Analysing Vote Choice Data

Assignment 2

Jacob Edenhofer*

10 May 2023

Preliminaries

Let us import the necessary packages and the data:

```
# packages
library(tidyverse)
library(here)
library(modelsummary)
library(janitor)
library(scales)
library(haven)
library(ggpubr)
library(coefplot)
library(MASS)
library(knitr)
library(kableExtra)
library(ggeffects)
library(margins)

# data
deV07 <- read_dta(paste0(here(), "/Data/ITANES1994_2001.dta"))
```

Exercise 1

1.1

In this exploratory phase, get also a sense of the missing values in the dataset: if you want to measure extremism, you need to make sure that you know the left-right placement of parties and respondents. Following De Vries (2007), create a dummy variable for parties and individuals with "extreme" righ-left self-)placements (Hint: you should ask yourself how much an individual/party differ from the average individual/party). Following De Vries (2007), build a measure of "issue salience" for unemployment (Hint: in order to measure issue salience, you

*jacob.edenhofer@some.ox.ac.uk

should ask yourself what is the share of respondents that think that that issue is very relevant. However, do not forget that this dataset is longitudinal, meaning that there are more years).

To get a feel for the data, I start by creating a table summarising all variables, save for `year` and `prtystd` since the former is co-extensive with `prtyst_name` and the latter only take three values (1994, 1996, 2001). To that end, I run:

```
deV07 %>%
  dplyr::select(-c(year, prtystd)) %>%
  datasummary_skim(histogram = F,
                   title = "Summary statistics") %>%
  kable_styling(latex_options = "hold_position")
```

Table 1: Summary statistics

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
place_self	6	6	2.9	1.3	1.0	3.0	5.0
place_votedprty	6	23	2.9	1.3	1.0	3.0	5.0
gndr	2	0	1.4	0.5	1.0	1.0	2.0
edu	8	0	4.5	1.6	1.0	5.0	7.0
unemp_issue	3	19	0.5	0.5	0.0	0.0	1.0
debate	4	0	2.0	0.8	1.0	2.0	3.0
relig	6	3	3.1	1.3	1.0	3.0	5.0

The second column indicates the number of levels for each factor, with the seventh and eighth columns indicating the range of the variables. The third column indicates that there are significant numbers of missing values for two variables in particular, namely `place_votedprty` and `unemp_issue`.

To create the dummy variable, I consider the median and standard deviations for `place_self` and `place_votedprty` respectively. Since the median is three for both for three variables and the standard deviation (roughly) one, I consider those parties/individuals extreme whose distance from the median exceeds one standard deviation. This applies for individuals and parties whose values for the respective variables are one and five respectively. Furthermore, I rescale the `gndr` variable to the 0/1 format. Finally, to create the unemployment salience variable, `unemp_salience_per_year`, I compute the share of respondents per year for whom unemployment is the most important issue. To wrangle the data in that way, I run:

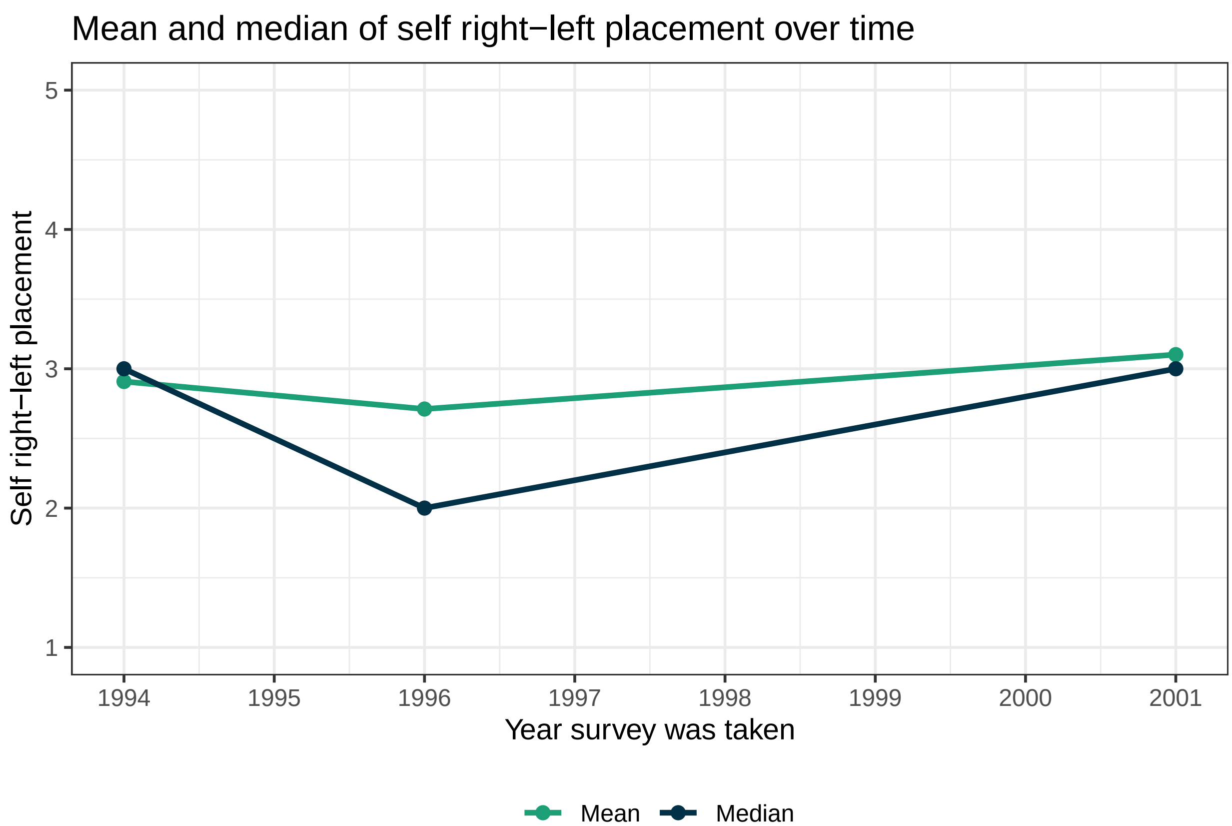
```
deV07_mod <- deV07 %>%
  mutate(gndr_dummy = ifelse(gndr == 1, 1, 0), # 1 for male, 0 for female
         extreme_dummy_individuals = ifelse(place_self %in% c(1, 5), 1, 0), # extreme dummy for individ
         extreme_dummy_parties = ifelse(place_votedprty %in% c(1, 5), 1, 0)) %>% # extreme dummy for pa
  group_by(year) %>% # group by year to account for longitudinal nature of data
  mutate(unemp_salience_per_year = sum(unemp_issue == 1, na.rm = T)/n(),
         n = n()) %>%
  ungroup()
```

1.2

Plot the mean right-left self-placement of individuals over years. Has the public become more or less (left-) right-wing? Can you actually say anything meaningful about this? Compare the mean and the median values for each year.

To create the desired plot, I run:

```
deV07_mod %>%
  group_by(year) %>%
  summarise(mean_lr = mean(place_self, na.rm = T),
            median_lr = median(place_self, na.rm = T)) %>%
  pivot_longer(cols = !year, names_to = "measure", values_to = "values") %>%
  ggplot(aes(x = year, y = values, colour = measure)) +
  geom_point(size = 2) +
  geom_line(linewidth = 1) +
  expand_limits(y = c(1, 5)) +
  scale_colour_manual("",
                      labels = c("mean_lr" = "Mean",
                                 "median_lr" = "Median"),
                      values = c("#1C9E77", "#023047")) +
  scale_x_continuous("Year survey was taken",
                     breaks = seq(1994, 2001, 1)) +
  labs(y = "Self right-left placement",
       title = "Mean and median of self right-left placement over time") +
  theme_bw() +
  theme(legend.position = "bottom")
```



We can see that, in 1994 and 2001, the mean and median are both roughly three, suggesting that in these two years the distribution of ideological self-placement is roughly normal. In 1996, however, the mean is higher than the median, i.e. the median has ideologically shifted to the right. Because of the labelling of `place_self` (1 == right, 5 == left), the distribution of `place_self` in that election year is, in statistical language, skewed to the left, which, substantively, is due to an increase in the share of fairly extreme right-wing respondents in that year.

Without taking into account the variance or standard deviation of the mean of `place_self`, we cannot determine whether the ideological right-ward shift in 1996 is a statistically significant one. To that end, I run the following code:

```
deV07_mod %>%
  group_by(year) %>%
  summarise(round(mean_cl_normal(place_self), digits = 2)) %>%
  kbl(booktabs = T,
      col.names = c("Year", "Mean", "Lower bound of 95% CI",
                    "Upper bound of 95% CI"),
      caption = "Mean of ideological shift") %>%
  kable_styling(latex_options = "hold_position")
```

Since the lower bound of the 1994 mean value is greater than the upper bound for the 1996 mean value, we can

Table 2: Mean of ideological shift

Year	Mean	Lower bound of 95% CI	Upper bound of 95% CI
1994	2.91	2.81	3.01
1996	2.71	2.65	2.77
2001	3.10	3.03	3.17

conclude[1] that the ideological right-ward shift is statistically significant at the 5% level. The overall conclusion is then that there is a significant right-ward shift from the 1994 to 1996 survey, which, however, is fully reversed by 2001.

Following De Vries (2007), measure the perceived distance betwen voters and their party of choice. Why would do you square it? And what is the main implication for results interpretation?

To create the desired variable, I subtract `place_self` from `place_votedprty` and then square that difference.

```
deV07_mod <- deV07_mod %>%
  mutate(squared_distance = (place_self-place_votedprty)^2)
```

The resulting difference is referred to as the *Euclidean* distance, which is minimised when the voter's self-placement equals the party's ideological position, as perceived by that voter. Consequently, increases in the Euclidean distance indicate increases in the perceived ideological distance between a voter-party pair, with the square ensuring that (i) left and right differences are treated identically, and (ii) that greater deviations are weighted more heavily, i.e. the possible differences are 0, 1, 4, 9, and 16.

1.3

What pushes individuals to vote for parties with extreme right-left placements? Estimate the appropriate model and report it a nice, tidy table. Explain why you controlled for the covariates that you picked.What model did you use in the previous point? Why? Plot the coefficients and comment their magnitude and level of statistical significance

My dependent variable is the `extreme_dummy_parties` variable created in 1.1, which is unity of voters perceive a party as either extreme right or left wing. To identify the correlates of extreme voting, I run six models.

- In the first model, I simply regress the dependent variable on `squared_distance`. This is motivated by the spatial theory of voting, which predicts that individuals will vote for those parties that they perceive to be ideologically closest to them. Hence, I expect the coefficient estimate to be *negative* since increases `squared_distance` mean lower ideological proximity, as explained in 1.2.
- In the second model, I add the `gndr_dummy`. This is motivated by recent findings (e.g. Anduiza and Rico 2022; Oshri et al. 2022) that men are, ceteris paribus, more likely than women to vote for radical right parties. Therefore, I expect the coefficient estimate to be *positive*[2], though this is more of a tentative hypothesis. This is because the dependent variable also includes radical left parties and the theoretical expectations are less clear for that party family.

[1] Non-overlapping confidence intervals imply significance, while the overlapping ones do not allow us to make any inferences about significance (Gailmard 2014).

[2] The dummy variable, recall, is unity for men and zero otherwise.

- The third model adds education, which is widely considered to be one of the most powerful predictors of vote choice. Given that higher values indicate greater education, I expect the coefficient estimate to be *negative*.
- The fourth model adds religiosity, reflecting theoretical arguments relating religiosity to extreme voting. Some argue that greater religiosity decreases the probability of voting for an extreme party by creating a strong attachment to religious parties, notably the Christian democrats (Arzheimer and Carter 2009). Others contend that religiosity can push individuals to the extreme right on account of promoting social conservatism (Marcinkiewicz and Dassonneville 2022). Hence, the sign is theoretically indeterminate.
- The fifth model adds the salience of unemployment in a given election year, with greater salience potentially pushing voters towards the extremes (*positive* coefficient estimate). On the other hand, voters may care more about competence in economically hard times, and elect moderate parties, provided, of course, that they believe these parties to be more competent.
- The sixth model adds a proxy for political interest. Given the coding of the variable, I expect a *positive* coefficient estimate, indicating that less interested individuals are more likely to vote for extreme parties, holding all other included variables constant.

```
extreme_party1 <- glm(extreme_dummy_parties ~ squared_distance,
                      family = binomial(link = "logit"),
                      data = deV07_mod)
extreme_party2 <- glm(extreme_dummy_parties ~ squared_distance + gndr_dummy,
                      family = binomial(link = "logit"),
                      data = deV07_mod)
extreme_party3 <- glm(extreme_dummy_parties ~ squared_distance + gndr_dummy + edu,
                      family = binomial(link = "logit"),
                      data = deV07_mod)
extreme_party4 <- glm(extreme_dummy_parties ~ squared_distance + gndr_dummy + edu + relig,
                      family = binomial(link = "logit"),
                      data = deV07_mod)
extreme_party5 <- glm(extreme_dummy_parties ~ squared_distance + gndr_dummy + edu + relig + unemp_salie
                      family = binomial(link = "logit"),
                      data = deV07_mod)
extreme_party6 <- glm(extreme_dummy_parties ~ squared_distance + gndr_dummy + edu + relig + unemp_salie
                      family = binomial(link = "logit"),
                      data = deV07_mod)
# modelsummary
modelsummary(list(extreme_party1, extreme_party2, extreme_party3, extreme_party4, extreme_party5, extrem
             estimate = "{estimate}{stars}",
             coef_map = c("squared_distance" = "Perceived squared distance",
                          "gndr_dummy" = "Gender dummy",
                          "edu" = "Education",
                          "relig" = "Religiosity",
                          "unemp_salience_per_year" = "Salience of unemployment",
                          "debate" = "Political interest"),
```

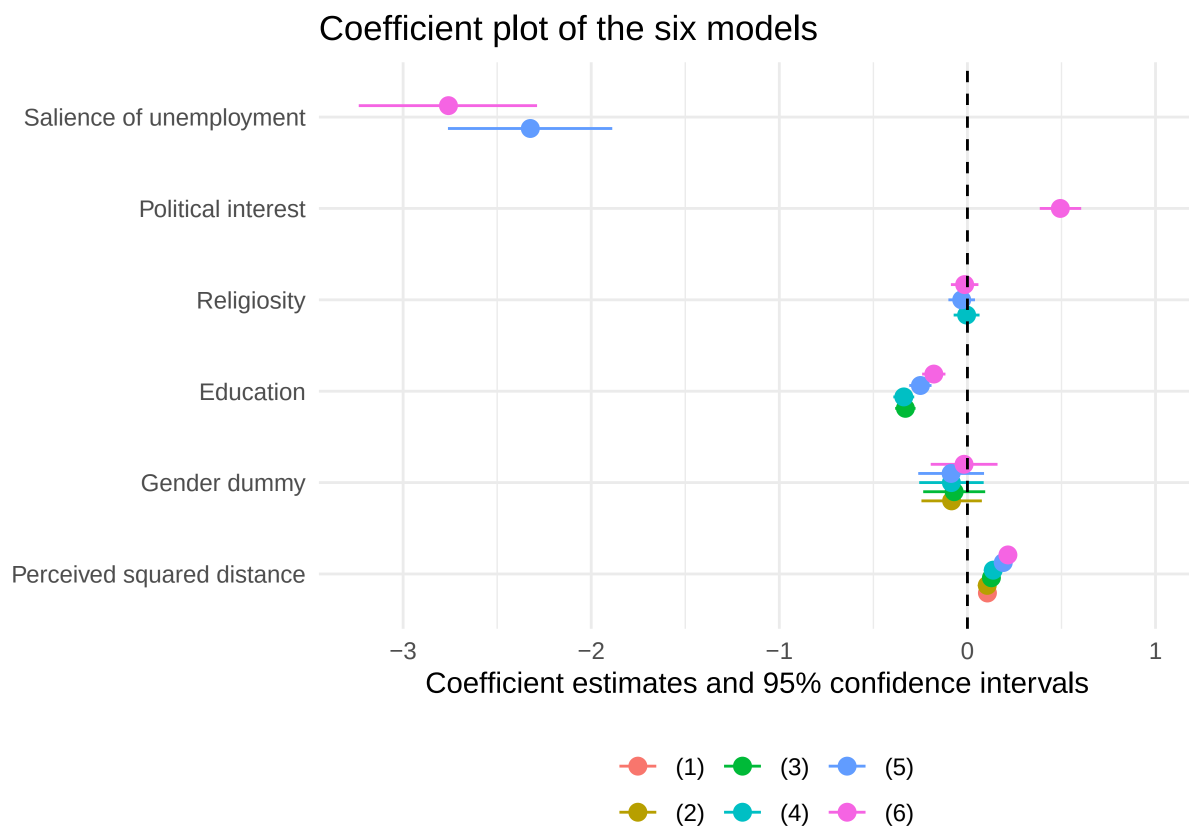
```
            title = "Correlates of extreme voting") %>%
  kable_styling(latex_options = "hold_position")
```

Table 3: Correlates of extreme voting

	(1)	(2)	(3)	(4)	(5)	(6)
Perceived squared distance	0.106***	0.105***	0.127***	0.137***	0.190***	0.215***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.017)	(0.018)
Gender dummy		−0.084	−0.071	−0.085	−0.087	−0.018
		(0.082)	(0.084)	(0.087)	(0.089)	(0.091)
Education			−0.330***	−0.338***	−0.250***	−0.179***
			(0.028)	(0.028)	(0.030)	(0.032)
Religiosity				−0.004	−0.030	−0.015
				(0.035)	(0.036)	(0.037)
Salience of unemployment					−2.324***	−2.759***
					(0.223)	(0.242)
Political interest						0.494***
						(0.056)
Num.Obs.	3050	3050	3049	2952	2952	2951
AIC	3610.4	3611.4	3463.0	3301.1	3193.3	3112.7
BIC	3622.5	3629.5	3487.1	3331.1	3229.3	3154.6
Log.Lik.	−1803.224	−1802.697	−1727.509	−1645.564	−1590.667	−1549.336
F	55.684	28.351	62.673	48.398	54.552	51.648
RMSE	0.45	0.45	0.44	0.43	0.43	0.42

To represent the results in a coefficient plot, I use the `modelplot()` function because the latter can be more easily integrated with the `kableExtra` package than the `coefplot()` function.

```
modelplot(list(extreme_party1, extreme_party2, extreme_party3, extreme_party4, extreme_party5, extreme_
         coef_map = c("squared_distance" = "Perceived squared distance",
                      "gndr_dummy" = "Gender dummy",
                      "edu" = "Education",
                      "relig" = "Religiosity",
                      "debate" = "Political interest",
                      "unemp_salience_per_year" = "Salience of unemployment")) +
  expand_limits(x = 1) +
  geom_vline(xintercept = 0, linetype = "dashed") +
  labs(title = "Coefficient plot of the six models") +
  theme(legend.position = "bottom")
```



The above plot and table 3 reveal four statistically significant correlates of extreme voting:

- The `squared_distance` term is significant, but the sign is the opposite of what spatial voting theory would lead us to expect. That is, as the squared distance increases by one unit (i.e. as there is less ideological proximity), the probability of extreme voting increases. Specifically, the log odds increase by roughly 24% $(100*(exp(0.215)-1))$, holding all other included covariates constant.
- Greater levels of education are, as expected, significantly associated with a lower probability of extreme voting, with the log odds decreasing by roughly 16.4% $(100*(exp(-0.179)-1))$ for a unit increase in the level of education.
- As the share of voters that regard unemployment as the most important issue rises, the probability of extreme voting decreases significantly - the log odds increase by approximately 95% $(100*(exp(-2.759)-1))$. Put simply: there is a strong inverse association between the salience of unemployment and extreme voting. This is consistent with the "competence" mechanism outlined above, though the results should not be interpreted as direct evidence for that mechanism.
- As expected, less politically interested voters are, ceteris paribus, more likely to vote for extreme parties, with the log odds increasing by roughly 64% $(100*(exp(0.494)-1))$ for a unit increase in political interest.

Exercise 2

2.1

What is the effect of gender on interest for politics (measured as watching TV debate)? Does it change when individuals are very concerned by unemployment? If so, how? Write down the formal models you have estimated, explain what the coefficients represent, and interpret the results.

To examine the relationship of interest, I regress the `debate` variable on the gender dummy in the first model, and add an interaction term between the latter and the unemployment issue dummy in the second model. The interaction terms allows us to examine whether the effect of gender on political interest varies by the importance respondents attach to unemployment as a political issue. In the third model, I add education and religiosity as covariates to probe the robustness of the results obtained from the previous models. My estimating equation for this final model is:

$$log(\frac{Interest_i}{1-Interest_i}) = \alpha + \beta_1 Gender_i + \beta_2 Unemployment_i + \beta_3 Gender_i * Unemployment_i$$
$$+\beta_4 Education_i + \beta_5 Religiosity_i + \epsilon_i$$

In R, I estimate this ordered logisitic regression by running:

```
# create dummy variable
deV07_mod <- deV07_mod %>%
  mutate(interest_dummy = ifelse(debate %in% c(1, 2), 1, 0),
         interest_factor = factor(debate),
         # factorise to increase interpretability
         gndr_dummy_f = factor(gndr_dummy,
                               levels = c("0", "1"),
                               labels = c("Female", "Male")),
         unemp_issue_f = factor(unemp_issue,
                               levels = c("0", "1"),
                               labels = c("No", "Yes")))

# estimate logit model
interest1 <- polr(interest_factor ~ gndr_dummy_f,
                 data = deV07_mod,
                 Hess = T)
interest2 <- polr(interest_factor ~ gndr_dummy_f*unemp_issue_f,
                 data = deV07_mod,
                 Hess = T)
interest3 <- polr(interest_factor ~ gndr_dummy_f*unemp_issue_f + edu + relig,
                 data = deV07_mod,
                 Hess = T)

# modelsummary
modelsummary(list(interest1, interest2, interest3),
```

```r
            estimate = "{estimate}{stars}",
            coef_map = c("1|2" = "1 vs 2",
                         "2|3" = "2 vs 3",
                         "gndr_dummy_fMale" = "Gender",
                         "unemp_issue_fYes" = "Unemployment most important issue",
                         "gndr_dummy_fMale:unemp_issue_fYes" = "Gender x Unemployment most important i
                         "edu" = "Education",
                         "relig" = "Religiosity"),
            title = "Association between interest in politics and gender") %>%
  kable_styling(latex_options = "hold_position")
```

Table 4: Association between interest in politics and gender

	(1)	(2)	(3)
1 vs 2	−0.682***	−1.033***	−3.894***
	(0.046)	(0.073)	(0.170)
2 vs 3	0.511***	0.024	−2.676***
	(0.045)	(0.070)	(0.163)
Gender	−0.342***	−0.296**	−0.359***
	(0.058)	(0.092)	(0.098)
Unemployment most important issue		−0.546***	−0.426***
		(0.097)	(0.102)
Gender x Unemployment most important issue		−0.022	0.043
		(0.130)	(0.137)
Education			−0.542***
			(0.024)
Religiosity			−0.077**
			(0.027)
Num.Obs.	4089	3299	3278
AIC	8902.8	7057.9	6450.6
BIC	8921.7	7088.4	6493.2
RMSE	1.83	1.91	1.91

To interpret the coefficient estimates, it is worth writing out the conditional expectations the above regression implies. To that end, let us define ϕ as a shorthand for $log(\frac{Interest_i}{1-Interest_i})$, and let $\hat{Educ}$ and $\hat{Relig}$ denote fixed values of these covariates. Then, we can write:

$$\mathbb{E}[\phi|Gender_i = 0, Unemployment_i = 0, \hat{Educ}_i, \hat{Relig}_i] = \alpha + \beta_4 \hat{Educ}_i + \beta_5 \hat{Relig}_i$$
$$\mathbb{E}[\phi|Gender_i = 1, Unemployment_i = 0, \hat{Educ}_i, \hat{Relig}_i] = \alpha + \beta_1 + \beta_4 \hat{Educ}_i + \beta_5 \hat{Relig}_i$$
$$\mathbb{E}[\phi|Gender_i = 0, Unemployment_i = 1, \hat{Educ}_i, \hat{Relig}_i] = \alpha + \beta_2 + \beta_4 \hat{Educ}_i + \beta_5 \hat{Relig}_i$$
$$\mathbb{E}[\phi|Gender_i = 1, Unemployment_i = 1, \hat{Educ}_i, \hat{Relig}_i] = \alpha + \beta_1 + \beta_2 + \beta_3 + \beta_4 \hat{Educ}_i + \beta_5 \hat{Relig}_i$$

Subtracting the first equation from the second gives us β_1. Substantively, β_1 therefore represents the marginal effect of gender on political interest, i.e. the change in the expected probability of watching TV debates for males (gender = 1) compared to females (gender = 0), holding the values of education and religiosity constant. Similarly, subtracting the third equation from the first gives us β_2. Hence, β_2 represents the marginal effect of unemployment importance on political interest, i.e. the change in the expected probability of watching TV debates for those viewing unemployment as the most important political issue, compared to those who do not do so, holding the values of education and religiosity constant. Finally, the coefficient on the interaction term is obtained by subtracting (i) equation three from four, (ii) subtracting one from two, and (iii) subtracting the second difference from the first. Intuitively, this means that the interaction term represents the difference in the marginal effect of gender between those who view unemployment as the most important issue, compared to those who do not, holding all other covariates constant. Finally, the intercepts represent the probability of "choosing" level one (high interest), as opposed to level two, and "choosing" level two, as opposed to level three (no interest).

Table 4 allows us to draw four lessons:

- Males exhibit significantly greater political interest than females, holding all other included covariates constant.
- Similarly, those considering unemployment the most important political issue exhibit, ceteris paribus, greater political interest than those who do consider unemployment as important.
- The marginal effect of gender on political interest does not significantly vary with the unemployment dummy, as the coefficient estimate in the third row shows.
- Both greater religiosity and higher education levels are positively and significantly associated with political interest.

2.2

Represent graphically: The predicted probabilities of gender on interest in politics when unemployment is the most important issue (or not), The marginal effects of gender on interest in politics when unemployment is the most important issue (or not), Put the two graphs in a singular imagine

To plot the predicted probabilities, I use the `ggeffects` package. The plot below includes connecting lines as visual aids for tracking how the effect of gender varies across unemployment.

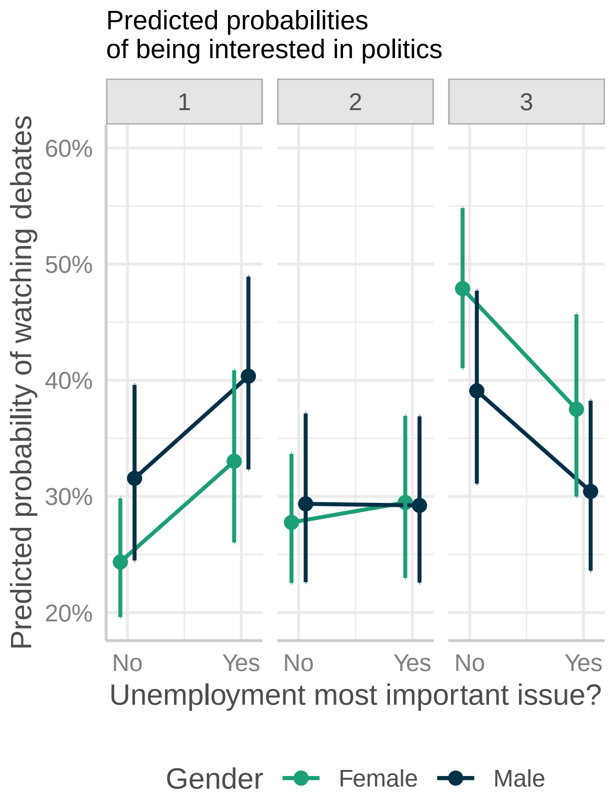
```
pp_plot <- plot(ggpredict(interest3, terms = c("unemp_issue_f", "gndr_dummy_f")),
     connect.lines = T) +
  scale_colour_manual("Gender",
                      labels = c("0" = "Female",
                                 "1" = "Male"),
                      values = c(c("#1C9E77", "#023047"))) +
  expand_limits(y = 0.6) +
  labs(y = "Predicted probability of watching debates",
       x = "Unemployment most important issue?",
       title = "Predicted probabilities\nof being interested in politics",
       caption = "Covariates include: education and religiosity.") +
  theme(legend.position = "bottom") +
  theme(plot.title = element_text(size = 10),
        plot.caption = element_text(hjust = 0, size = 6))
```

To compute the marginal effect, I use the `slopes()` function from the `marginaleffects` package.

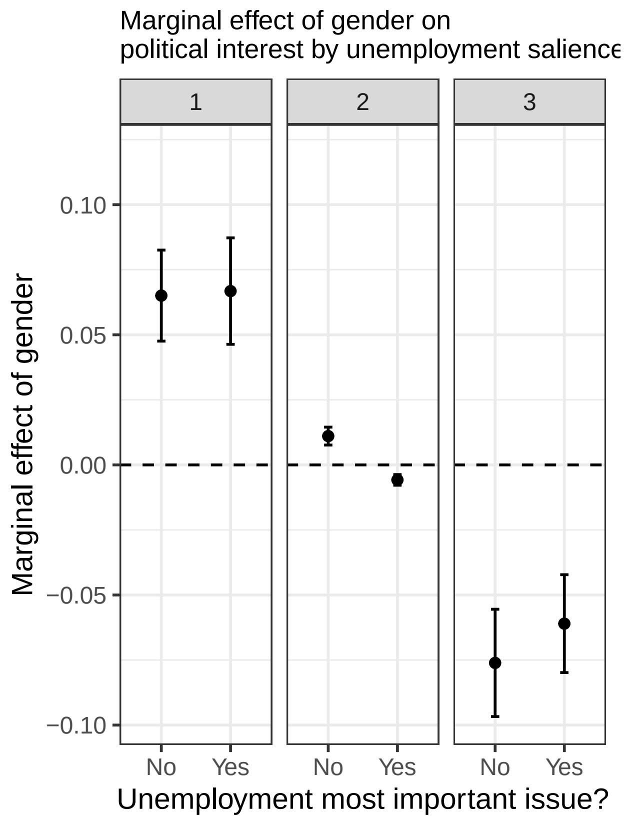
```
# compute marginal effect
interest4 <- MASS::polr(interest_factor ~ gndr_dummy*unemp_issue_f + edu + relig,
                 data = deV07_mod,
                 Hess = T)

# marginal effect
me_plot <- marginaleffects::slopes(interest4,
                        variables = "gndr_dummy", by = "unemp_issue_f") %>%
  ggplot(aes(x = unemp_issue_f, y = estimate,
             group = factor(term))) +
  geom_point() +
  geom_errorbar(aes(ymin = estimate - std.error, ymax = estimate + std.error),
                width = 0.12) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  expand_limits(y = 0.12) +
  facet_wrap(~group) +
  labs(y = "Marginal effect of gender",
       x = "Unemployment most important issue?",
       title = "Marginal effect of gender on\npolitical interest by unemployment salience",
       caption = "Covariates include: education and religiosity.") +
  theme_bw() +
  theme(plot.title = element_text(size = 10),
        plot.caption = element_text(hjust = 0, size = 6))
```

I combine the two plots using the `ggarrange()` function.

```
ggarrange(pp_plot, me_plot, nrow = 1)
```



Covariates include: education and religiosity.



Covariates include: education and religiosity.

As explained above, the marginal effect of gender is the vertical difference in predicted probabilities for males and females respectively. Analogously, the marginal effect of unemployment is the horizontal difference between the blue and green point estimates respectively. The two horizontal differences are almost identical. This, in combination with the almost completely overlapping confidence intervals for the marginal effects at levels one and two, confirms our conclusion that the effect of gender does not change significantly with unemployment.

2.3

Estimate a multinomial model: who do individuals with an extreme right-left self-placement vote for? Report it in a nice, tidy table.

To run a multinomial model, I start by factorising the `prtystd_name` variable (with the right-wing Lega as my reference category), and regressing it on the `extreme_dummy_individuals` dummy as well as gender, education, religiosity and the unemployment importance dummy.

```
# factorise DV
deV07_mod <- deV07_mod %>%
  mutate(prtystd_name_f = relevel(factor(prtystd_name), "Lega"))

# model
multinomial <- nnet::multinom(prtystd_name_f ~ extreme_dummy_individuals + gndr_dummy + edu + relig + u
                     data = deV07_mod, trace = F)
```

```
# modelsummary
modelsummary(multinomial,
             estimate = "{estimate}{stars}",
             title = "Multinomial model for predicting vote choice",
             output = "data.frame") %>%
  filter(grepl("extreme_dummy_individuals|(Intercept)", term),
         grepl("estimate", statistic)) %>%
  dplyr::select(-1) %>%
  kbl(booktabs = T) %>%
  kable_styling(latex_options = "hold_position") %>%
  add_footnote(label = "The model includes the following covariates: gender, education, religiosity and
```

term	statistic	(1)
(Intercept)	estimate	-0.196
(Intercept)	estimate	-1.093
(Intercept)	estimate	0.000
(Intercept)	estimate	1.025
(Intercept)	estimate	3.113***
(Intercept)	estimate	3.683***
extreme_dummy_individuals	estimate	0.381
extreme_dummy_individuals	estimate	0.719
extreme_dummy_individuals	estimate	0.975*
extreme_dummy_individuals	estimate	0.987+
extreme_dummy_individuals	estimate	1.303**
extreme_dummy_individuals	estimate	2.000***

The model includes the following covariates: gender, education, religiosity and unemployment importance.

Given the reference category, the coefficient estimates show that extreme individuals, as opposed to moderate ones, are more likely to vote for RC and UCD than the Lega.[3]

2.4

What is a Conditional Logit Model, and does it differ from a Multinomial Logit Model? What types of data should be used (or need to be used) for this type of model? What is the best advantage of a Conditional Logit Model?

A conditional logit model is essentially a fixed-effects model with a binary outcome variable. For multinomial logit models, by contrast, the dependent variable is a multi-level factor variable, rather than a binary one. Given the fixed effects, conditional logit models require repeated cross-sectional or panel data.

The most significant advantage of such models, as with all fixed effects models, is that they allow us to restrict attention to variation within a given time period. In this way, we can control for unobserved time-invariant unobservables, which, generally, strengthens our confidence that the relationship we observe is not spurious. Despite that, fixed effects are usually not sufficient to claim causality since we often cannot rule out the presence of unobservable, time-varying confounders.

[3] See the appendix for the ugly summary output, which includes the parties' names.

Appendix

```r
summary(nnet::multinom(prtystd_name_f ~ extreme_dummy_individuals,
                       data = deV07_mod, trace = F))
```

```
## Call:
## nnet::multinom(formula = prtystd_name_f ~ extreme_dummy_individuals,
##     data = deV07_mod, trace = F)
##
## Coefficients:
##     (Intercept) extreme_dummy_individuals
## AN     1.130424                 0.5885418
## FI     2.549992                -0.2266472
## PDS    1.993200                 0.6826851
## PPI    2.467315                 0.6188516
## RC    -1.082562                 2.2324194
## UDC    2.188262                 0.2238060
##
## Std. Errors:
##     (Intercept) extreme_dummy_individuals
## AN    0.1460749                 0.2887893
## FI    0.1318678                 0.2741743
## PDS   0.1353790                 0.2730906
## PPI   0.1322787                 0.2693127
## RC    0.2524841                 0.3647497
## UDC   0.1339328                 0.2743771
##
## Residual Deviance: 13473.09
## AIC: 13497.09
```

References

Anduiza, Eva, and Guillem Rico. 2022. "Sexism and the Far-Right Vote: The Individual Dynamics of Gender Backlash." *American Journal of Political Science*.

Arzheimer, Kai, and Elisabeth Carter. 2009. "Christian Religiosity and Voting for West European Radical Right Parties." *West European Politics* 32 (5): 985–1011.

Gailmard, Sean. 2014. *Statistical Modeling and Inference for Social Science*. Cambridge: Cambridge University Press.

Marcinkiewicz, Kamil, and Ruth Dassonneville. 2022. "Do Religious Voters Support Populist Radical Right Parties? Opposite Effects in Western and East-Central Europe." *Party Politics* 28 (3): 444–56.

Oshri, Odelia, Liran Harsgor, Reut Itzkovitch-Malka, and Or Tuttnauer. 2022. "Risk Aversion and the Gender Gap in the Vote for Populist Radical Right Parties." *American Journal of Political Science*.