Analysing Vote Choice Data

Final assignment

Jacob Edenhofer*

09 May 2023

Preliminaries

```
# load relevant packages
library(tidyverse)
library(haven)
library(modelsummary)
library(survey)
library(here)
library(geffects)
library(margins)

# import data
gles <- read_dta(pasteO(here(), "/Data/german_longitudinal_election_study_cross_section_post_election20
gles1 <- read_dta(pasteO(here(), "/Data/gles_panel_wave20.dta"))</pre>
```

Next, we will create some new variables:

^{*}jacob.edenhofer@some.ox.ac.uk

```
labels = c("male", "female")),
year_born1 = as.numeric(as.character(year_born)),
age = 2021 - as.numeric(as.character(year_born)),
spd_21 = ifelse(btw21_zweitstimme == 4, 1, 0),
union_21 = ifelse(btw21_zweitstimme == 1, 1, 0),
gruene_21 = ifelse(btw21_zweitstimme == 6, 1, 0),
fdp_21 = ifelse(btw21_zweitstimme == 5, 1, 0),
afd_21 = ifelse(btw21_zweitstimme == 322, 1, 0),
linke_21 = ifelse(btw21_zweitstimme == 7, 1, 0),
spd_to_switch = ifelse(btw21_zweitstimme == 4 & btw17_zweitstimme != 4, 1, 0),
afd_away_switch = ifelse(btw17_zweitstimme == 322 & btw21_zweitstimme != 322, 1, 0),
constituency_centric_rep = ifelse(q63a < 0, NA, q63a),</pre>
party_centric_rep = ifelse(q63c < 0, NA, q63c),</pre>
household_income = ifelse(d63 < 0, NA, d63),</pre>
household_income_factor = as.factor(household_income),
bachelor_dummy = ifelse(d8j1 < 0, NA, d8j1),</pre>
school = ifelse(d7 < 0, NA, d7),
abitur = ifelse(d7 == 5, 1, 0),
abitur_factor = ifelse(abitur == 1, "abitur", "no_abitur"),
urban_rural = ifelse(wum6 < 0, NA, wum6),</pre>
urban rural factor = as.factor(urban rural),
subjective_class = ifelse(d38 < 0, NA, d38),</pre>
left_right_self = ifelse(q37 < 0, NA, q37),</pre>
left right self factor = as.factor(left right self),
left_right_cdu = ifelse(q35b < 0, NA, q35b),</pre>
left_right_cdu_factor = as.factor(left_right_cdu),
distance_cdu = (left_right_cdu-left_right_self)^2,
left_right_csu = ifelse(q35c < 0, NA, q35c),</pre>
left_right_csu_factor = as.factor(left_right_csu),
distance_csu = (left_right_csu-left_right_self)^2,
left_right_spd = ifelse(q35d < 0, NA, q35d),</pre>
left_right_spd_factor = as.factor(left_right_spd),
distance_spd = (left_right_spd-left_right_self)^2,
left_right_afd = ifelse(q35h < 0, NA, q35h),</pre>
left_right_afd_factor = as.factor(left_right_afd),
distance_afd = (left_right_afd-left_right_self)^2,
left_right_fdp = ifelse(q35e < 0, NA, q35e),</pre>
left right fdp factor = as.factor(left right fdp),
distance_fdp = (left_right_fdp-left_right_self)^2,
left_right_green = ifelse(q35f < 0, NA, q35f),</pre>
left_right_green_factor = as.factor(left_right_green),
distance_green = (left_right_green-left_right_self)^2,
left_right_linke = ifelse(q35g < 0, NA, q35g),</pre>
```

```
left_right_linke_factor = as.factor(left_right_linke),
distance_linke = (left_right_linke-left_right_self)^2,
scholz_love = ifelse(q18b < 0, NA, q18b),</pre>
scholz_love_factor = as.factor(scholz_love),
finzanz_abgehangt_subjektiv = ifelse(q46a < 0, NA, q46a),</pre>
finzanz_abgehangt_subjektiv_factor = as.factor(finzanz_abgehangt_subjektiv),
arbeit_abgehant_subjektiv = ifelse(q46b < 0, NA, q46b),</pre>
arbeit_abgehant_subjektiv_factor = as.factor(arbeit_abgehant_subjektiv),
cancel_culture_subjektiv = ifelse(q46d < 0, NA, q46d),</pre>
cancel culture subjektiv factor = as.factor(cancel culture subjektiv),
infrastruktur subjektiv = ifelse(q46c < 0, NA, q46c),</pre>
infrastruktur_subjektiv_factor = as.factor(infrastruktur_subjektiv),
unemployed_last10_yrs = ifelse(d17a < 0, NA, d17a),</pre>
unemployed_last10yrs_months = ifelse(d17b < 0, NA, d17b),</pre>
unemployed_last10yrs_weeks = ifelse(d17c < 0, NA, d17c),</pre>
unemployed_dummy = ifelse(unemployed_last10_yrs != 0, 1, 0),
unemployed_dummy_factor = as.factor(unemployed_dummy),
trust_in_politicians = ifelse(q79d < 0, NA, q79d),</pre>
trust_in_politicians_factor = as.factor(trust_in_politicians),
trust_in_parliament = ifelse(q79b < 0, NA, q79b),</pre>
trust_in_parliament_factor = as.factor(trust_in_parliament),
trust_in_parties = ifelse(q79c < 0, NA, q79c),</pre>
trust_in_parties_factor = as.factor(trust_in_parties),
trust in public broadcast = ifelse(q79i < 0, NA, q79i),
trust_in_public_broadcast_factor = as.factor(trust_in_public_broadcast),
trust_general = ifelse(q78 < 0, NA, q78),</pre>
trust_general_factor = as.factor(trust_general),
out_group_minorities_assim = ifelse(q125a < 0, NA, q125a),</pre>
out_group_minorities_assim_factor = as.factor(out_group_minorities_assim),
out_group_majority_will = ifelse(q125b < 0, NA, q125b),</pre>
out_group_majority_will_factor = as.factor(out_group_majority_will),
out_group_immig_econ_good = ifelse(q125c < 0, NA, q125c),</pre>
out_group_immig_econ_good_factor = as.factor(out_group_immig_econ_good),
out_group_immig_culture_threat = ifelse(q125d < 0, NA, q125d),</pre>
out_group_immig_culture_threat_factor = as.factor(out_group_immig_culture_threat),
out_group_immig_crime = ifelse(q125e < 0, NA, q125e),</pre>
out_group_immig_crime_factor = as.factor(out_group_immig_crime),
scale pol lasceht = ifelse(q18a < 0, NA, q18a),</pre>
scale_pol_scholz = ifelse(q18b < 0, NA, q18b),</pre>
scale_pol_baerbock = ifelse(q18c < 0, NA, q18c))</pre>
```

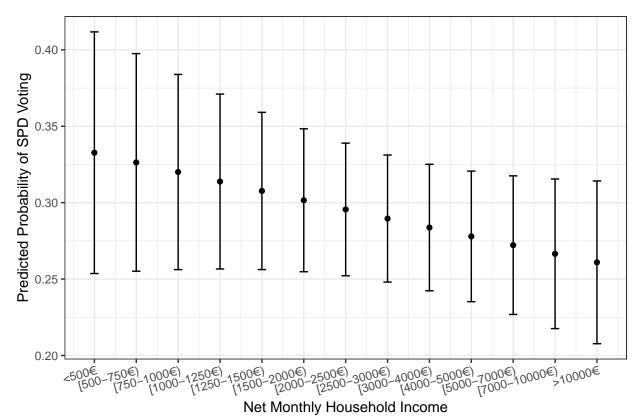
Some tentative analysis

SPD

Socio-Demographic Correlates

Relationship between household income and SPD voting

```
spd_income <- glm(spd_21 ~ household_income + age + abitur_factor + sex1 + urban_rural_factor + ostwest</pre>
# plot
cplot(spd_income, x = "household_income",
      xvals = seq(1, 13, 1), draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_continuous("Net Monthly Household Income",
                     breaks = seq(1, 13, 1),
                     labels = c("<500€", "[500-750€)",</pre>
                                "[750-1000€)", "[1000-1250€)",
                                "[1250-1500€)", "[1500-2000€)",
                                "[2000-2500€)", "[2500-3000€)",
                                "[3000-4000]", "[4000-5000]",
                                "[5000-7000€)", "[7000-10000€)",
                                ">10000€")) +
  labs(y = "Predicted Probability of SPD Voting",
       caption = "Covariates include: age, education, gender and rurality of place of residence.") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 15, hjust = 1))
```

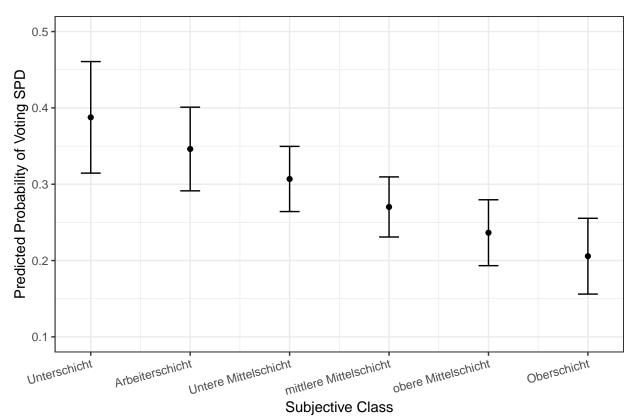


Covariates include: age, education, gender and rurality of place of residence.

There is no robust relationship between net monthly household income and voting for the SPD.

Relationship between subjective class and SPD voting

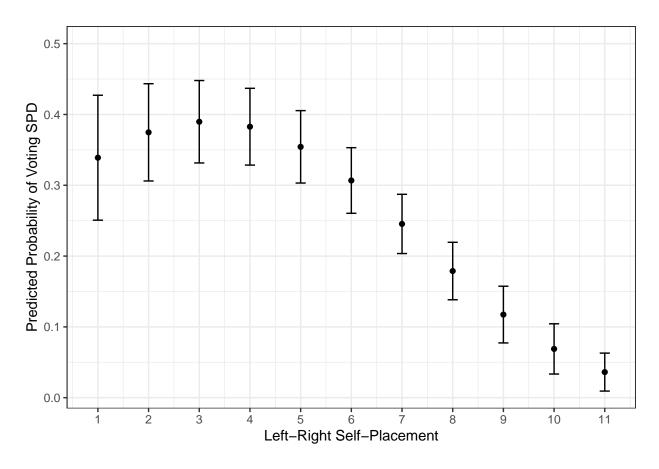
```
spd_sclass <- glm(spd_21 ~ subjective_class + age + abitur_factor + sex1 + urban_rural_factor + ostwest
# plot
cplot(spd_sclass, x = "subjective_class",
      xvals = seq(1, 6, 1), draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_continuous("Subjective Class",
                     breaks = seq(1, 6, 1),
                     labels = c("Unterschicht", "Arbeiterschicht",
                                "Untere Mittelschicht", "mittlere Mittelschicht",
                                "obere Mittelschicht", "Oberschicht")) +
  labs(y = "Predicted Probability of Voting SPD",
       caption = "Covariates include: age, education, gender and rurality of place of residence.") +
  ylim(c(0.1, 0.5)) +
  theme bw() +
  theme(axis.text.x = element_text(angle = 15, hjust = 1))
```



Covariates include: age, education, gender and rurality of place of residence.

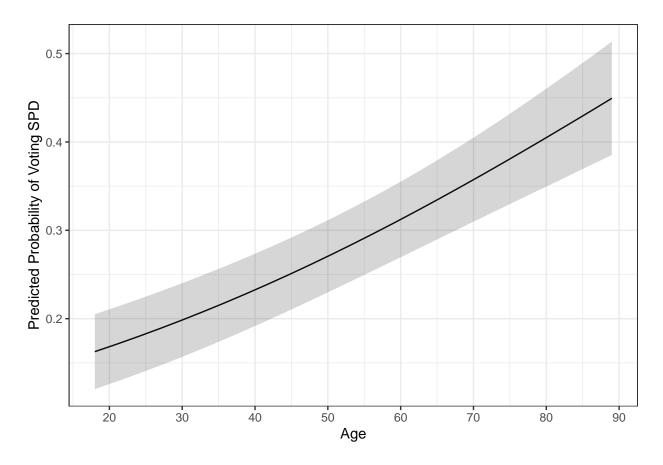
What is the relationship between left-right self-placement and SPD voting?

```
spd_left_right_self <- glm(spd_21 ~ left_right_self + I(left_right_self^2) + household_income + age + a
```



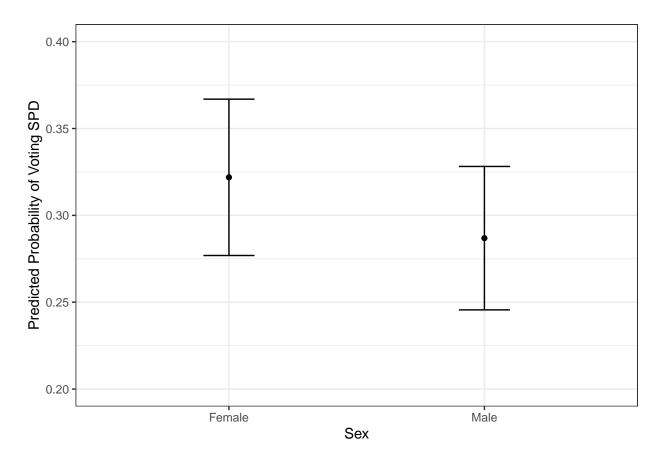
Relationship between age and SPD voting

```
spd_age <- glm(spd_21 ~ household_income + age + abitur_factor + sex1 + urban_rural_factor + ostwest_fa
# plot
cplot(spd_age, x = "age", draw = F) %>%
    as_tibble() %>%
    ggplot(aes(x = xvals)) +
    geom_line(aes(y = yvals)) +
    geom_ribbon(aes(ymin = lower, ymax = upper), alpha = 0.2) +
    scale_x_continuous("Age", breaks = seq(20, 90, 10)) +
    labs(y = "Predicted Probability of Voting SPD") +
    theme_bw()
```



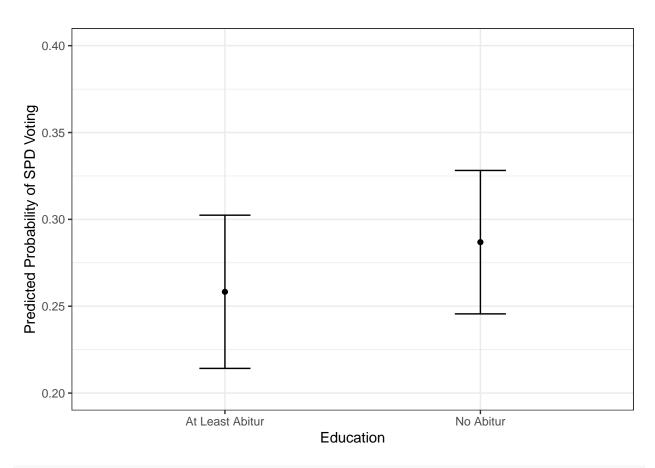
Relationship between sex and SPD voting

```
cplot(spd_income, x = "sex1", draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  ylim(c(0.2, 0.4)) +
  scale_x_discrete("Sex", labels = c("Female", "Male")) +
  labs(y = "Predicted Probability of Voting SPD") +
  theme_bw()
```

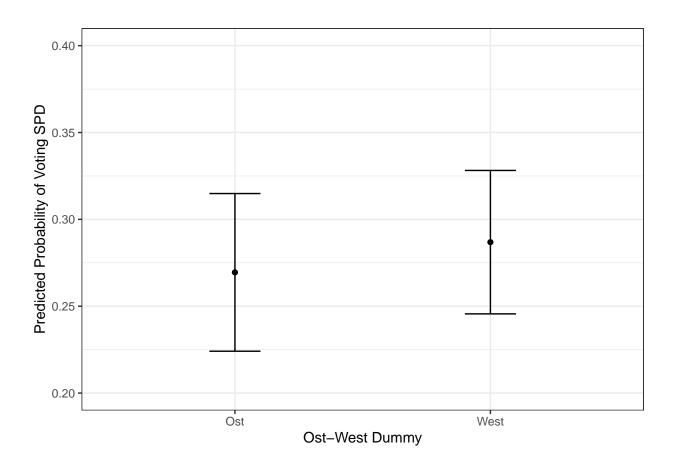


Relationship between education and SPD

```
cplot(spd_income, x = "abitur_factor", draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_discrete("Education", labels = c("At Least Abitur", "No Abitur")) +
  ylim(c(0.2, 0.4)) +
  labs(y = "Predicted Probability of SPD Voting") +
  theme_bw()
```

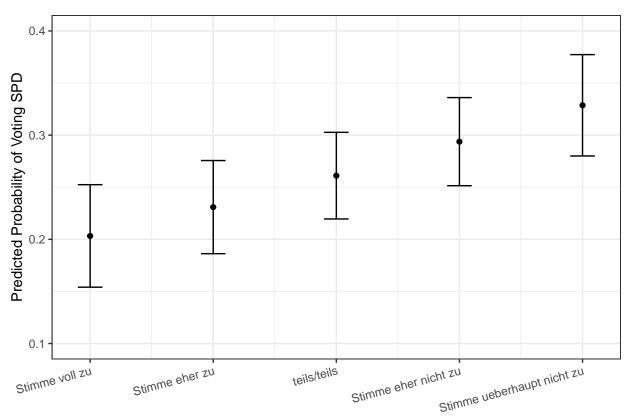


```
cplot(spd_income, x = "ostwest_factor", draw = F) %>%
    as_tibble() %>%
    ggplot(aes(x = xvals)) +
    geom_point(aes(y = yvals)) +
    geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
    scale_x_discrete("Ost-West Dummy", labels = c("Ost", "West")) +
    ylim(c(0.2, 0.4)) +
    labs(y = "Predicted Probability of Voting SPD") +
    theme_bw()
```



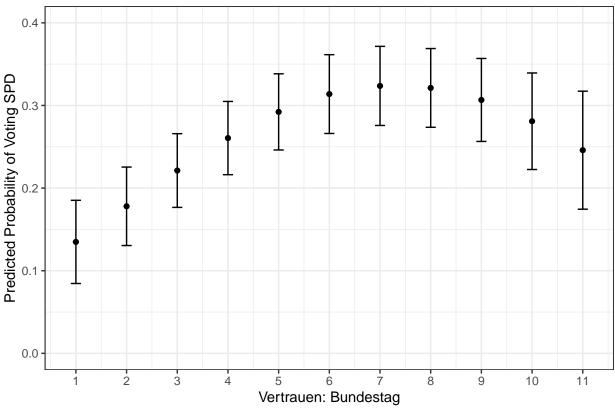
Attiudinal Correlates

```
# none of the other abgehaengt variables is significant
spd_cancel_culture <- glm(spd_21 ~ cancel_culture_subjektiv + household_income + age + abitur_factor +</pre>
# plot
cplot(spd_cancel_culture, x = "cancel_culture_subjektiv",
      xvals = seq(1, 5, 1), draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_continuous("Subjektiv: Keine freie Meinungsaeusserung moeglich",
                   breaks = seq(1, 5, 1),
                   labels = c("Stimme voll zu", "Stimme eher zu",
                              "teils/teils", "Stimme eher nicht zu",
                              "Stimme ueberhaupt nicht zu")) +
  labs(y = "Predicted Probability of Voting SPD") +
  ylim(c(0.1, 0.4)) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 15, hjust = 1))
```

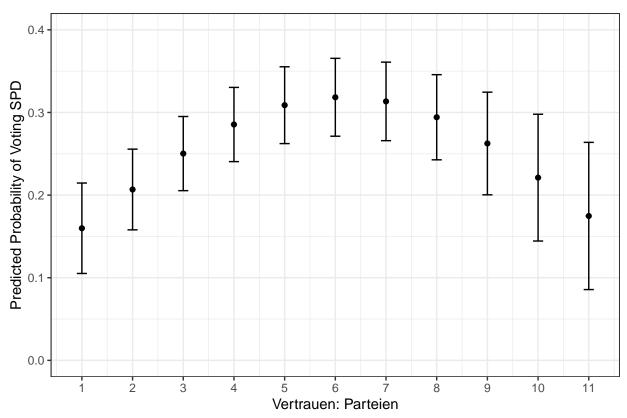


Subjektiv: Keine freie Meinungsaeusserung moeglich

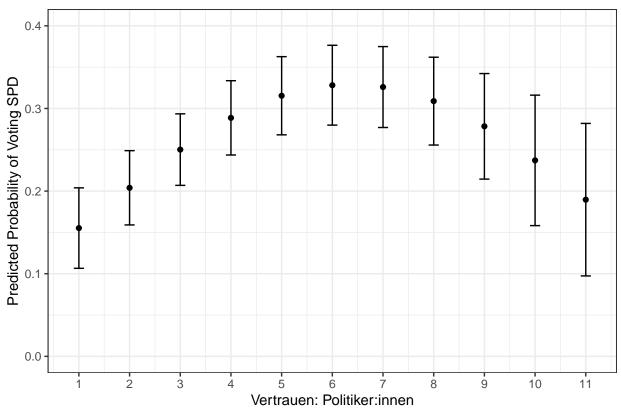
```
# general trust is not significant
# trust in parliament is significant
spd_trust_parliament <- glm(spd_21 ~ trust_in_parliament + I(trust_in_parliament^2) + household_income</pre>
# plot
cplot(spd_trust_parliament, x = "trust_in_parliament",
      xvals = seq(1, 11, 1), draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_continuous("Vertrauen: Bundestag",
                     breaks = seq(1, 11, 1)) +
 labs(y = "Predicted Probability of Voting SPD",
       caption = "'1' indicates 'no trust', while 11 indicates 'full trust'.") +
  ylim(c(0, 0.4)) +
  theme_bw()
```



'1' indicates 'no trust', while 11 indicates 'full trust'.

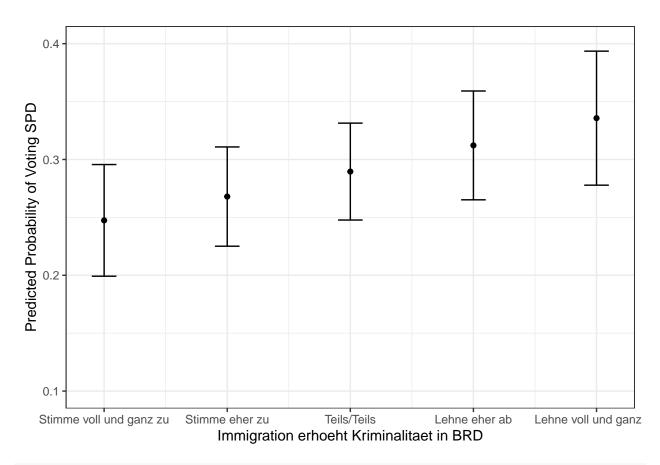


'1' indicates 'no trust', while 11 indicates 'full trust'.



'1' indicates 'no trust', while 11 indicates 'full trust'.

```
# immigrants bring crime is significant
spd_immig_crime <- glm(spd_21 ~ out_group_immig_crime + household_income + age + abitur_factor + sex1 +</pre>
# plot
cplot(spd_immig_crime, x = "out_group_immig_crime",
      xvals = seq(1, 5, 1), draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_continuous("Immigration erhoeht Kriminalitaet in BRD",
                     breaks = seq(1, 5, 1),
                     labels = c("Stimme voll und ganz zu", "Stimme eher zu",
                                "Teils/Teils", "Lehne eher ab",
                                "Lehne voll und ganz ab")) +
  labs(y = "Predicted Probability of Voting SPD") +
  ylim(c(0.1, 0.4)) +
 theme_bw()
```



```
# immigrants pose cultural threat is not significant at 5% level
# immigrants are good for economics is not significant at 5% level
# majority will is paramount is not significant
# outgroups should assimilate not significant
```

Gruene

Valence -> motivation: Baerbock's campaign

How can we measure valence vs spatial distance?

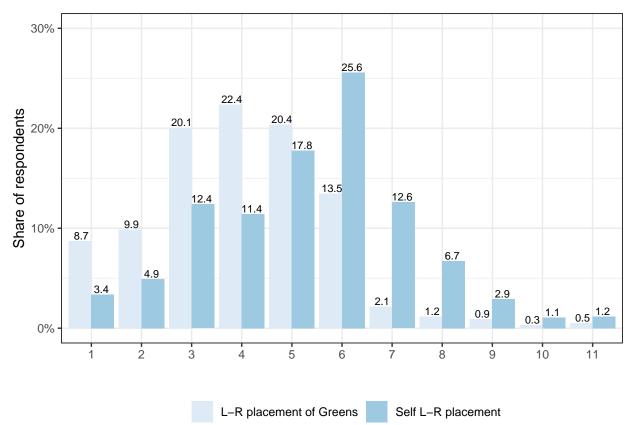
Who punished the Greens because of Baerbock? Who punished the CDU/CSU because of Laschet? (Lacher and so on) -> Those who are struggling / hard times.

-> egoistic vs sociotropic motivations/evaluations? -> egoistic evaluations matter more when one is ideologically closer to a candidate; spell this out -> sociotropic evaluations matter more when one thinks highly of a candidate

Spatial distance

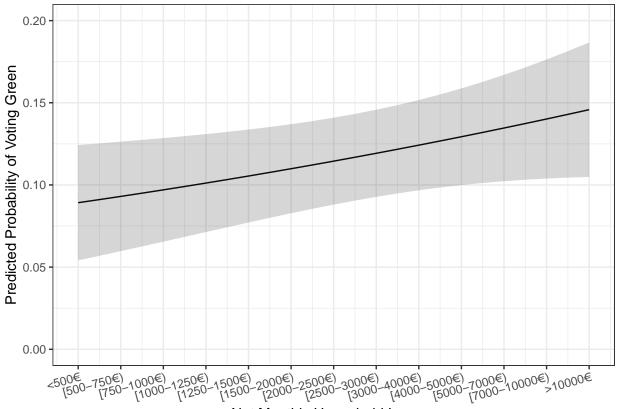
```
gruene_space <- glm(gruene_21 ~ distance_green + household_income + age + abitur_factor + sex1 + urban_s
summary(gruene_space)</pre>
```

```
##
## Call:
## glm(formula = gruene_21 ~ distance_green + household_income +
##
      age + abitur_factor + sex1 + urban_rural_factor + ostwest_factor,
##
      family = binomial(link = "logit"), data = gles_mod)
##
## Deviance Residuals:
      Min
                    Median
##
                10
                                 30
                                        Max
## -1.5342 -0.7083 -0.3949 -0.0132
                                      4.0312
##
## Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
                                   0.333388 0.490
## (Intercept)
                         0.163515
                                                     0.6238
## distance_green
                        -0.218754
                                  0.021620 -10.118 < 2e-16 ***
## household income
                         0.045328 0.026401
                                             1.717
                                                     0.0860 .
## age
                        ## abitur_factorno_abitur -0.750826   0.123852   -6.062 1.34e-09 ***
## sex1female
                         0.296052 0.115518 2.563
                                                     0.0104 *
                        -0.195854 0.182306 -1.074
## urban rural factor2
                                                     0.2827
## urban_rural_factor3
                        -0.326921 0.151903 -2.152
                                                     0.0314 *
## urban_rural_factor4
                        0.0149 *
## urban_rural_factor5
                         0.158309
                                   0.576306 0.275
                                                     0.7835
## ostwest_factorwest
                         0.596282
                                   0.138086
                                              4.318 1.57e-05 ***
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##
      Null deviance: 2334.8 on 2243 degrees of freedom
## Residual deviance: 1848.4 on 2233 degrees of freedom
     (1180 observations deleted due to missingness)
## ATC: 1870.4
##
## Number of Fisher Scoring iterations: 7
# placement of greens
gles_mod %>%
 select(left_right_green_factor, left_right_self_factor) %>%
 filter(!is.na(left_right_green_factor) & !is.na(left_right_self_factor)) %%
 pivot_longer(cols = everything(), names_to = "type", values_to = "value") %>%
 count(type, value) %>%
 group_by(type) %>%
 mutate(share = 100*(n/sum(n))) \%\%
 ggplot(aes(x = value, y = share, fill = type)) +
```

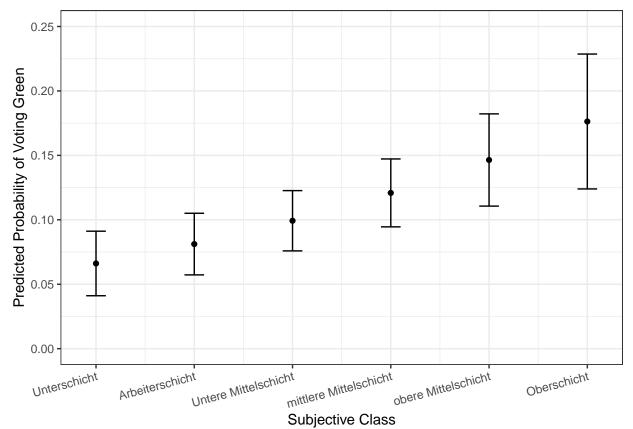


Socio-demographic Correlates

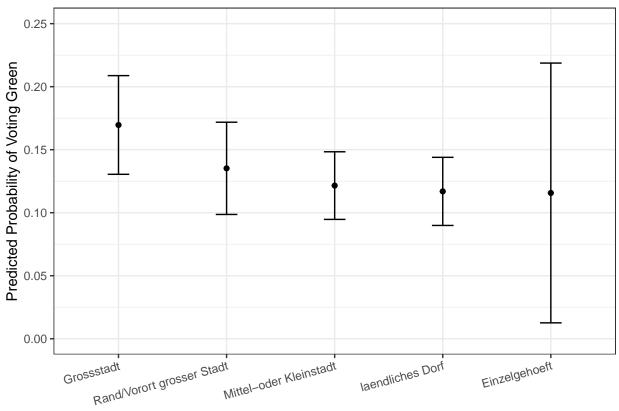
```
gruene_income <- glm(gruene_21 ~ household_income + age + abitur_factor + sex1 + urban_rural_factor + o
# plot
cplot(gruene_income, x = "household_income", draw = F) %>%
    as_tibble() %>%
    ggplot(aes(x = xvals)) +
    geom_line(aes(y = yvals)) +
    geom_ribbon(aes(ymin = lower, ymax = upper), alpha = 0.2) +
```



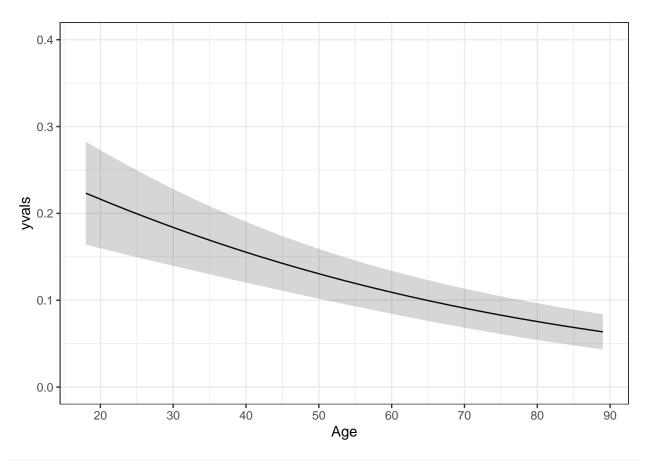
Net Monthly Household Income

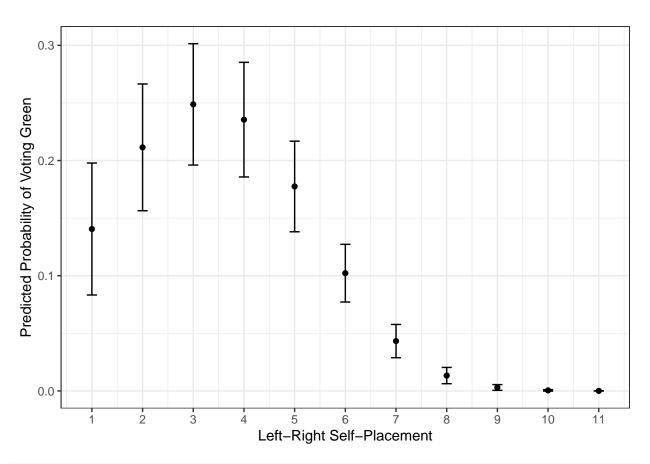


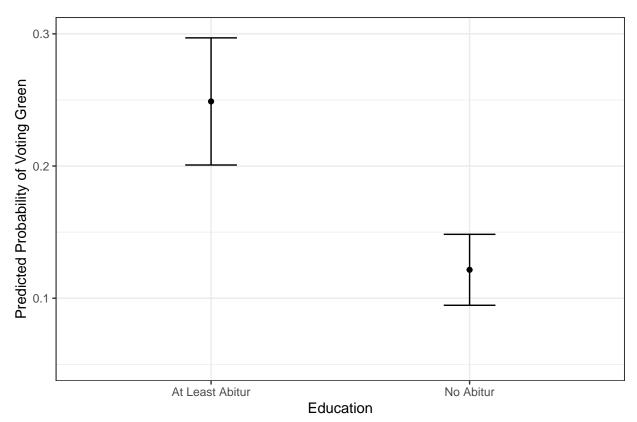
```
ylim(c(0, 0.25)) +
theme_bw() +
theme(axis.text.x = element_text(angle = 15, hjust = 1))
```



Urban-Rural

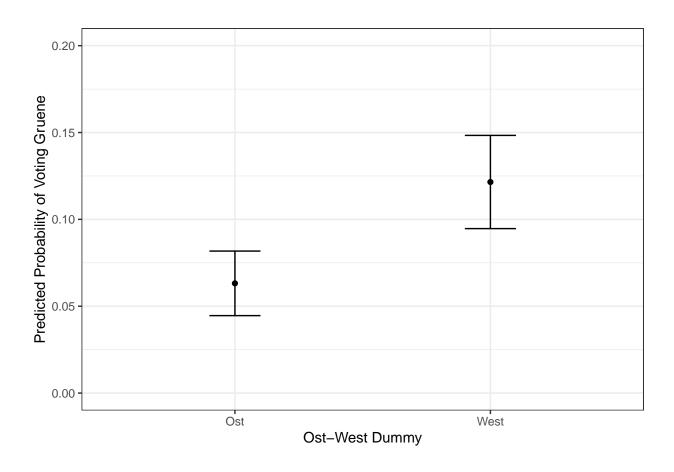






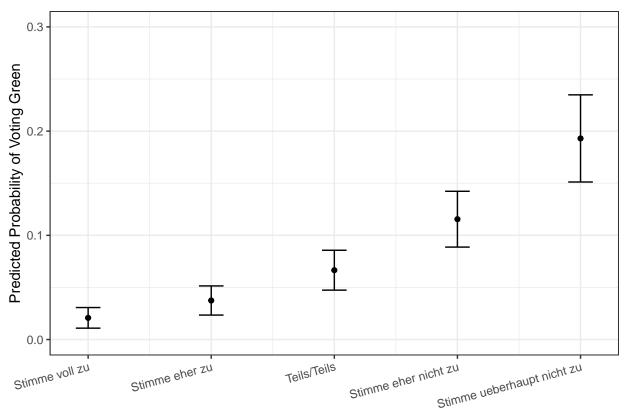
Covariates include: age, household income, sex, rurality of place of residence and an east-west dummy.

```
cplot(gruene_income, x = "ostwest_factor", draw = F) %>%
   as_tibble() %>%
   ggplot(aes(x = xvals)) +
   geom_point(aes(y = yvals)) +
   geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
   scale_x_discrete("Ost-West Dummy", labels = c("Ost", "West")) +
   ylim(c(0, 0.2)) +
   labs(y = "Predicted Probability of Voting Gruene") +
   theme_bw()
```

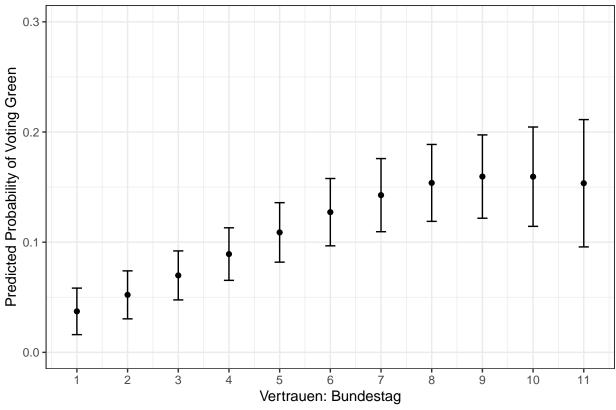


Attiudinal Correlates

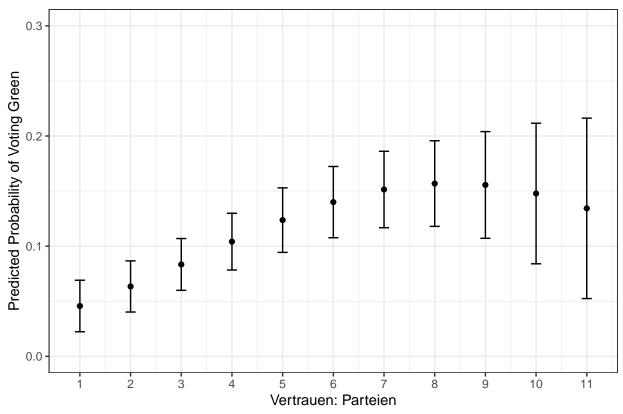
```
gruene_cancel <- glm(gruene_21 ~ cancel_culture_subjektiv + household_income + age + abitur_factor + se
# plot
cplot(gruene_cancel, x = "cancel_culture_subjektiv",
      xvals = seq(1, 5, 1), draw = F) %>%
 as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_continuous("Subjektiv: Keine freie Meinungsaeusserung moeglich",
                   breaks = seq(1, 5, 1),
                   labels = c("Stimme voll zu", "Stimme eher zu",
                              "Teils/Teils", "Stimme eher nicht zu",
                              "Stimme ueberhaupt nicht zu")) +
  labs(y = "Predicted Probability of Voting Green") +
  ylim(c(0, 0.3)) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 15, hjust = 1))
```



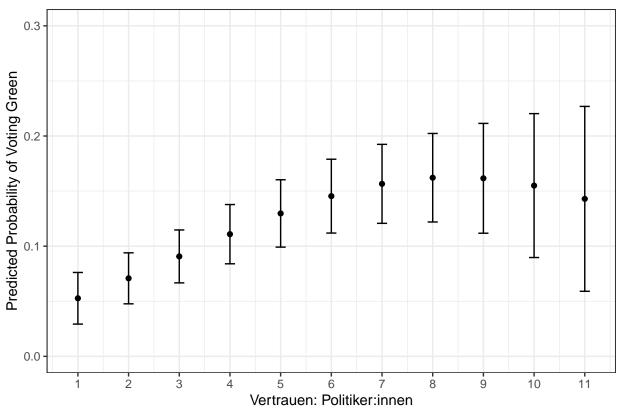
Subjektiv: Keine freie Meinungsaeusserung moeglich



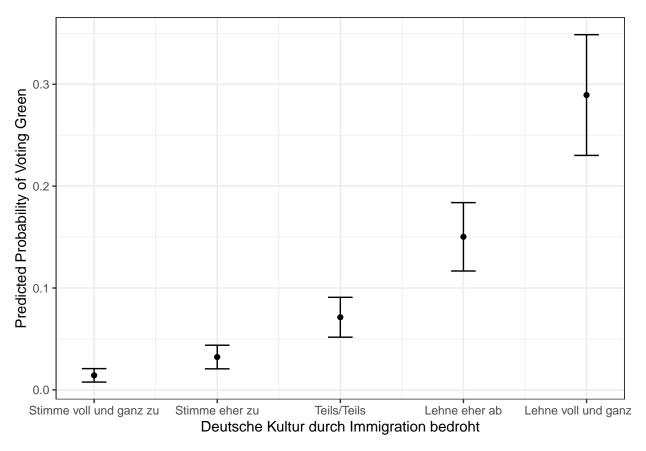
'1' indicates 'no trust', while 11 indicates 'full trust'.



'1' indicates 'no trust', while 11 indicates 'full trust'.



'1' indicates 'no trust', while 11 indicates 'full trust'.



```
gruene_immig_crime <- glm(gruene_21 ~ out_group_immig_crime + I(out_group_immig_crime^2) + household_in</pre>
# plot
cplot(gruene_immig_crime, x = "out_group_immig_crime",
      xvals = seq(1, 5, 1), draw = F) %>%
  as_tibble() %>%
  ggplot(aes(x = xvals)) +
  geom_point(aes(y = yvals)) +
  geom_errorbar(aes(ymin = lower, ymax = upper), width = 0.2) +
  scale_x_continuous("Immigration erhoeht Kriminalitaet in BRD",
                     breaks = seq(1, 5, 1),
                     labels = c("Stimme voll und ganz zu", "Stimme eher zu",
                                "Teils/Teils", "Lehne eher ab",
                                "Lehne voll und ganz ab")) +
  labs(y = "Predicted Probability of Voting Green") +
  ylim(c(0, 0.4)) +
  theme_bw()
```

