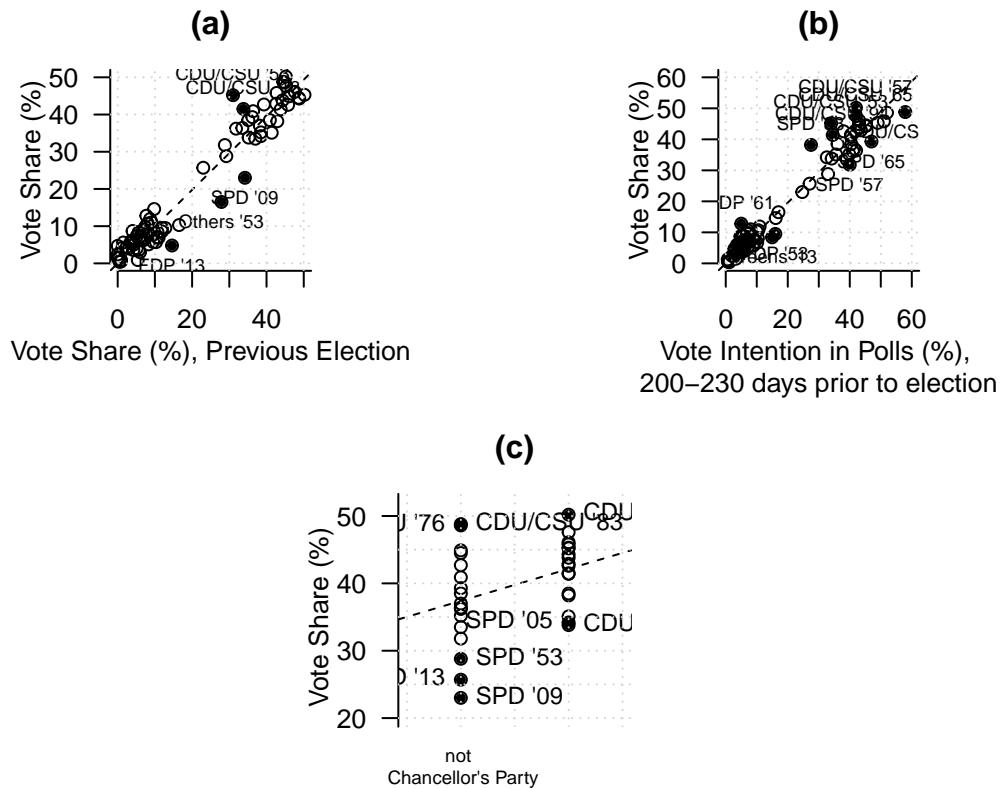
```r
source("code/R/Figure2.R")
```

Replication - Supplementary Material - Germany

B.1.2 Figure 1

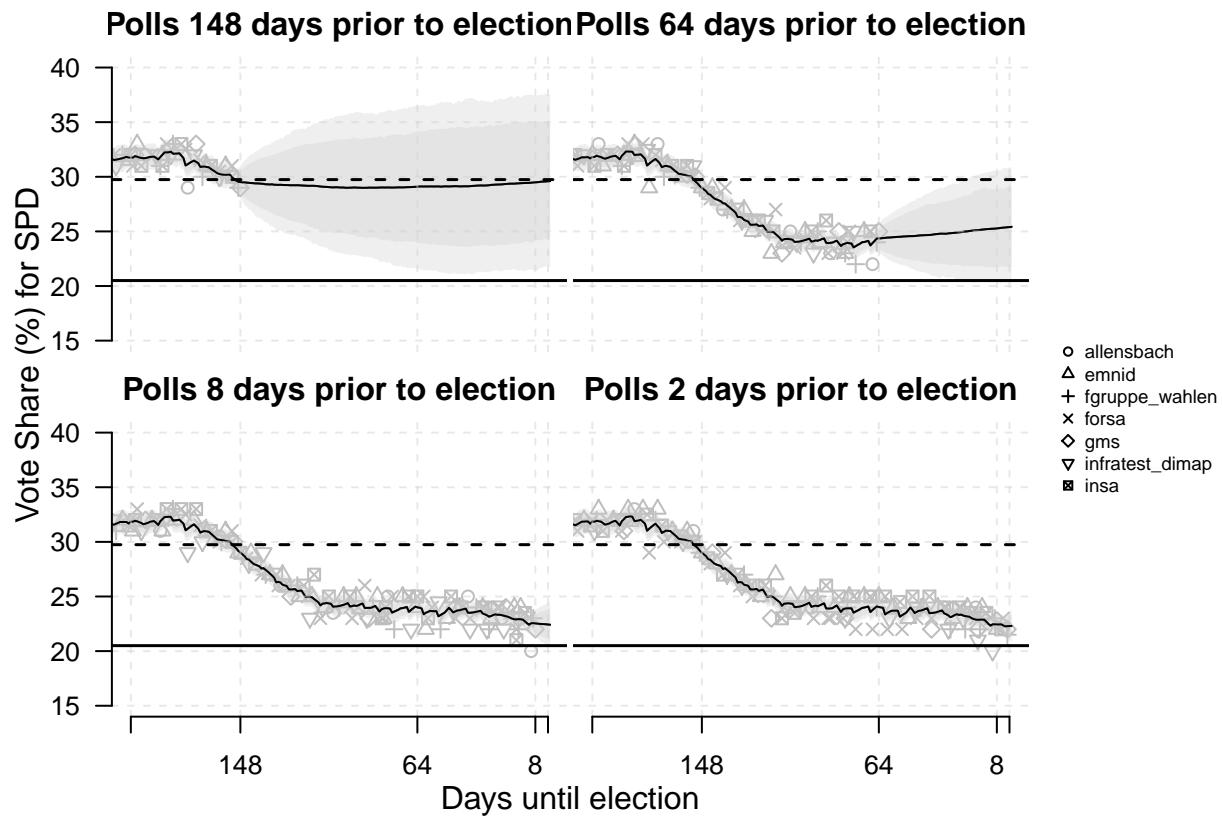```
source("code/R/Figure1B12.R")
```

**(a)**

**(b)**

**(c)**

### B.1.3 Figure 2

```r
source("code/R/Figure2B13.R")
```


Vote Share Previous Election

Vote Intention in Polls, 200−230 days prior to election

Chancellor's Party

Regression Coefficient
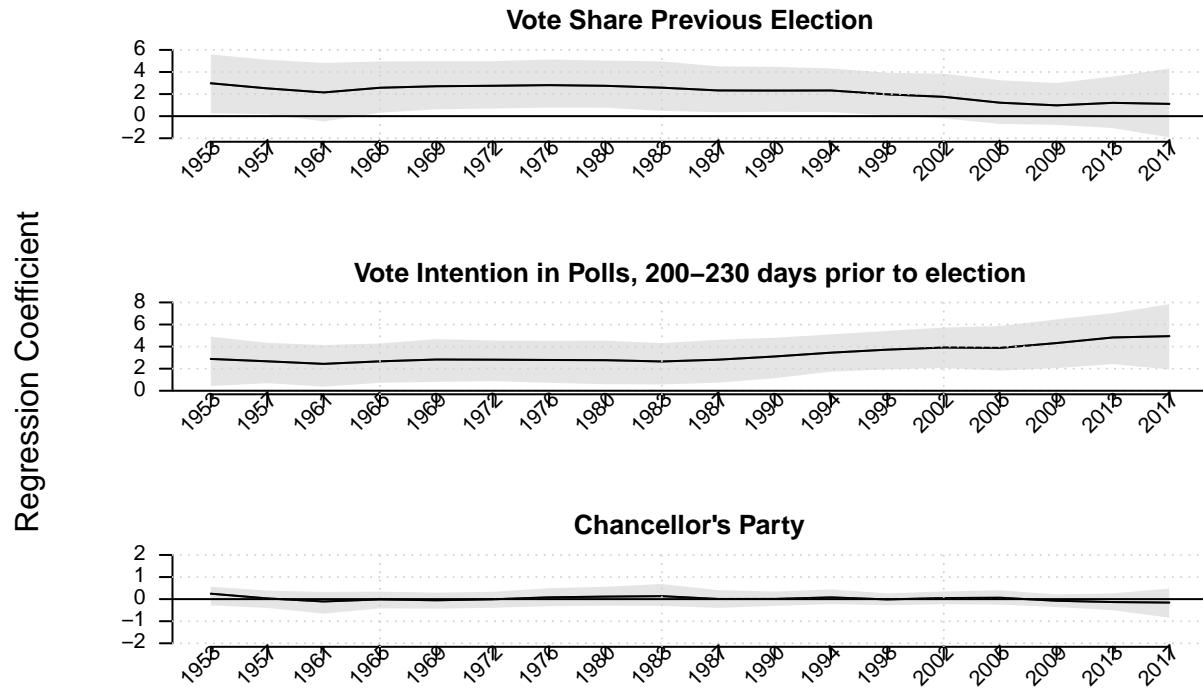
B.2 Figure 3

```
source("code/R/Figure3B2.R")
```


B.3.1 Table 1

```
source("code/R/Table1B31.R")
```

```
## % latex table generated in R 3.4.3 by xtable 1.8-2 package
## % Wed Jun  6 21:19:43 2018
## \begin{table}[ht]
## \centering
## \begin{tabular}{rrrrrrr}
##    \hline
##  & 2002-17 & 2002 & 2005 & 2009 & 2013 & 2017 \\
##    \hline
## Structural & 3.88 & 2.29 & 4.66 & 4.28 & 3.50 & 4.22 \\
##    2 & 1.79 & 1.27 & 2.96 & 0.92 & 1.20 & 1.88 \\
##    8 & 1.95 & 0.92 & 3.12 & 1.12 & 1.67 & 2.13 \\
##    36 & 2.88 & 2.03 & 4.34 & 1.52 & 2.33 & 3.29 \\
##    64 & 3.07 & 1.84 & 5.02 & 1.37 & 2.38 & 3.36 \\
##    92 & 3.33 & 1.91 & 5.67 & 1.49 & 2.69 & 3.31 \\
##    116 & 3.31 & 2.58 & 5.29 & 1.74 & 2.46 & 3.42 \\
##    148 & 3.45 & 2.15 & 5.14 & 1.93 & 2.42 & 4.34 \\
##    \hline
## \end{tabular}
## \end{tabular}
```

```
## \end{table}
```

B.3.2 Table 2

```
source("code/R/Table2B32.R")
```

```
## % latex table generated in R 3.4.3 by xtable 1.8-2 package
## % Wed Jun  6 21:19:51 2018
## \begin{table}[ht]
## \centering
## \begin{tabular}{rrrr}
##   \hline
##  & Full Fundamentals Model & Only last vote share Fundamentals Model & No House \\
##   \hline
## Structural & 4.22 & 5.59 & 4.22 \\
##   2 & 1.88 & 2.02 & 2.01 \\
##   8 & 2.13 & 2.40 & 2.06 \\
##   36 & 3.29 & 4.32 & 3.12 \\
##   64 & 3.36 & 4.37 & 3.14 \\
##   92 & 3.31 & 4.59 & 3.39 \\
##   116 & 3.42 & 4.54 & 3.47 \\
##   148 & 4.34 & 5.19 & 3.92 \\
##    \hline
## \end{tabular}
## \end{table}
```

B.3.3 Table 3

```
source("code/R/Table3B33.R")
```

```
## % latex table generated in R 3.4.3 by xtable 1.8-2 package
## % Wed Jun  6 21:20:03 2018
## \begin{table}[ht]
## \centering
## \begin{tabular}{rrrrrrr}
##   \hline
##  & 2002-17 & 2002 & 2005 & 2009 & 2013 & 2017 \\
##   \hline
## Combined Model 2 & 1.79 & 1.27 & 2.96 & 0.92 & 1.20 & 1.88 \\
##   Avg. of Polls 2 & 2.07 & 1.23 & 3.30 & 1.17 & 1.42 & 2.41 \\
##   Combined Model 8 & 1.95 & 0.92 & 3.12 & 1.12 & 1.67 & 2.13 \\
##   Avg. of Polls 8 & 2.23 & 1.34 & 3.44 & 1.27 & 1.69 & 2.60 \\
##   Combined Model 36 & 2.88 & 2.03 & 4.34 & 1.52 & 2.33 & 3.29 \\
##   Avg. of Polls 36 & 2.89 & 2.14 & 4.45 & 1.52 & 2.27 & 3.20 \\
##   Combined Model 64 & 3.07 & 1.84 & 5.02 & 1.37 & 2.38 & 3.36 \\
##   Avg. of Polls 64 & 3.10 & 1.87 & 5.16 & 1.39 & 2.36 & 3.31 \\
##   Combined Model 92 & 3.33 & 1.91 & 5.67 & 1.49 & 2.69 & 3.31 \\
##   Avg. of Polls 92 & 3.27 & 2.42 & 5.48 & 1.46 & 2.33 & 3.29 \\
##   Combined Model 116 & 3.31 & 2.58 & 5.29 & 1.74 & 2.46 & 3.42 \\
##   Avg. of Polls 116 & 3.32 & 2.96 & 4.95 & 1.83 & 2.52 & 3.58 \\
##   Combined Model 148 & 3.45 & 2.15 & 5.14 & 1.93 & 2.42 & 4.34 \\
##   Avg. of Polls 148 & 3.53 & 2.06 & 4.91 & 1.65 & 2.79 & 4.74 \\
```

```
##     \hline
## \end{tabular}
## \end{table}
```

B.3.4 Table 4

```
source("code/R/Table4B34.R")
```

```
## Warning in data.row.names(row.names, rowsi, i): some row.names duplicated:
## 3,4,5,6,7,8,9,10,11,12 --> row.names NOT used

## % latex table generated in R 3.4.3 by xtable 1.8-2 package
## % Wed Jun  6 21:20:09 2018
## \begin{table}[ht]
## \centering
## \begin{tabular}{rllllll}
##    \hline
##  & afd & cdu & fdp & gru & lin & oth \\
##    \hline
## 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
##    2 & (1; 1) & (0; 0) & (0; 0) & (0; 0) & (0; 0) & (0; 0) \\
##    3 & -0.21 & 0.97 & 0 & 0 & 0 & 0 \\
##    4 & (-0.39; -0.02) & (0.92; 1) & (0; 0) & (0; 0) & (0; 0) & (0; 0) \\
##    5 & 0.14 & 0.06 & 0.98 & 0 & 0 & 0 \\
##    6 & (-0.04; 0.32) & (-0.1; 0.24) & (0.94; 1) & (0; 0) & (0; 0) & (0; 0) \\
##    7 & -0.04 & 0.28 & 0.06 & 0.94 & 0 & 0 \\
##    8 & (-0.23; 0.15) & (0.11; 0.44) & (-0.12; 0.23) & (0.88; 0.99) & (0; 0) & (0; 0) \\
##    9 & 0.29 & 0.01 & 0.13 & 0.04 & 0.93 & 0 \\
##    10 & (0.11; 0.47) & (-0.18; 0.19) & (-0.05; 0.3) & (-0.14; 0.23) & (0.86; 0.98) & (0; 0) \\
##    11 & -0.29 & 0.25 & -0.01 & -0.02 & -0.04 & 0.9 \\
##    12 & (-0.45; -0.12) & (0.08; 0.42) & (-0.18; 0.16) & (-0.2; 0.15) & (-0.23; 0.14) & (0.82; 0.96) \'
##     \hline
## \end{tabular}
## \end{table}
```

Replication - Supplementary Material - New Zealand

Forecast and RMSE in Text (C.1)

```
df <- readRDS(file="output/nz/draws/combined_model/draws_forcast_levels_2017_2.RDS")

forecast <- df$forecast

adjustOrder <- match(c("nat", "lab", "gre", "nzf", "oth"), df$party_names)


forecast <- forecast[,adjustOrder]


# Party Names

plot_names <- c("National", "Labour", "Green", "NZ First", "Others")
```

```r
df_forecast <- data.frame(y = apply(forecast, 2,  mean),
                          ci = t(apply(forecast, 2, function(x) quantile(x, c(1/12, 11/12)))),
                          ci95 = t(apply(forecast,2, function(x) quantile(x, c(0.025, 0.975)))))

rownames(df_forecast) <- plot_names
colnames(df_forecast) <- c("value", "low", "high", "low95", "high95")

df_forecast <- round(df_forecast*100, 1)

# Numbers in text

df_forecast
```

```
##           value  low high low95 high95
## National  45.9 38.9 53.0  34.6   57.4
## Labour    35.3 28.6 42.1  24.5   46.5
## Green      8.3  6.2 10.6   5.2   12.6
## NZ First   6.3  4.7  8.1   3.9    9.3
## Others     4.2  3.5  4.9   3.2    5.4
```

```r
round(sqrt(mean((round(df_forecast,1)[,1] -  c(44.45, 36.89, 6.27, 7.2, 5.19))^2)),2)
```

```
## [1] 1.45
```

```r
# Quantities of Interest in text

# Probability of National becoming strongest party
round(sum(t(apply(-forecast[,1:5], 1, rank))[,1] == 1) / nrow(forecast) * 100, 1)
```

```
## [1] 86.9
```

```r
# Probability for Greens to clear the 5 percent threshold
round(sum(forecast[,3] >= 0.05)  / nrow(forecast) * 100, 1)
```
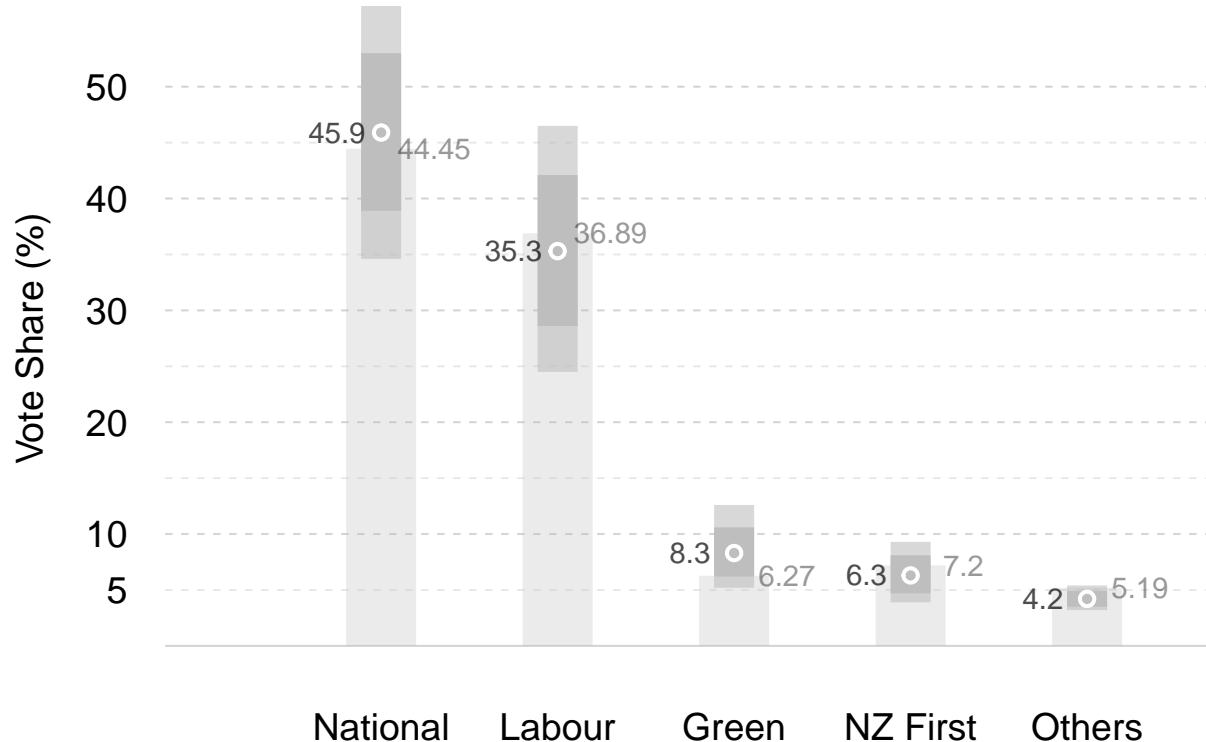
```
## [1] 97.9
```

```r
# Probability for NZ First to clear the 5 percent threshold
round(sum(forecast[,4] >= 0.05)  / nrow(forecast) * 100, 1)
```
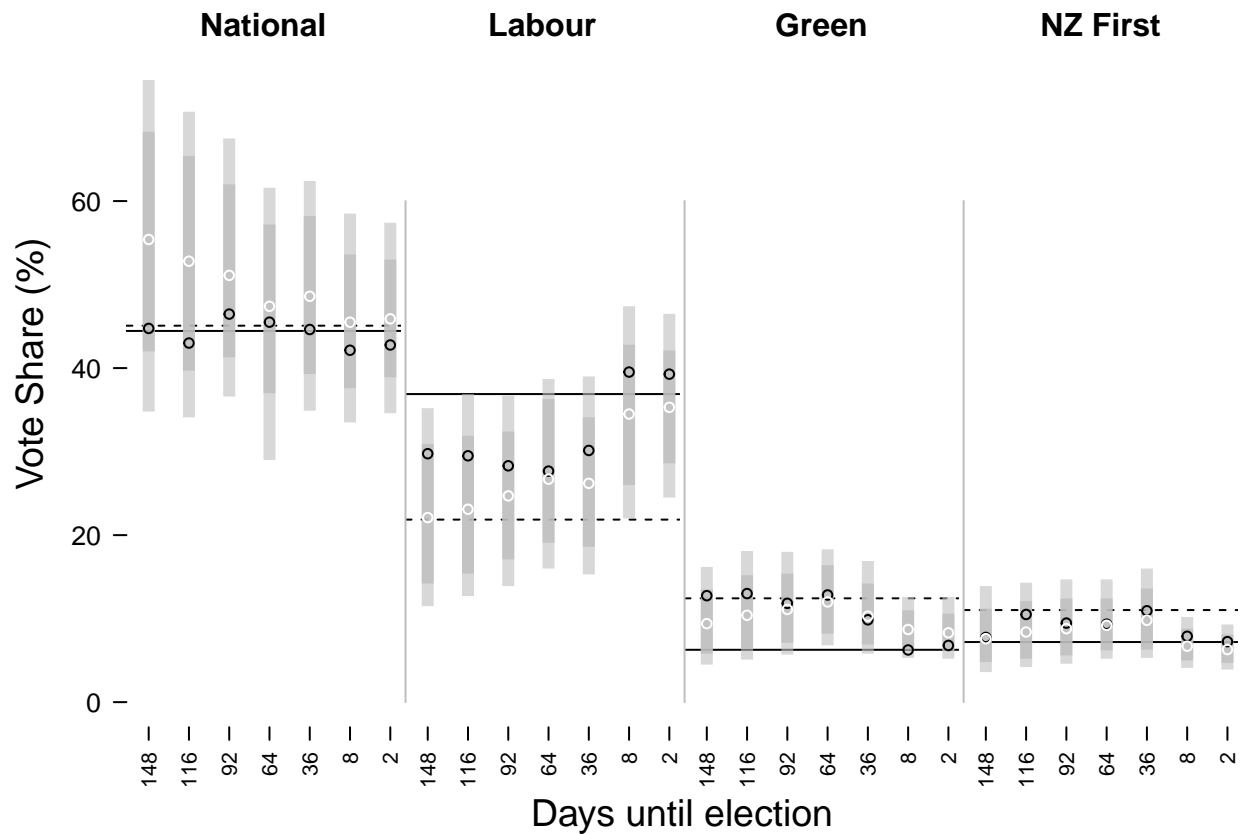
```
## [1] 86.1
```

C.1 Figure 4
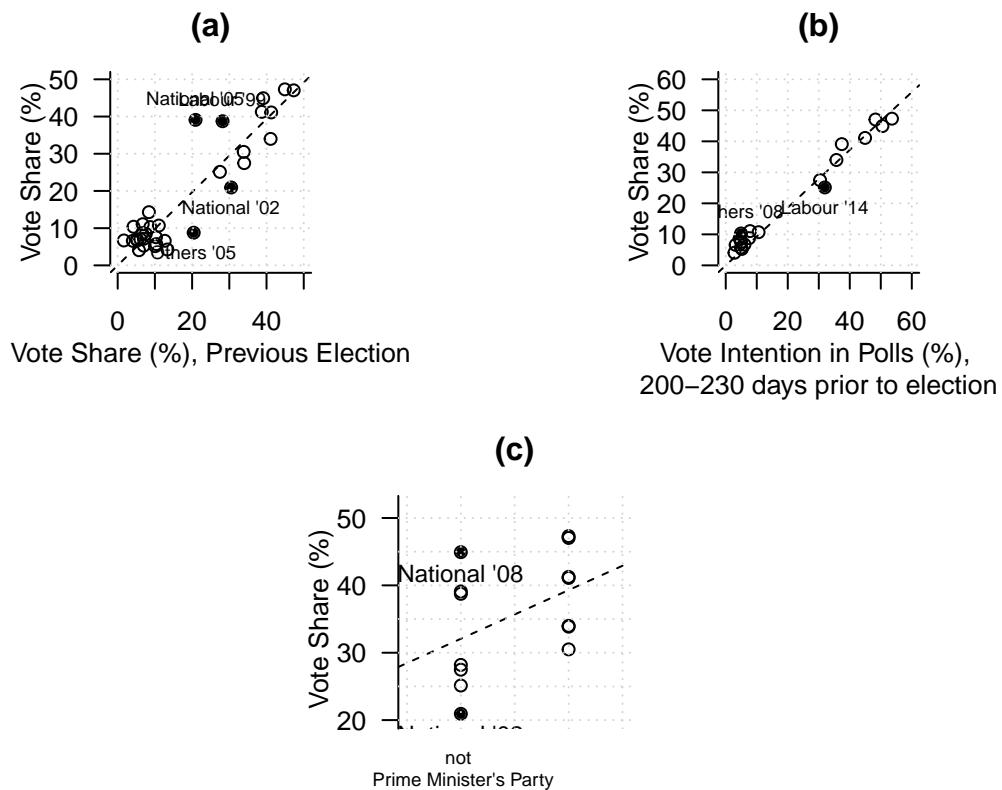
```r
source("code/R/Figure4C1.R")
```

C.1 Figure 5

```r
source("code/R/Figure5C1.R")
```
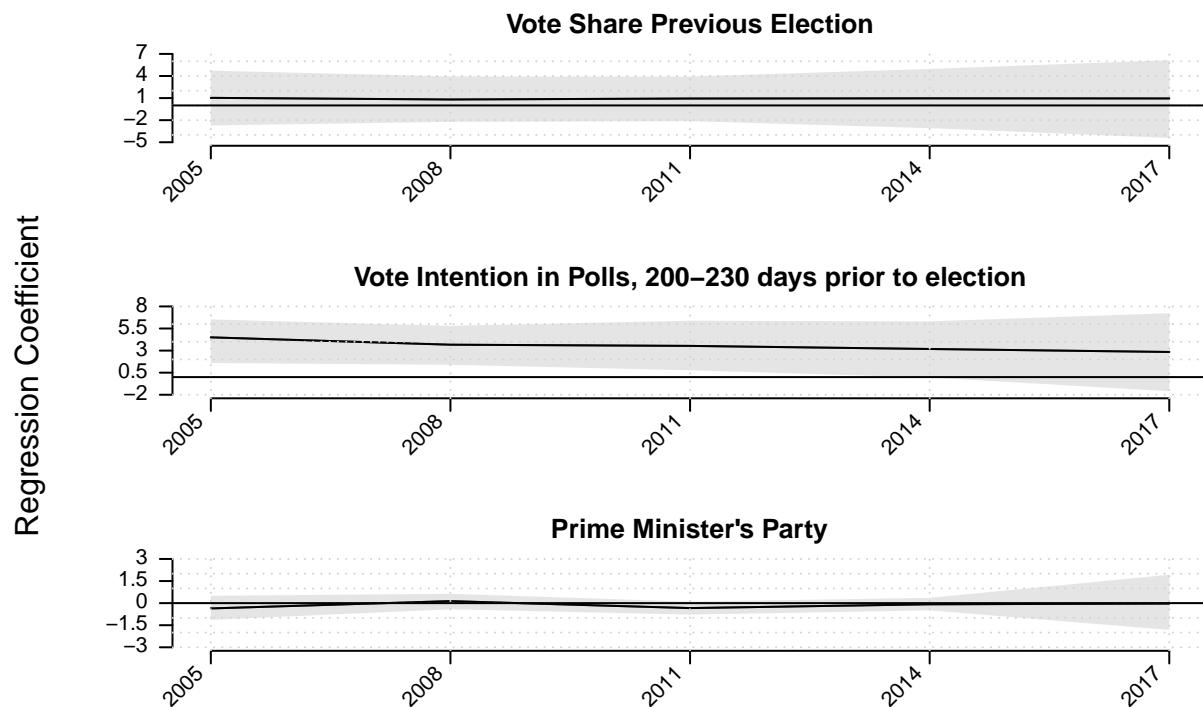
Figure with four panels (National, Labour, Green, NZ First) showing Vote Share (%) on the y-axis and Days until election on the x-axis (148, 116, 92, 64, 36, 8, 2).

C.2.2 Figure 6

```r
source("code/R/Figure6C22.R")
```

(a)

Vote Share (%)
National '05 / Labour '05 / '02
National '02
Others '05
Vote Share (%), Previous Election

(b)

Vote Share (%)
Others '08 / Labour '14
Vote Intention in Polls (%),
200–230 days prior to election

(c)

Vote Share (%)
National '08
not
Prime Minister's Party

C.2.3 Figure 7

```
source("code/R/Figure7C23.R")
```


Vote Share Previous Election

Vote Intention in Polls, 200–230 days prior to election

Prime Minister's Party

Regression Coefficient

C.3 Figure 8

```
source("code/R/Figure8C3.R")
```
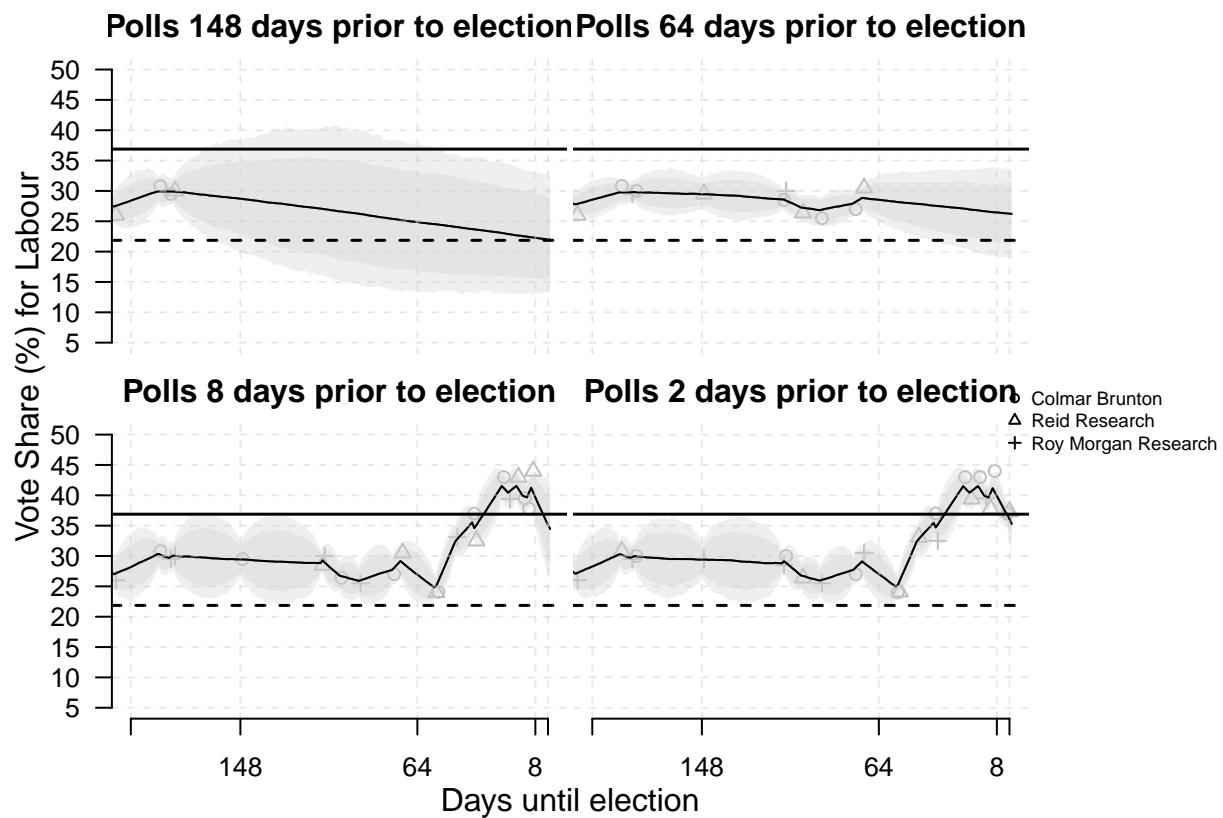

C.4.1 Table 5

```
source("code/R/Table5C41.R")
```

```
## % latex table generated in R 3.4.3 by xtable 1.8-2 package
## % Wed Jun  6 21:20:27 2018
## \begin{table}[ht]
## \centering
## \begin{tabular}{rrrrr}
##   \hline
##  & 2011-17 & 2011 & 2014 & 2017 \\
##   \hline
## Structural & 6.37 & 7.60 & 2.03 & 7.73 \\
##   2 & 1.84 & 2.42 & 1.48 & 1.46 \\
##   8 & 1.99 & 2.82 & 1.15 & 1.63 \\
##   36 & 4.66 & 5.00 & 3.00 & 5.58 \\
##   64 & 4.82 & 5.85 & 2.34 & 5.48 \\
##   92 & 5.37 & 5.83 & 2.99 & 6.60 \\
##   116 & 5.98 & 6.88 & 2.04 & 7.47 \\
##   148 & 6.73 & 7.88 & 1.96 & 8.36 \\
##    \hline
## \end{tabular}
```

```
## \end{table}
```

C.4.2 Table 6

```
source("code/R/Table6C42.R")
```

```
## % latex table generated in R 3.4.3 by xtable 1.8-2 package
## % Wed Jun  6 21:20:32 2018
## \begin{table}[ht]
## \centering
## \begin{tabular}{rrrrr}
##    \hline
##  & 2011-17 & 2011 & 2014 & 2017 \\
##    \hline
## Combined Model 2 & 1.84 & 2.42 & 1.48 & 1.46 \\
##    Avg. of Polls 2 & 2.12 & 2.92 & 1.82 & 1.30 \\
##    Combined Model 8 & 1.99 & 2.82 & 1.15 & 1.63 \\
##    Avg. of Polls 8 & 2.41 & 3.35 & 1.88 & 1.65 \\
##    Combined Model 36 & 4.66 & 5.00 & 3.00 & 5.58 \\
##    Avg. of Polls 36 & 3.68 & 4.15 & 2.95 & 3.83 \\
##    Combined Model 64 & 4.82 & 5.85 & 2.34 & 5.48 \\
##    Avg. of Polls 64 & 4.17 & 3.99 & 3.09 & 5.18 \\
##    Combined Model 92 & 5.37 & 5.83 & 2.99 & 6.60 \\
##    Avg. of Polls 92 & 4.09 & 3.71 & 3.65 & 4.81 \\
##    Combined Model 116 & 5.98 & 6.88 & 2.04 & 7.47 \\
##    Avg. of Polls 116 & 4.07 & 4.07 & 3.22 & 4.78 \\
##    Combined Model 148 & 6.73 & 7.88 & 1.96 & 8.36 \\
##    Avg. of Polls 148 & 3.99 & 4.20 & 3.38 & 4.32 \\
##     \hline
## \end{tabular}
## \end{table}
```