Replicating Adams et al. (2023)

(Can't We All Just Get Along?)

Daniel Benenson Markovits; Nadjim Fréchet; Rémi Thériault; Chris Yurris

2024-02-22

Report & Conclusions

This replication study was preregistered at: https://osf.io/wd5fk. The source code for this report/project is available on GitHub: https://github.com/Nadjim10/replication_game_mtl_2023.

We replicated the main results of Adams et al (2023) for the Montreal Replication Games on June 14, 2023. The main results replicated with the authors' initial code and with a simplification of the code using the `tidyverse` (the initial code was written in base R, see section Tidyverse Code Simplification). Our main substantive findings besides this replication is that the results of the paper were driven by left wing respondents and that the main findings were otherwise robust to various additional tests beyond those carried out by the authors.

We examined the results subsetted by the party family and by left and right blocks. We find highly significant results for left wing parties while right wing and centrist parties exhibit clear null results with some positively and others negatively signed. Substantively we believe this finding matters because of recent literature on democratic backsliding - the threat to democracy today comes mostly from the right Berman (2022) and it is right wing voters who are mostly willing to sacrifice democracy for instrumental reasons (Svolik 2023). It matters then that the portion of women MPs affects the attitudes of left wing voters and not the attitudes of the voters most likely to undermine democracy. We explain this treatment effect heterogeneity because of differences in party level attitudes towards female representation - left wing parties tend to have gender equality in representation written into their platforms. The conditional effect by evaluator party family is below (see section Affective Polarization by Party).

We performed a country and year level version of leave one out cross-validation. The results are robust to the exclusion of any country and any year (1996-2007) in the dataset. The exclusion of minor parties from the dataset slightly reduces the main effect size but does not affect their statistical significance or substantive interpretation.

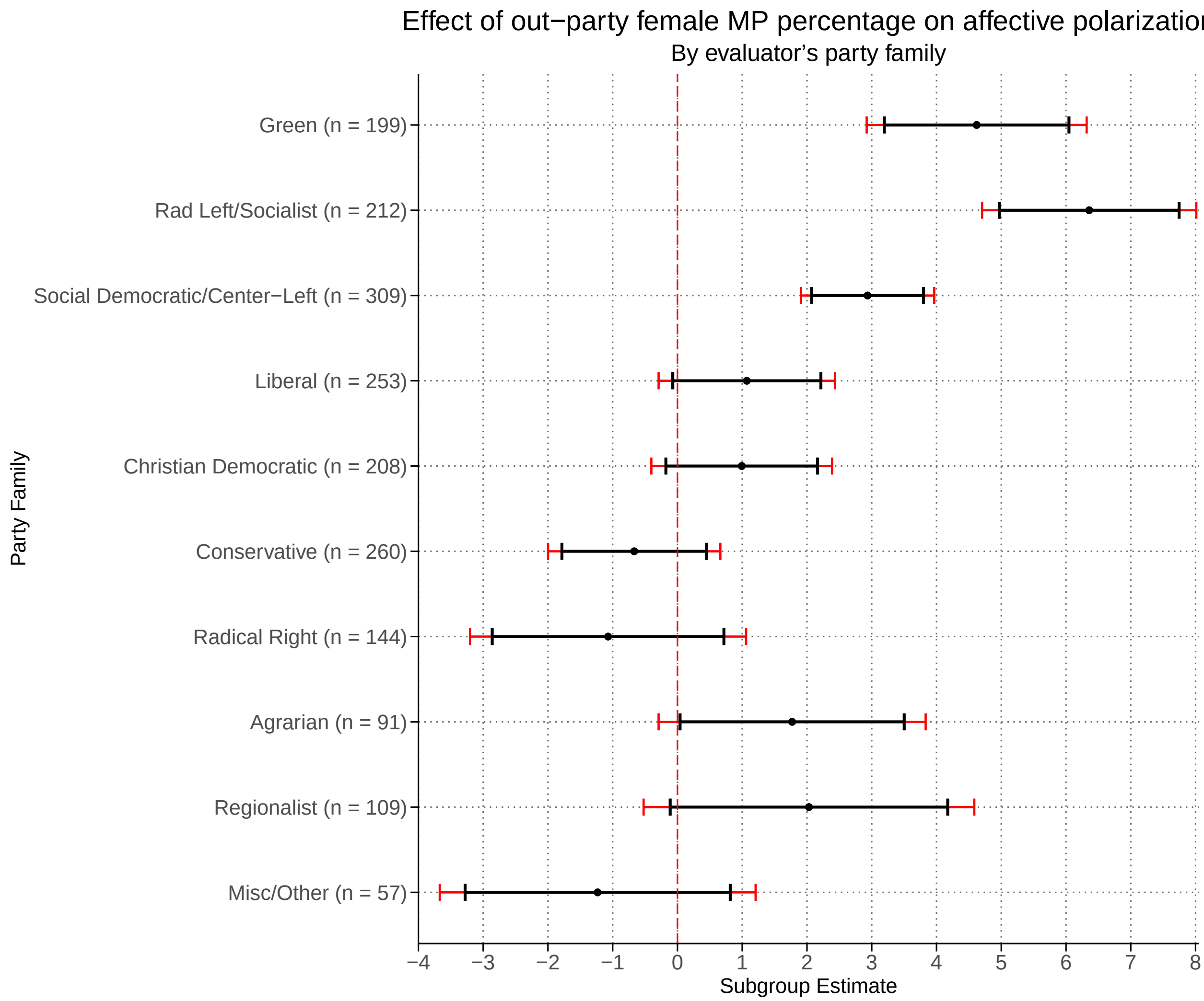
In addition, we wanted to clarify some aspects of the structure of the data that were not entirely clear from the paper, which is understandable given the paper is a short letter. The CSES data is an individual level dataset with more than 300,000 respondents. Weights are not consistently available for these data sets because of the structure of the CSES. Some country-election year iterations of the data have election weights (where respondents are weighted to election results) while others have demographic weights. In the absence of these weights the dataset is transformed into an (N ~= 2000) average of country-year-dyad affective scores - where each data point is an average of the out-party warmth scores in country I at time T from party A's assessment of party B. The main analyses are performed on this data. We note that the absence of weighting may affect the results in a number of ways.

Finally, we also observed that `na.omit()` was used to exclude rows with missing observations even when the missing data were not involved in the analyses. However, listwise deletion is almost always most biased than a good imputation technique (van Ginkel, et al., 2020). Thus, it seems that at least for Table 1, the original authors excluded observations unnecessarily. In the last section, we impute the missing data through random forests before reproducing Table 1. We find that imputation does not affect the table coefficients statistical significance or substantive interpretation (see section Imputing Missing Data).

Affective Polarization by Party

```
party_spec <- function(x) {
  table1.2 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
    prior_opposition + as.factor(cntryyr), data = dta %>%
    subset(from_parfam == x), group = "Party")
  fig <- broom::tidy(table1.2) %>%
    mutate(par_fam = x, group = "Party") %>%
    subset(term == "to_pfeml")
  obs <- nobs(table1.2)
  cbind(fig, obs)
}

load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data

# creating the country-year fixed effects
dta$cntryyr <- paste(dta$country, dta$year, sep = "")

## Removing smaller parties
dta <- subset(dta, dta$to_prior_seats >= 4)

vars <- c(
  "rile_distance_s", "prior_coalition", "prior_opposition", "econ_distance_s",
  "society_distance_s", "year", "country", "party_dislike", "party_like",
  "cntryyr", "to_pfeml", "to_prior_seats", "to_mp_number", "from_parfam",
  "from_left_bloc"
)
dta <- dta[vars]
dta <- na.omit(dta)

parties <- unique(dta$from_parfam)

bind_rows(map(parties, party_spec)) %>%
  mutate(
    par_fam2 = ifelse(par_fam == "10", "Green", par_fam),
    par_fam2 = ifelse(par_fam2 == "20", "Rad Left/Socialist", par_fam2),
    par_fam2 = ifelse(par_fam2 == "30", "Social Democratic/Center-Left",
                      par_fam2),
    par_fam2 = ifelse(par_fam2 == "40", "Liberal", par_fam2),
    par_fam2 = ifelse(par_fam2 == "50", "Christian Democratic", par_fam2),
    par_fam2 = ifelse(par_fam2 == "60", "Conservative", par_fam2),
    par_fam2 = ifelse(par_fam2 == "70", "Radical Right", par_fam2),
    par_fam2 = ifelse(par_fam2 == "80", "Agrarian", par_fam2),
    par_fam2 = ifelse(par_fam2 == "90", "Regionalist", par_fam2),
    par_fam2 = ifelse(par_fam2 == "95", "Misc/Other", par_fam2),
    par_fam2 = paste(par_fam2, " (n = ", obs, ")", sep = ""),
    par_fam2 = fct_reorder(par_fam2, desc(par_fam))
```

```r
) %>%
ggplot(aes(
  x = estimate, y = par_fam2, xmin = estimate - std.error * 1.96,
  xmax = estimate + std.error * 1.96
)) +
geom_point() +
geom_errorbar(size = .75, color = "red", width = .2) +
labs(
  x = "Subgroup Estimate", y = "Party Family",
  title = "Effect of out-party female MP percentage on affective polarization",
  subtitle = "By evaluator's party family"
) +
geom_point(size = 2) +
geom_hline(yintercept = 0, linetype = "longdash", color = "red") +
geom_errorbar(
  aes(
    xmin = estimate - std.error * 1.645,
    xmax = estimate + std.error * 1.645
  ),
  width = 0.2, size = 1, color = "black"
) +
geom_vline(xintercept = 0, color = "red", linetype = "longdash") +
theme(
  panel.spacing = unit(.5, "lines"),
  strip.background = element_blank(),
  strip.placement = "outside",
  axis.text.x = element_text(size = 15),
  axis.title.y = element_text(size = 15),
  axis.title.x = element_text(size = 15),
  legend.position = "bottom",
  plot.caption = ggtext::element_markdown(color = "black", size = 17),
  axis.text.y = element_text(size = 15),
  plot.title = ggtext::element_markdown(size = 20, color = "black",
                                        hjust = 0.5),
  plot.subtitle = ggtext::element_markdown(size = 17, color = "black",
                                           hjust = 0.5),
  panel.grid.major.y = element_line(colour = "#6d6d6d", size = 0.5,
                                    linetype = "dotted"),
  panel.grid.major.x = element_line(colour = "#6d6d6d", size = 0.5,
                                    linetype = "dotted"),
  panel.background = element_rect(fill = "white"),
  axis.ticks.length.x = unit(.2, "cm"),
  axis.ticks.x = element_line(color = "black"),
  axis.ticks.y = element_line(color = "black"),
  axis.ticks.length.y = unit(.2, "cm"),
  panel.border = element_blank(),
  axis.line = element_line(color = "black"),
  strip.text.x = element_text(size = 15)
) +
scale_x_continuous(limits = c(-4, 8.05), breaks = seq(-4, 8, 1),
                   expand = c(0, 0))
```



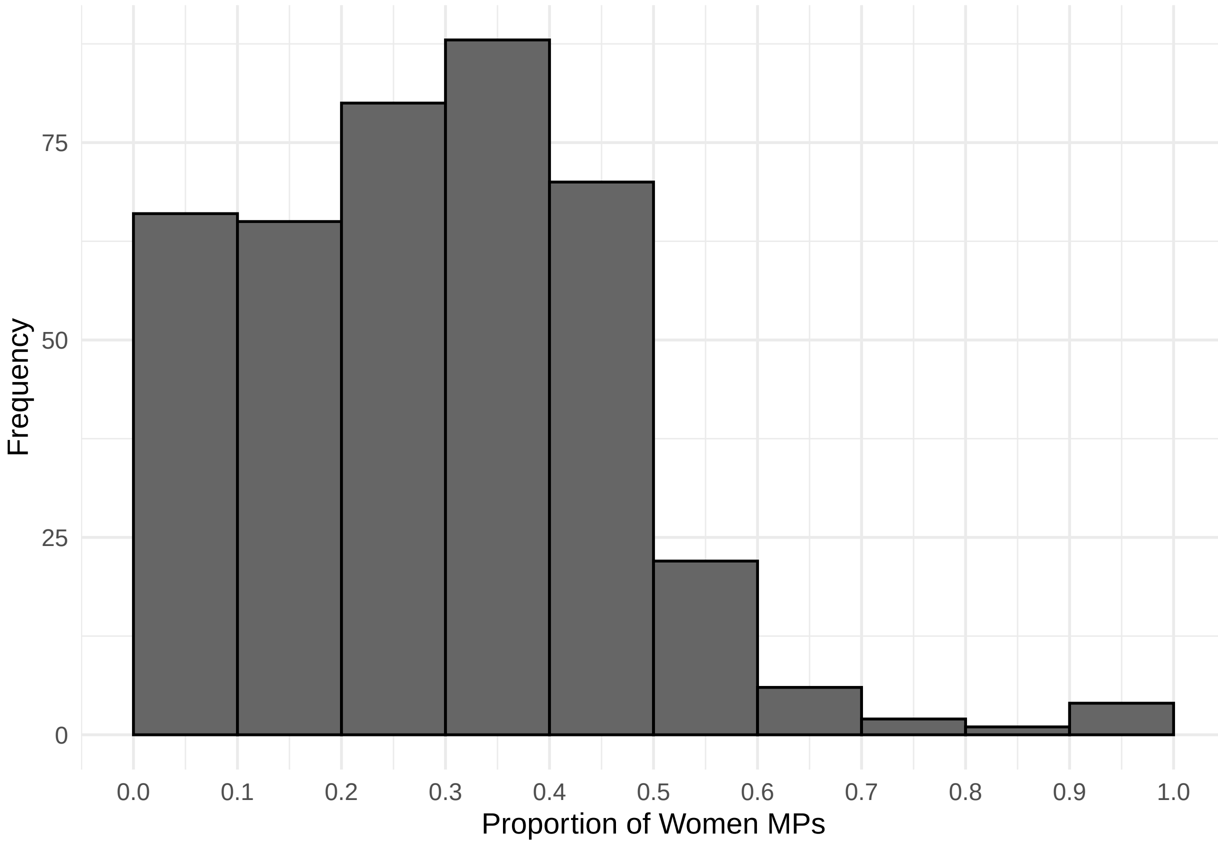
Tidyverse Code Simplification

```
###############################################
############## CREATING FIGURE 1 ###########
###############################################

load("_data/dyadic_data_1-4-22.Rdata") # Load data

dta_unique <- updated_data |>
  select(country, year, to_mp_number, to_rile, to_economy, to_society,
         to_femaleleader, to_pfeml) |> # Remove unneeded variables
  na.omit() |>
  distinct() # Keep only the unique parties being evaluated
```

Figure 1

```
# Graph 1
ggplot(dta_unique, aes(x = to_pfeml)) +
  geom_histogram(color = "black", fill = "grey40", binwidth = 0.1,
```

```r
                    center = 0.25) +
  scale_x_continuous(breaks = seq(0, 1, 0.1)) +
  theme_minimal() +
  theme(plot.title = element_text(size = 12)) +
  ylab("Frequency") +
  xlab("Proportion of Women MPs")
```




```r
ggsave("_graphs/fig1.pdf", width = 12, height = 10)
```

Table 1

```r
#############################################
###### CREATING TABLE 1 COLUMNS 1 & 2 ######
#############################################

dta <- updated_data |>
  mutate(cntryyr = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  filter(to_prior_seats >= 4) |> # Removing smaller parties
  select(
    year, country, rile_distance_s, prior_coalition, prior_opposition,
    econ_distance_s, society_distance_s, party_dislike, party_like, cntryyr,
    to_pfeml, to_prior_seats, to_mp_number
  ) |> # selecting variables
  na.omit()
```

```r
table1.1 <- lm(party_like ~ to_pfeml + as.factor(cntryyr), data = dta)
table1.2 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                 prior_opposition + as.factor(cntryyr), data = dta)

# With clustered ses
stargazer(
  type = "latex", table1.1, table1.2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table1.1, table1.2,
    clusters = dta$country
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:17 PM

Table 1:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	1.887^{***}	1.726^{***}
	(0.429)	(0.519)
	p = 0.00002	p = 0.001
rile_distance_s		-0.596^{***}
		(0.092)
		p = 0.000
prior_coalition		0.943^{***}
		(0.245)
		p = 0.0002
prior_opposition		0.371^{***}
		(0.108)
		p = 0.001
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,842	1,842
R^2	0.191	0.340
Adjusted R^2	0.154	0.308
Residual Std. Error	1.403 (df = 1760)	1.268 (df = 1757)
F Statistic	5.136^{***} (df = 81; 1760)	10.776^{***} (df = 84; 1757)
Note:		*p<0.1; **p<0.05; ***p<0.01

```r
############################################
###### CREATING TABLE 2 COLUMNS 3 & 4 ######
############################################

# Note in gendered data, the party_like and party_dislike variable indicate
# mean levels of like/dislike for party by ALL partisans
# the "dislike" variable indicates level of dislike towards out-party
# by partisans of specified gender

dta <- readRDS("_data/gender_disagregated_8-8-21.rds") |>
  mutate(countryyear = paste(country, dta$year, sep = "")) |>
  # creating the country-year fixed effects
  filter(to_prior_seats >= 4) |> # Removing smaller parties
  mutate(like = 10 - dislike) |>
  # Create Like variable for gendered data from dislike
  select(
    year, country, rile_distance_s, prior_coalition, prior_opposition,
    econ_distance_s, society_distance_s, to_pfeml,
    countryyear, gender, like, dislike, to_prior_seats
  ) |>
  na.omit()

# Only men subset
dta_male <- dta |> filter(gender == 1)
dta_female <- dta |> filter(gender == 2)

dta_female$countryyear <- as.factor(dta_female$countryyear)

tableS2.3 <- lm(like ~ to_pfeml + rile_distance_s + prior_coalition +
  prior_opposition + countryyear, data = dta_female)
tableS2.4 <- lm(like ~ to_pfeml + rile_distance_s + prior_coalition +
  prior_opposition + countryyear, data = dta_male)
```

Table S2 (women)

```r
# With clustered ses - women
stargazer(
  type = "latex", tableS2.3,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Out-Party Fixed Effects?", "Yes"),
                   c("Country-Level Clustered ses?", "Yes")),
  se = starprep(tableS2.3,
    clusters = dta_female$country
  ),
  keep = c("to_pfeml", "rile_distance_s", "prior_coalition",
           "prior_opposition", "to_pfeml2"),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:18 PM

Table 2:

	Dependent variable:
	like
to_pfeml	2.149***
	(0.479)
	p = 0.00001
rile_distance_s	−0.601***
	(0.093)
	p = 0.000
prior_coalition	0.976***
	(0.215)
	p = 0.00001
prior_opposition	0.356***
	(0.104)
	p = 0.001
Country-Year Fixed Effects?	Yes
Out-Party Fixed Effects?	Yes
Country-Level Clustered ses?	Yes
Observations	1,836
R^2	0.343
Adjusted R^2	0.252
Residual Std. Error	1.405 (df = 1613)
F Statistic	3.792*** (df = 222; 1613)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table S2 (men)

```
# With clustered ses - men
stargazer(
  type = "latex", tableS2.4,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Out-Party Fixed Effects?", "Yes"),
                   c("Country-Level Clustered ses?", "Yes")),
  se = starprep(tableS2.4,
    clusters = dta_male$country
  ),
  keep = c("to_pfeml", "rile_distance_s", "prior_coalition",
           "prior_opposition", "to_pfeml2"),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:19 PM

Table 3:

	Dependent variable:
	like
to_pfeml	1.189***
	(0.428)
	p = 0.006
rile_distance_s	−0.617***
	(0.080)
	p = 0.000
prior_coalition	0.988***
	(0.231)
	p = 0.00002
prior_opposition	0.342***
	(0.109)
	p = 0.002
Country-Year Fixed Effects?	Yes
Out-Party Fixed Effects?	Yes
Country-Level Clustered ses?	Yes
Observations	1,833
R^2	0.351
Adjusted R^2	0.260
Residual Std. Error	1.361 (df = 1605)
F Statistic	3.830*** (df = 227; 1605)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
############################################
############ CREATING TABLE S3 #############
############################################
```

```r
load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  mutate(cntryyr = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  select(
    year, country, rile_distance_s, prior_coalition, prior_opposition,
    econ_distance_s, society_distance_s, party_dislike, party_like, cntryyr,
    to_pfeml, from_rile, to_rile, from_left_bloc, from_right_bloc, to_left_bloc,
    to_right_bloc, from_parfam, to_parfam, to_prior_seats
  ) |>
  na.omit()

dta_nrr <- dta |> filter(to_parfam != 70)
dta_nrr <- dta_nrr |> filter(from_parfam != 70)

# Remove small parties, with fewer than 4 seats

dta_small_nrr <- dta_nrr |> filter(to_prior_seats >= 4)

table.S3 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                 prior_opposition + as.factor(country), data = dta_small_nrr)
```

Table S3 (stargazer)

```r
stargazer(
  type = "latex", table.S3,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered ses?", "Yes")),
  se = starprep(table.S3,
    clusters = dta_small_nrr$country
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:19 PM

```r
############################################
############ CREATING TABLE S3B ############
############################################

# Load
load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
```

Table 4:

	Dependent variable:
	party_like
to_pfeml	0.893***
	(0.260)
	p = 0.001
rile_distance_s	−0.576***
	(0.082)
	p = 0.000
prior_coalition	0.952***
	(0.214)
	p = 0.00001
prior_opposition	0.470***
	(0.093)
	p = 0.00000
Country-Year Fixed Effects?	Yes
Country-Level Clustered ses?	Yes
Observations	1,590
R^2	0.311
Adjusted R^2	0.301
Residual Std. Error	1.228 (df = 1566)
F Statistic	30.761*** (df = 23; 1566)
Note:	*p<0.1; **p<0.05; ***p<0.01

```r
    party_dislike, party_like, to_parfam, to_left_bloc,
    to_right_bloc, cntryyr, to_pfeml, to_prior_seats
  ) |>
  na.omit()

# Remove small parties, with fewer than 4 seats
dta_small <- dta |> filter(to_prior_seats >= 4)

table.3B.1 <- lm(party_like ~ to_pfeml + as.factor(cntryyr), data = dta_small)
table.3B.2 <- lm(party_like ~ to_pfeml + rile_distance_s + to_left_bloc +
                   prior_coalition + prior_opposition + as.factor(cntryyr),
                 data = dta_small)
```

Table S3B

```r
# With clustered ses
stargazer(
  type = "latex", table.3B.1, table.3B.2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.3B.1, table.3B.2,
    clusters = dta_small$country
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "to_left_bloc", "prior_coalition",
    "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:19 PM

```r
############################################
############ CREATING TABLE S4 ############
############################################

# Read in data
load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  mutate(countryyear = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  select(
    year, country, rile_distance_s, prior_coalition, prior_opposition,
    econ_distance_s, society_distance_s, party_dislike, party_like,
    countryyear, to_pfeml, from_rile, to_rile, logDM, to_left_bloc,
    to_prior_seats
  ) |>
  na.omit()

# Split by year, 1996-2006 and 2007-2017
dta_early <- dta |> filter(year <= 2006)
```

Table 5:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	1.887^{***}	1.300^{**}
	(0.429)	(0.622)
	p = 0.00002	p = 0.037
rile_distance_s		-0.603^{***}
		(0.094)
		p = 0.000
to_left_bloc		0.221^{*}
		(0.115)
		p = 0.054
prior_coalition		0.962^{***}
		(0.246)
		p = 0.0001
prior_opposition		0.366^{***}
		(0.109)
		p = 0.001
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,842	1,842
R^2	0.191	0.344
Adjusted R^2	0.154	0.312
Residual Std. Error	1.403 (df = 1760)	1.265 (df = 1756)
F Statistic	5.136^{***} (df = 81; 1760)	10.814^{***} (df = 85; 1756)
Note:		*p<0.1; **p<0.05; ***p<0.01

```r
dta_late <- dta |> filter(year >= 2007)

table.early <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                    prior_opposition + as.factor(countryyear), data = dta_early)
table.late <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                   prior_opposition + as.factor(countryyear), data = dta_late)

# Without small parties
dta_early_small <- dta_early |> filter(to_prior_seats >= 4)
dta_late_small <- dta_late |> filter(to_prior_seats >= 4)

table.4.1 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                  prior_opposition + as.factor(countryyear),
                data = dta_early_small)
table.4.2 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                  prior_opposition + as.factor(countryyear),
                data = dta_late_small)
```

Table 4.1 (stargazer)

```r
# With clustered ses
stargazer(
  type = "latex", table.4.1,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Out-Party Fixed Effects?", "Yes"),
                   c("Country-Level Clustered ses?", "Yes")),
  se = starprep(table.4.1,
    clusters = dta_early_small$country
  ),
  keep = c("to_pfeml", "rile_distance_s", "prior_coalition",
           "prior_opposition"),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:20 PM

Table 4.2 (stargazer)

```r
# With clustered ses
stargazer(
  type = "latex", table.4.2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Out-Party Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.4.2,
    clusters = dta_late_small$country
  ),
  keep = c("to_pfeml", "rile_distance_s", "prior_coalition",
           "prior_opposition"),
  report = "vc*sp"
)
```

Table 6:

	Dependent variable:
	party_like
to_pfeml	1.134**
	(0.548)
	p = 0.039
rile_distance_s	−0.697***
	(0.114)
	p = 0.000
prior_coalition	0.879***
	(0.281)
	p = 0.002
prior_opposition	0.371**
	(0.170)
	p = 0.029
Country-Year Fixed Effects?	Yes
Out-Party Fixed Effects?	Yes
Country-Level Clustered ses?	Yes
Observations	848
R^2	0.365
Adjusted R^2	0.332
Residual Std. Error	1.197 (df = 805)
F Statistic	11.017*** (df = 42; 805)
Note:	*p<0.1; **p<0.05; ***p<0.01

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:20 PM

Table 7:

	Dependent variable:
	party_like
to_pfeml	2.063***
	(0.688)
	p = 0.003
rile_distance_s	−0.505***
	(0.133)
	p = 0.0002
prior_coalition	0.966***
	(0.290)
	p = 0.001
prior_opposition	0.366*
	(0.191)
	p = 0.056
Country-Year Fixed Effects?	Yes
Out-Party Fixed Effects?	Yes
Country-Level Clustered SEs?	Yes
Observations	994
R^2	0.328
Adjusted R^2	0.296
Residual Std. Error	1.322 (df = 948)
F Statistic	10.278*** (df = 45; 948)
Note:	*p<0.1; **p<0.05; ***p<0.01

```r
############################################
############ CREATING TABLE S5 #############
############################################

load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  mutate(countryyear = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  select(
    year, country, rile_distance_s, prior_coalition, prior_opposition,
    econ_distance_s, society_distance_s, party_dislike, party_like, countryyear,
    to_pfeml, from_pfeml, diff_pfeml, to_prior_seats
  ) |>
  na.omit()

# Remove small parties, with fewer than 4 seats
dta_small <- dta |> filter(to_prior_seats >= 4)
```

```r
table.S5.1 <- lm(party_like ~ to_pfeml + from_pfeml + diff_pfeml +
                   as.factor(countryyear), data = dta_small)
table.S5.2 <- lm(party_like ~ to_pfeml + from_pfeml + diff_pfeml +
                   rile_distance_s + prior_coalition + prior_opposition +
                   as.factor(countryyear), data = dta_small)
```

Table 5

```r
# With clustered ses
stargazer(
  type = "latex", table.S5.1, table.S5.2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.S5.1, table.S5.2,
    clusters = dta_small$country
  ),
  keep = c(
    "to_pfeml", "from_pfeml", "diff_pfeml", "rile_distance_s",
    "prior_coalition", "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:20 PM

```r
############################################
########### CREATING FIG. S1 ############
############################################

# All values of Out-Party % women

plot_1 <- dta |>
  distinct(to_pfeml) |>
  mutate(
    from_pfeml = mean(dta$from_pfeml, na.rm = T) + sd(dta$from_pfeml, na.rm = T),
    # All 1 Sd above mean of in-party % women
    diff_pfeml = abs(to_pfeml - from_pfeml),
    # Create difference between in-and out-party women
    # Select other values (mean RILE distance, opposition together,
    # France 2012 country year)
    rile_distance_s = mean(dta$rile_distance_s, na.rm = T),
    prior_coalition = 0,
    prior_opposition = 1,
    countryyear = "France2012",
    to_mp_number = "31320",
    group = "above_mean"
  )

# All values of Out-Party % women
plot_2 <- dta |>
  distinct(to_pfeml) |>
  mutate(
```

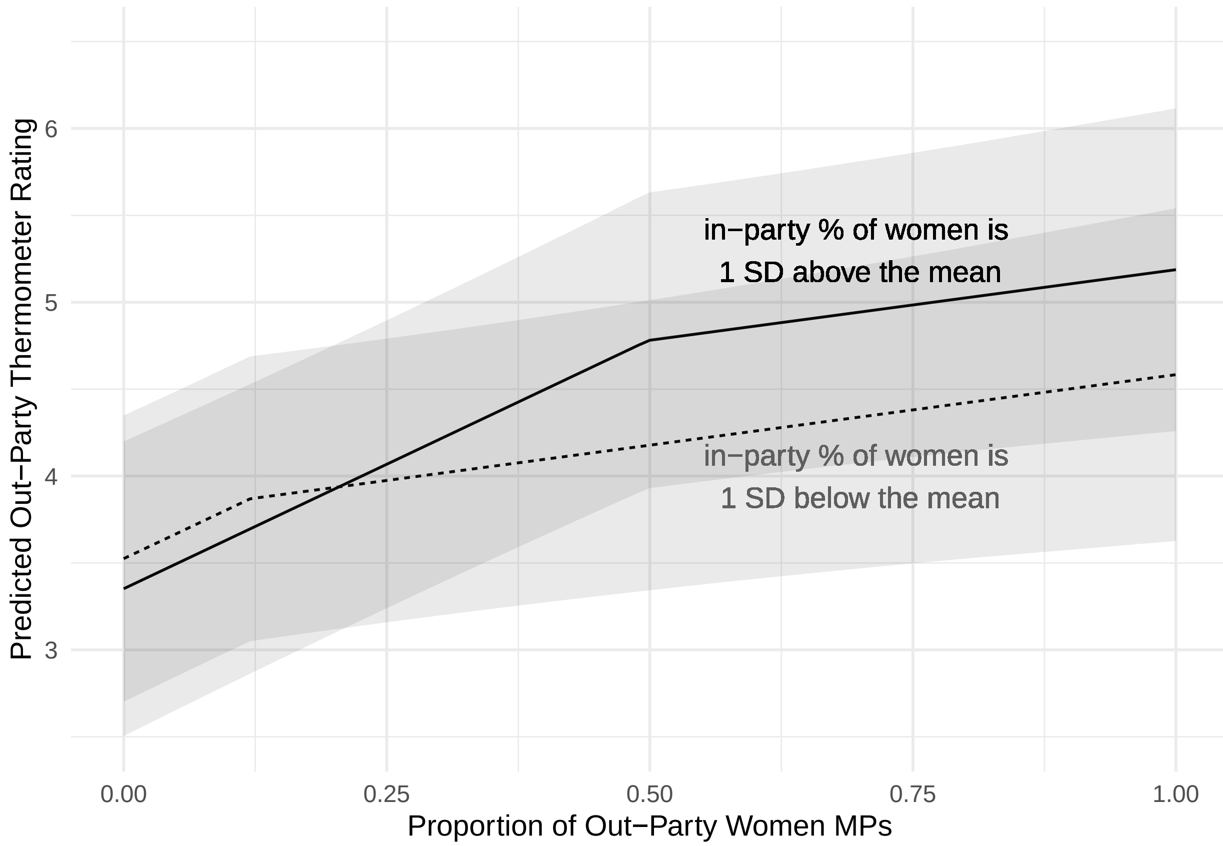
Table 8:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	2.079***	1.838***
	(0.473)	(0.554)
	p = 0.00002	p = 0.001
from_pfeml	0.773***	0.569***
	(0.227)	(0.215)
	p = 0.001	p = 0.009
diff_pfeml	−1.621**	−1.026**
	(0.664)	(0.519)
	p = 0.015	p = 0.048
rile_distance_s		−0.596***
		(0.076)
		p = 0.000
prior_coalition		0.891***
		(0.264)
		p = 0.001
prior_opposition		0.384***
		(0.104)
		p = 0.0003
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,593	1,593
R^2	0.216	0.359
Adjusted R^2	0.174	0.323
Residual Std. Error	1.379 (df = 1510)	1.248 (df = 1507)
F Statistic	5.076*** (df = 82; 1510)	9.924*** (df = 85; 1507)
Note:		*p<0.1; **p<0.05; ***p<0.01

```r
    from_pfeml = mean(dta$from_pfeml, na.rm = T) - sd(dta$from_pfeml, na.rm = T),
    # All 1 Sd below mean of In-party % women
    diff_pfeml = abs(to_pfeml - from_pfeml),
    # Create difference between in-and out-party women
    # Select other values (opposition together, France 2012 country year)
    rile_distance_s = mean(dta$rile_distance_s, na.rm = T),
    prior_coalition = 0,
    prior_opposition = 1,
    countryyear = "France2012",
    to_mp_number = "31320",
    group = "below_mean"
  )

plot_dta <- bind_rows(plot_1, plot_2)

## Plot based on table.S5.2

figureS1.data <- as.data.frame(predict(table.S5.2, newdata = plot_dta,
                                       interval = "confidence"))

plot_dta <- plot_dta |>
  mutate(
    fit = figureS1.data$fit,
    lwr = figureS1.data$lwr,
    upr = figureS1.data$upr
  )
```

Figure S1

```r
ggplot(plot_dta, aes(x = to_pfeml, y = fit, lty = group)) +
  geom_line() +
  geom_ribbon(
    aes(
      x = to_pfeml, y = fit, ymin = lwr,
      ymax = upr
    ),
    lwd = 1 / 2, alpha = 0.1
  ) +
  theme_minimal() +
  theme(plot.title = element_text(size = 12)) +
  ylab("Predicted Out-Party Thermometer Rating") +
  xlab("Proportion of Out-Party Women MPs") +
  theme(legend.position = "none") +
  geom_text(x = 0.70, y = 5.3,
            label = "in-party % of women is \n1 SD above the mean") +
  geom_text(x = 0.70, y = 4.0,
            label = "in-party % of women is \n1 SD below the mean",
            color = "grey37") +
  ylim(c(2.5, 6.5))
```




```r
ggsave("_graphs/figS1.pdf", width = 12, height = 10)

############################################
############ CREATING TABLE S6 #############
############################################

## Read in data
load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  mutate(countryyear = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
    party_dislike, party_like, countryyear, to_pfeml,
    from_rile, to_rile, logDM, to_left_bloc, to_prior_seats
  ) |>
  na.omit()

dta_small <- dta |> filter(to_prior_seats >= 4)

table.S6.1 <- lm(party_like ~ to_pfeml + rile_distance_s + logDM +
                   prior_coalition + prior_opposition + as.factor(year),
                 data = dta_small)
table.S6.2 <- lm(party_like ~ to_pfeml * logDM + rile_distance_s +
                   prior_coalition + prior_opposition + as.factor(year),
```

```r
                  data = dta_small)
table.S6.3 <- lm(party_like ~ to_pfeml * logDM + rile_distance_s * logDM +
                   prior_coalition * logDM + prior_opposition * logDM +
                   as.factor(year), data = dta_small)
```

Table S6

```r
stargazer(
  type = "latex", table.S6.1, table.S6.2, table.S6.3,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.S6.1, table.S6.2, table.S6.3,
    clusters = dta_small$country
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "logDM", "prior_coalition",
    "prior_opposition", "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:23 PM

```r
cat("### Table S6.1")

table.S6.1 |>
  report_table() |>
  filter(!grepl("year|Intercept", Parameter)) |>
  nice_table(highlight = TRUE, short = TRUE,
             col.format.custom = 2, format.custom = "format3")
```

```r
############################################
############ CREATING TABLE S7 #############
############################################

# Out party % women, non-clustered SEs
load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  mutate(cntryyr = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
    party_dislike, party_like, cntryyr, to_pfeml, to_prior_seats
  ) |>
  na.omit() |>
  mutate(to_pfeml2 = to_pfeml^2) |>
  # Creating squared term for out-party % women
  filter(to_prior_seats >= 4)

table.S7.1 <- lm(party_like ~ to_pfeml + to_pfeml2 +
                   as.factor(cntryyr), data = dta)
```

Table 9:

	Dependent variable:		
	party_like		
	(1)	(2)	(3)
to_pfeml	1.752***	2.119***	2.136***
	(0.416)	(0.813)	(0.818)
	p = 0.00003	p = 0.010	p = 0.010
rile_distance_s	−0.526***	−0.530***	−0.430***
	(0.076)	(0.075)	(0.150)
	p = 0.000	p = 0.000	p = 0.005
logDM	0.067	0.102*	0.130***
	(0.041)	(0.059)	(0.050)
	p = 0.103	p = 0.084	p = 0.010
prior_coalition	1.057***	1.060***	1.205**
	(0.203)	(0.204)	(0.482)
	p = 0.00000	p = 0.00000	p = 0.013
prior_opposition	0.339***	0.336***	0.451**
	(0.115)	(0.114)	(0.226)
	p = 0.004	p = 0.004	p = 0.047
to_pfeml:logDM		−0.126	−0.148
		(0.223)	(0.230)
		p = 0.573	p = 0.521
logDM:rile_distance_s			−0.038
			(0.059)
			p = 0.522
logDM:prior_coalition			−0.054
			(0.121)
			p = 0.655
logDM:prior_opposition			−0.039
			(0.063)
			p = 0.533
Country-Year Fixed Effects?	Yes		
Country-Level Clustered SEs?	Yes		
Observations	1,842	1,842	1,842
R^2	0.262	0.263	0.265
Adjusted R^2	0.252	0.252	0.252
Residual Std. Error	1.320 (df = 1815)	1.320 (df = 1814)	1.319 (df = 1811)
F Statistic	24.796*** (df = 26; 1815)	23.916*** (df = 27; 1814)	21.713*** (df = 30; 1811)
Note:			*p<0.1; **p<0.05; ***p<0.01

```r
table.S7.2 <- lm(party_like ~ to_pfeml + to_pfeml2 + rile_distance_s +
                   prior_coalition + prior_opposition + as.factor(cntryyr),
                 data = dta)
```

Table S7

```r
# With clustered ses
stargazer(
  type = "latex", table.S7.1, table.S7.2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.S7.1, table.S7.2,
    clusters = dta$country
  ),
  keep = c(
    "to_pfeml", "to_pfeml2", "rile_distance_s", "prior_coalition",
    "prior_opposition", "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:24 PM

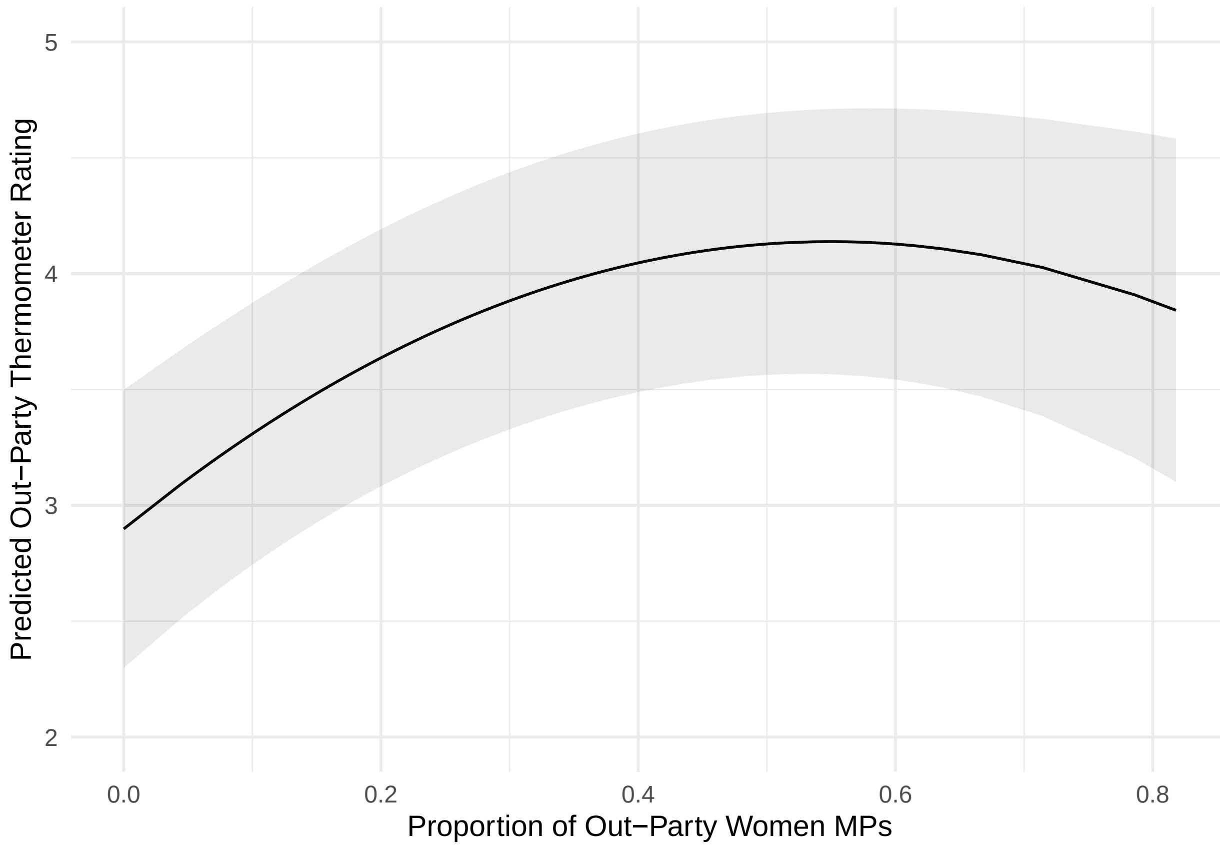
```r
############################################
############ CREATING FIG. S2 #############
############################################

## Create Plot Data

# All values of Out-Party % women

plot_S2 <- dta |>
  distinct(to_pfeml) |>
  mutate(
    to_pfeml2 = to_pfeml^2, # Create difference between in-and out-party women
    ## Select other values (mean RILE distance, opposition together,
    # France 2012 country year)
    rile_distance_s = mean(dta$rile_distance_s, na.rm = T),
    prior_coalition = 0,
    prior_opposition = 1,
    cntryyr = "France2012",
    to_mp_number = "31320"
  )

figureS2.data <- data.frame(predict(table.S7.2, newdata = plot_S2,
                                    interval = "confidence"))

plot_S2 <- plot_S2 |>
  mutate(
    fit = figureS2.data$fit,
    lwr = figureS2.data$lwr,
    upr = figureS2.data$upr
  )
```

Table 10:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	4.403***	4.513***
	(1.235)	(1.382)
	p = 0.0004	p = 0.002
to_pfeml2	−3.705***	−4.106***
	(1.421)	(1.511)
	p = 0.010	p = 0.007
rile_distance_s		−0.600***
		(0.092)
		p = 0.000
prior_coalition		0.927***
		(0.247)
		p = 0.0002
prior_opposition		0.385***
		(0.104)
		p = 0.0003
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,842	1,842
R^2	0.196	0.346
Adjusted R^2	0.159	0.315
Residual Std. Error	1.399 (df = 1759)	1.263 (df = 1756)
F Statistic	5.242*** (df = 82; 1759)	10.947*** (df = 85; 1756)
Note:		*p<0.1; **p<0.05; ***p<0.01

Figure S2

```
ggplot(plot_S2, aes(x = to_pfeml, y = fit)) +
  geom_line() +
  geom_ribbon(
    aes(
      x = to_pfeml, y = fit, ymin = lwr,
      ymax = upr
    ),
    lwd = 1 / 2, alpha = 0.1
  ) +
  theme_minimal() +
  theme(plot.title = element_text(size = 12)) +
  ylab("Predicted Out-Party Thermometer Rating") +
  xlab("Proportion of Out-Party Women MPs") +
  ylim(c(2, 5))
```




```
ggsave("_graphs/figS2.pdf", width = 12, height = 10)

#############################################
############ CREATING TABLE S8 #############
#############################################

# Women-led parties
load("_data/dyadic_data_1-4-22.Rdata")

dta_womenlead <- updated_data |>
```

```r
  filter(to_femaleleader == 1) |>
  mutate(countryyear = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
    party_dislike, party_like, countryyear, to_pfeml, to_prior_seats
  ) |>
  na.omit() |>
  filter(to_prior_seats >= 4)

table.S8A1 <- lm(party_like ~ to_pfeml + as.factor(countryyear),
                 data = dta_womenlead)
table.S8A2 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                   prior_opposition + as.factor(countryyear),
                 data = dta_womenlead)
```

Table S8A

```r
# With clustered SEs
stargazer(
  type = "latex", table.S8A1, table.S8A2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.S8A1, table.S8A2,
    clusters = dta_womenlead$country
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:26 PM

```r
# Male-led parties

load("_data/dyadic_data_1-4-22.Rdata")

dta_malelead <- updated_data |>
  filter(to_femaleleader == 0) |>
  mutate(countryyear = paste(country, year, sep = "")) |>
  # creating the country-year fixed effects
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
    party_dislike, party_like, countryyear, to_pfeml, to_prior_seats
  ) |>
  na.omit() |>
  filter(to_prior_seats >= 4) # Exclude small parties

table.S8B1 <- lm(party_like ~ to_pfeml + as.factor(countryyear),
```

Table 11:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	0.918	0.390
	(0.852)	(0.801)
	p = 0.282	p = 0.627
rile_distance_s		-0.821^{***}
		(0.148)
		p = 0.00000
prior_coalition		0.824^{***}
		(0.236)
		p = 0.0005
prior_opposition		0.399^{*}
		(0.225)
		p = 0.077
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	479	479
R^2	0.221	0.455
Adjusted R^2	0.132	0.389
Residual Std. Error	1.312 (df = 429)	1.101 (df = 426)
F Statistic	2.477^{***} (df = 49; 429)	6.844^{***} (df = 52; 426)
Note:		*p<0.1; **p<0.05; ***p<0.01

```
                 data = dta_malelead)
table.S8B2 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                   prior_opposition + as.factor(countryyear),
                 data = dta_malelead)
```

Table S8B

```
### With clustered SEs
stargazer(
  type = "latex", table.S8B1, table.S8B2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.S8B1, table.S8B2,
    clusters = dta_malelead$country
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:27 PM

```
############################################
############ CREATING TABLE S9 #############
############################################

## Read in data
load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  mutate(
    countryyear = paste(country, year, sep = ""),
    # creating the country-year fixed effects
    partydyad = paste(from_mp_number, to_mp_number, sep = "")
  ) |> # creating the party fixed effects / cluster
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
    party_dislike, party_like, countryyear, to_pfeml,
    from_rile, to_rile, to_mp_number, partydyad, to_prior_seats
  ) |>
  na.omit() |>
  filter(to_prior_seats >= 4) # Exclude small parties

table.S9 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                 prior_opposition + as.factor(countryyear), data = dta)
```

Table S9 (Stargazer)

Table 12:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	2.905^{***}	2.928^{***}
	(0.568)	(0.674)
	p = 0.00000	p = 0.00002
rile_distance_s		-0.546^{***}
		(0.090)
		p = 0.000
prior_coalition		0.979^{***}
		(0.261)
		p = 0.0002
prior_opposition		0.414^{***}
		(0.115)
		p = 0.0004
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,357	1,357
R^2	0.228	0.360
Adjusted R^2	0.179	0.318
Residual Std. Error	1.413 (df = 1275)	1.288 (df = 1272)
F Statistic	4.642^{***} (df = 81; 1275)	8.514^{***} (df = 84; 1272)
Note:		*p<0.1; **p<0.05; ***p<0.01

```r
### With clustered SEs
stargazer(
  type = "latex", table.S9,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Out-Party Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.S9,
    clusters = dta$partydyad
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:30 PM

Table 13:

	Dependent variable:
	party_like
to_pfeml	1.726^{***}
	(0.317)
	p = 0.00000
rile_distance_s	-0.596^{***}
	(0.044)
	p = 0.000
prior_coalition	0.943^{***}
	(0.125)
	p = 0.000
prior_opposition	0.371^{***}
	(0.073)
	p = 0.00000
Country-Year Fixed Effects?	Yes
Out-Party Fixed Effects?	Yes
Country-Level Clustered SEs?	Yes
Observations	1,842
R^2	0.340
Adjusted R^2	0.308
Residual Std. Error	1.268 (df = 1757)
F Statistic	10.776^{***} (df = 84; 1757)
Note:	*p<0.1; **p<0.05; ***p<0.01

```r
############################################
############ CREATING TABLE 10 #############
############################################
```

```r
load("_data/dyadic_data_1-4-22.Rdata")

dta <- updated_data |>
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
    party_dislike, to_pfeml, party_like, to_prior_seats
  ) |>
  na.omit() |>
  filter(to_prior_seats >= 4) # Remove small parties, with fewer than 4 seats

table.S10.1 <- lm(party_like ~ to_pfeml + as.factor(country), data = dta_small)
table.S10.2 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                    prior_opposition + as.factor(country), data = dta_small)
```

Table S10

```r
### With clustered SEs
stargazer(
  type = "latex", table.S10.1, table.S10.2,
  add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                   c("Country-Level Clustered SEs?", "Yes")),
  se = starprep(table.S10.1, table.S10.2,
    clusters = dta_small$country
  ),
  keep = c(
    "to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
    "econ_distance_s", "society_distance_s"
  ),
  report = "vc*sp"
)
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:13:31 PM

```r
##################################################
############ CREATING TABLES 11 & 12 #############
##################################################

load("_data/multilevel_1-5-22.Rdata")

indiv_data <- multilevel_data |>
  select(
    year, country, rile_distance_s, prior_coalition,
    prior_opposition, econ_distance_s, society_distance_s,
    cntryyr, to_pfeml, from_pfeml, thermometer_score, ID,
    party_to, party_from, from_partyname, to_partyname, to_left_bloc,
    from_left_bloc, to_right_bloc, from_right_bloc, gender, to_parfam,
    from_parfam, to_prior_seats, from_mp_number, to_mp_number
  ) |>
  mutate(
    gender = case_when(
      gender == "1" ~ "male",
```

Table 14:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	1.833***	1.656***
	(0.362)	(0.427)
	p = 0.00000	p = 0.0002
rile_distance_s		−0.550***
		(0.087)
		p = 0.000
prior_coalition		0.932***
		(0.248)
		p = 0.0002
prior_opposition		0.376***
		(0.105)
		p = 0.0004
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,842	1,842
R^2	0.142	0.282
Adjusted R^2	0.133	0.273
Residual Std. Error	1.421 (df = 1821)	1.301 (df = 1818)
F Statistic	15.071*** (df = 20; 1821)	31.014*** (df = 23; 1818)
Note:		*p<0.1; **p<0.05; ***p<0.01

```r
      gender == "2" ~ "female"
    ), # Create gender variable
    gender = as.factor(gender)
  ) |>
  filter(is.na(to_pfeml) == F) |>
  mutate(dyad = paste(from_mp_number, to_mp_number, sep = "_to_"))

### Create Table 11, column 1, with Standard errors clustered at country-year,
# party-dyad, and individual levels
table11A.1.1 <- feols(thermometer_score ~ to_pfeml | ID, data = indiv_data,
                      cluster = ~cntryyr)
table11A.1.2 <- feols(thermometer_score ~ to_pfeml | ID, data = indiv_data,
                      cluster = ~dyad)
table11A.1.3 <- feols(thermometer_score ~ to_pfeml | ID, data = indiv_data,
                      cluster = ~ID)

### Create Table 11, column 2, with Standard errors clustered at country-year,
# party-dyad, and individual levels
table11A.2.1 <- feols(thermometer_score ~ to_pfeml + +rile_distance_s +
                        prior_coalition + prior_opposition | ID,
                      data = indiv_data, cluster = ~cntryyr)
table11A.2.2 <- feols(thermometer_score ~ to_pfeml + +rile_distance_s +
                        prior_coalition + prior_opposition | ID,
                      data = indiv_data, cluster = ~dyad)
table11A.2.3 <- feols(thermometer_score ~ to_pfeml + +rile_distance_s +
                        prior_coalition + prior_opposition | ID,
                      data = indiv_data, cluster = ~ID)

### Create Table 11B, column 1, with Standard errors clustered at country-year,
# party-dyad, and individual levels
table11B.1.1 <- feols(thermometer_score ~ to_pfeml | cntryyr,
                      data = indiv_data, cluster = ~cntryyr)
table11B.1.2 <- feols(thermometer_score ~ to_pfeml | cntryyr,
                      data = indiv_data, cluster = ~dyad)
table11B.1.3 <- feols(thermometer_score ~ to_pfeml | cntryyr,
                      data = indiv_data, cluster = ~ID)

### Create Table 11B, column 2, with Standard errors clustered at country-year,
# party-dyad, and individual levels
table11B.2.1 <- feols(thermometer_score ~ to_pfeml + +rile_distance_s +
                        prior_coalition + prior_opposition | cntryyr,
                      data = indiv_data, cluster = ~cntryyr)
table11B.2.2 <- feols(thermometer_score ~ to_pfeml + +rile_distance_s +
                        prior_coalition + prior_opposition | cntryyr,
                      data = indiv_data, cluster = ~dyad)
table11B.2.3 <- feols(thermometer_score ~ to_pfeml + +rile_distance_s +
                        prior_coalition + prior_opposition | cntryyr,
                      data = indiv_data, cluster = ~ID)
```

Note: Tables 11 and 12 are not actually displayed in the original code.

Testing Different Ideologies

```
#### ___ ####

#### 0.2 - Load data ####

load("_data/multilevel_1-5-22.Rdata")

Data <- multilevel_data |>
  mutate(
    gender = case_when(
      gender == "1" ~ "male",
      gender == "2" ~ "female"
    ), # Create gender variable
    gender = as.factor(gender)
  ) |>
  filter(is.na(to_pfeml) == F) |>
  mutate(dyad = paste(from_mp_number, to_mp_number, sep = "_to_"))

#### ___ ####

#### 1 - Regression (basic) ####

model_1 <- lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                prior_opposition + as.factor(cntryyr), data = Data)
```

```
model_1_data <- broom::tidy(model_1)

model_1_conf_95 <- confint(model_1, level = 0.95) %>%
  data.frame() %>%
  rename(
    "conf.low_95" = "X2.5..",
    "conf.high_95" = "X97.5.."
  )

model_1_data_clean <- bind_cols(model_1_data, model_1_conf_95) |>
  filter(term %in% c("to_pfeml", "rile_distance_s", "prior_coalition",
                     "prior_opposition")) |>
  mutate(model = "rile_distance_s")

#### 1.1 - Regression (Other ideological scales) ####

#### nationalism_distance_s ####

model_2 <- lm(party_like ~ to_pfeml + nationalism_distance_s +
                prior_coalition + prior_opposition + as.factor(cntryyr),
              data = Data)
```

```
model_2_data <- broom::tidy(model_2)

model_2_conf_95 <- confint(model_2, level = 0.95) %>%
  data.frame() %>%
  rename(
    "conf.low_95" = "X2.5..",
```

```r
    "conf.high_95" = "X97.5.."
  )

model_2_data_clean <- bind_cols(model_2_data, model_2_conf_95) |>
  filter(term %in% c("to_pfeml", "nationalism_distance_s", "prior_coalition",
                     "prior_opposition")) |>
  mutate(model = "nationalism_distance_s")

#### econ_distance_s ####

model_3 <- lm(party_like ~ to_pfeml + econ_distance_s + prior_coalition +
                prior_opposition + as.factor(cntryyr), data = Data)
```

```r
model_3_data <- broom::tidy(model_3)

model_3_conf_95 <- confint(model_3, level = 0.95) %>%
  data.frame() %>%
  rename(
    "conf.low_95" = "X2.5..",
    "conf.high_95" = "X97.5.."
  )

model_3_data_clean <- bind_cols(model_3_data, model_3_conf_95) |>
  filter(term %in% c("to_pfeml", "econ_distance_s", "prior_coalition",
                     "prior_opposition")) |>
  mutate(model = "econ_distance_s")


#### galtan_distance_s ####
model_4 <- lm(party_like ~ to_pfeml + galtan_distance_s + prior_coalition +
                prior_opposition + as.factor(cntryyr), data = Data)
```

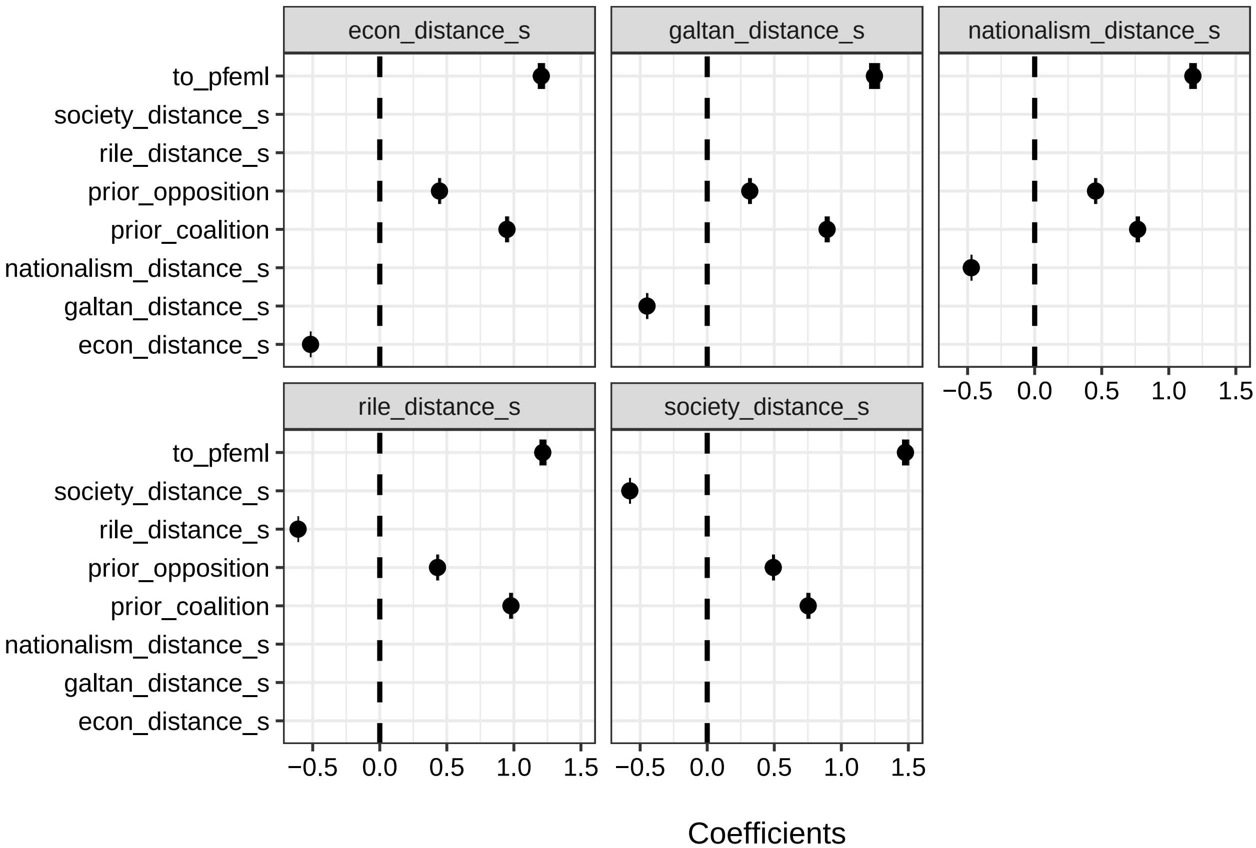
```r
model_4_data <- broom::tidy(model_4)

model_4_conf_95 <- confint(model_4, level = 0.95) %>%
  data.frame() %>%
  rename(
    "conf.low_95" = "X2.5..",
    "conf.high_95" = "X97.5.."
  )

model_4_data_clean <- bind_cols(model_4_data, model_4_conf_95) |>
  filter(term %in% c("to_pfeml", "galtan_distance_s", "prior_coalition",
                     "prior_opposition")) |>
  mutate(model = "galtan_distance_s")

#### society_distance_s ####
model_5 <- lm(party_like ~ to_pfeml + society_distance_s + prior_coalition +
                prior_opposition + as.factor(cntryyr), data = Data)
```

```r
model_5_data <- broom::tidy(model_5)

model_5_conf_95 <- confint(model_5, level = 0.95) %>%
  data.frame() %>%
```

```r
  rename(
    "conf.low_95" = "X2.5..",
    "conf.high_95" = "X97.5.."
  )

model_5_data_clean <- bind_cols(model_5_data, model_5_conf_95) |>
  filter(term %in% c("to_pfeml", "society_distance_s", "prior_coalition",
                     "prior_opposition")) |>
  mutate(model = "society_distance_s")

#### Bind dataset ####
Bind_model_data <- bind_rows(
  model_1_data_clean,
  model_2_data_clean,
  model_3_data_clean,
  model_4_data_clean,
  model_5_data_clean
)
```

Final Graph

```r
### Final graph ###
ggplot(Bind_model_data, aes(x = term, y = estimate)) +
  geom_point(size = 6) +
  facet_wrap(~model) +
  coord_flip() +
  geom_hline(yintercept = 0, colour = "black", lty = 2, size = 2) +
  scale_x_discrete("") +
  scale_y_continuous("\nCoefficients") +
  geom_point(aes(x = term, y = estimate)) +
  geom_linerange(aes(x = term, ymin = conf.low_95, ymax = conf.high_95),
                 lwd = 1 / 2) +
  geom_errorbar(aes(ymin = conf.low_95, ymax = conf.high_95), size = 10,
                width = 0) +
  theme_bw(base_size = 25) +
  theme(
    plot.title = element_text(size = 22, hjust = 0.5, colour = "black"),
    axis.text = element_text(size = 21, colour = "black"),
    plot.caption = element_text(size = 21, hjust = 0, colour = "black")
  )
```




```r
ggsave("_graphs/ideological_scale.png", width = 12, height = 8)
```

Imputing Missing Data

```r
load("_data/dyadic_data_1-4-22.Rdata")
```

First, let us check the number of rows in the data provided by the original authors.

```r
nrow(updated_data)
```

```
## [1] 2266
```

```r
dta2 <- updated_data |>
  mutate(cntryyr = paste(country, year, sep = "")) |> # creating the country-year fixed effects
  filter(to_prior_seats >= 4) |>  # Removing smaller parties
  select(year, country, rile_distance_s, prior_coalition, prior_opposition,
         econ_distance_s, society_distance_s, party_dislike, party_like, cntryyr,
         to_pfeml, to_prior_seats, to_mp_number) # selecting variables

dta <- dta2 |>
  na.omit()
```

Then the number of rows in transformed data.

```r
nrow(dta)
```

```
## [1] 1842
```

And the difference between the two, so how many rows were lost.

```r
nrow(updated_data) - nrow(dta)
```

```
## [1] 424
```

We have thus lost 424 observations. Yet, we only have 177 missing observations as part of the variables necessary for Table 1.1, as demonstrated below.

```r
updated_data |>
  select(party_like, to_pfeml, cntryyr) |>
  na.omit() |>
  nrow() - nrow(dta)
```

```
## [1] 177
```

For Table 1.2, we also have only 177 missing observations.

```r
true.missing <- updated_data |>
  select(party_like, to_pfeml, cntryyr, rile_distance_s,
         prior_coalition, prior_opposition) |>
  na.omit() |>
  nrow() - nrow(dta)
true.missing
```

```
## [1] 177
```

From the 424 originally excluded observations, we can remove the 177 true missing to see how many observations were excluded unnecessarily.

```r
nrow(updated_data) - nrow(dta) - true.missing
```

```
## [1] 247
```

Thus, the original authors excluded 247 observations unnecessarily. Yet, the row difference betweeen the two data sets is still only 56, not 247, probably because we are losing some rows (368) when removing smaller parties with `filter(to_prior_seats >= 4)` when creating dta2.

```r
nrow(dta2) - nrow(dta)
```

```
## [1] 56
```

Tables Comparison

Let us now compare results with and without the missing data.

```r
table1.1 <-lm(party_like ~ to_pfeml + as.factor(cntryyr), data = dta)
table1.2 <-lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                prior_opposition + as.factor(cntryyr), data = dta)

table1.1_2 <-lm(party_like ~ to_pfeml + as.factor(cntryyr), data = dta2)
table1.2_2 <-lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                  prior_opposition + as.factor(cntryyr), data = dta2)
```

```r
# With clustered ses
stargazer(type = "latex", table1.1, table1.2,
          add.lines = list(c("Country-Year Fixed Effects?", "Yes"), c("Country-Level Clustered SEs?", "
          se = starprep(table1.1, table1.2,
                        clusters = dta$country),
          keep = c("to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
                   "econ_distance_s", "society_distance_s"),
  report = "vc*sp")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:14:00 PM

Table 15:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	1.887^{***}	1.726^{***}
	(0.429)	(0.519)
	p = 0.00002	p = 0.001
rile_distance_s		-0.596^{***}
		(0.092)
		p = 0.000
prior_coalition		0.943^{***}
		(0.245)
		p = 0.0002
prior_opposition		0.371^{***}
		(0.108)
		p = 0.001
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,842	1,842
R^2	0.191	0.340
Adjusted R^2	0.154	0.308
Residual Std. Error	1.403 (df = 1760)	1.268 (df = 1757)
F Statistic	5.136^{***} (df = 81; 1760)	10.776^{***} (df = 84; 1757)
Note:		*p<0.1; **p<0.05; ***p<0.01

```r
stargazer(type = "latex", table1.1_2, table1.2_2,
          add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
                           c("Country-Level Clustered SEs?", "Yes")),
          se = starprep(table1.1_2, table1.2_2,
                        clusters = dta$country),
          keep = c("to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
                   "econ_distance_s", "society_distance_s"),
  report = "vc*sp")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:14:00 PM

> The conclusion thus far is that the results from the two data sets are perfectly identical, even though they differ on 56 observations (out of 1898, ~3% of the data). Thus, using `na.omit()` was unnecessary since it did not generate an error and did not change the results.

At the same time, I have noticed that the number of clusters used in the `se` argument of the stargazer function above still uses the `dta` dataset (with `na.omit`), not the one with the missing values. It is likely why the results are identical.

Table 16:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	1.887^{***}	1.726^{***}
	(0.429)	(0.519)
	p = 0.00002	p = 0.001
rile_distance_s		-0.596^{***}
		(0.092)
		p = 0.000
prior_coalition		0.943^{***}
		(0.245)
		p = 0.0002
prior_opposition		0.371^{***}
		(0.108)
		p = 0.001
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,842	1,842
R^2	0.191	0.340
Adjusted R^2	0.154	0.308
Residual Std. Error	1.403 (df = 1760)	1.268 (df = 1757)
F Statistic	5.136^{***} (df = 81; 1760)	10.776^{***} (df = 84; 1757)
Note:		*p<0.1; **p<0.05; ***p<0.01

se a list of numeric vectors that will replace the default coefficient values for each model. Behaves exactly like the argument coef.

I have tried changing it to `dta2`, but this leads to an error:

```
"Error in commarobust(x, se_type = se_type, clusters = clusters, alpha = alpha) :
  `clusters` must be the same length as the model data"
```

It is likely that this is because `lm()` uses `na.omit()` on the data, but not on the cluster. Indeed, na.omit is the default treatment of missing values in lm:

> `na.action` a function which indicates what should happen when the data contain NAs. The default is set by the `na.action`

Setting `na.action = NULL` in the lm models leads to an error, so that is not an option. Our only choice then might be to impute the missing data. However, this at least explains why the original authors used `na.omit()`.

```
"Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :
  NA/NaN/Inf in 'x'"
```

Imputing Missing Data

...

What if we impute the missing data? Here, we will impute missing data with the `missForest` package, as it is one of the best imputation methods.

Imputation

```
# Need logical and character variables as factors for missForest
# "Error: Can not handle categorical predictors with more than 53 categories."
new.data <- dta2 %>%
  select(-cntryyr) %>% # Too many categories (> 53)
  mutate(across(c(where(is.character), where(is.logical)), as.factor)) %>%
  as.data.frame()

# Parallel processing
registerDoParallel(cores = 4)

# Variables
set.seed(100)
data.imp <- missForest(new.data, verbose = TRUE, parallelize = "variables")
```

```
##   parallelizing over the variables of the input data matrix 'xmis'
##   missForest iteration 1 in progress...done!
##     estimated error(s): 0.000002240226 0
##     difference(s): 0.0000000000002255581 0
##     time: 1.86 seconds
##
##   missForest iteration 2 in progress...done!
##     estimated error(s): 0.000002178801 0
##     difference(s): 0.00000000000007116891 0
##     time: 1.61 seconds
##
##   missForest iteration 3 in progress...done!
##     estimated error(s): 0.00000224641 0
```

```
##     difference(s): 0.000000000000009959122 0
##     time: 1.56 seconds
```

```
# Extract imputed dataset
dta3 <- data.imp$ximp

# Add back country-year
dta3 <- dta3 %>%
  mutate(cntryyr = dta2$cntryyr)
```

Details

Why impute the data? van Ginkel explains,

> Regardless of the missingness mechanism, multiple imputation is always to be preferred over listwise deletion. Under MCAR it is preferred because it results in more statistical power, under MAR it is preferred because besides more power it will give unbiased results whereas listwise deletion may not, and under NMAR it is also the preferred method because it will give less biased results than listwise deletion.

van Ginkel, J. R., Linting, M., Rippe, R. C. A., & van der Voort, A. (2020). Rebutting existing misconceptions about multiple imputation as a method for handling missing data. *Journal of Personality Assessment*, *102*(3), 297-308. https://doi.org/10.1080/00223891.2018.1530680

Why `missForest`? It outperforms other imputation methods, including the popular MICE (multiple imputation by chained equations). You also don't end up with several datasets, which makes it easier for following analyses. Finally, it can be applied to mixed data types (missings in numeric & categorical variables).

Waljee, A. K., Mukherjee, A., Singal, A. G., Zhang, Y., Warren, J., Balis, U., ... & Higgins, P. D. (2013). Comparison of imputation methods for missing laboratory data in medicine. *BMJ open*, *3*(8), e002847. https://doi.org/10.1093/bioinformatics/btr597

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*(1), 112-118. https://doi.org/10.1093/bioinformatics/btr597

Final Tables Comparison

Let us now compare results with and without the imputed data.

```
table1.1_3 <-lm(party_like ~ to_pfeml + as.factor(cntryyr), data = dta3)
table1.2_3 <-lm(party_like ~ to_pfeml + rile_distance_s + prior_coalition +
                  prior_opposition + as.factor(cntryyr), data = dta3)

# With clustered ses
stargazer(type = "latex", table1.1, table1.2,
          add.lines = list(c("Country-Year Fixed Effects?", "Yes"), c("Country-Level Clustered SEs?", "Y
          se = starprep(table1.1, table1.2,
                        clusters = dta$country),
          keep = c("to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
                   "econ_distance_s", "society_distance_s"),
  report = "vc*sp")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:14:06 PM

```
stargazer(type = "latex", table1.1_3, table1.2_3,
          add.lines = list(c("Country-Year Fixed Effects?", "Yes"),
```

Table 17:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	1.887***	1.726***
	(0.429)	(0.519)
	p = 0.00002	p = 0.001
rile_distance_s		−0.596***
		(0.092)
		p = 0.000
prior_coalition		0.943***
		(0.245)
		p = 0.0002
prior_opposition		0.371***
		(0.108)
		p = 0.001
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,842	1,842
R^2	0.191	0.340
Adjusted R^2	0.154	0.308
Residual Std. Error	1.403 (df = 1760)	1.268 (df = 1757)
F Statistic	5.136*** (df = 81; 1760)	10.776*** (df = 84; 1757)
Note:		*p<0.1; **p<0.05; ***p<0.01

```
                              c("Country-Level Clustered SEs?", "Yes")),
          se = starprep(table1.1_3, table1.2_3,
                        clusters = dta3$country),
          keep = c("to_pfeml", "rile_distance_s", "prior_coalition", "prior_opposition",
                   "econ_distance_s", "society_distance_s"),
  report = "vc*sp")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: Thu, Feb 22, 2024 - 5:14:07 PM

Table 18:

	Dependent variable:	
	party_like	
	(1)	(2)
to_pfeml	1.890***	1.741***
	(0.443)	(0.527)
	p = 0.00002	p = 0.001
rile_distance_s		−0.605***
		(0.088)
		p = 0.000
prior_coalition		0.907***
		(0.245)
		p = 0.0003
prior_opposition		0.366***
		(0.104)
		p = 0.0005
Country-Year Fixed Effects?	Yes	
Country-Level Clustered SEs?	Yes	
Observations	1,898	1,898
R^2	0.198	0.347
Adjusted R^2	0.162	0.316
Residual Std. Error	1.403 (df = 1816)	1.267 (df = 1813)
F Statistic	5.536*** (df = 81; 1816)	11.453*** (df = 84; 1813)
Note:		*p<0.1; **p<0.05; ***p<0.01

After imputation, the numbers differ a little bit (good sign that they are not identical!). However, the results appear pretty robust since they are very similar and nothing has changed signed or significance threshold.