Replication to Journal of Politics: What Happens When Women Win Elections? The Electoral Returns to Increased Representation of Women

Frederik K. Kjøller Benjamin C.K. Egerod Lene H. Pedersen
Jens Olav Dahlgaard

2024-08-13

Setting up pathways

```
library(lfe); library(PanelMatch); library(purrr); library(tidyverse);
library(ggplot2); library(rdrobust); library(stargazer); library(haven);
library(DescTools); library(estimatr); library(broom); library(modelsummary)
library(lubridate); library(rdrobust); library(interflex); library(cowplot)
library(rddensity)

options(scipen = 999) # avoid scientific notation
work_data <- "e:/workdata/706687/"

knitr::opts_chunk$set(warning = FALSE, message = FALSE)

# create "not in"-operator
'%!in%' <- function(x,y)!('%in%'(x,y))
```

Data wrangling

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")


# standardize different spellings of the same municipalities
all_candidates <- all_candidates %>%
  mutate(KOMMUNE = case_when(KOMMUNE=="751 Århus"~"751 Aarhus",
                             KOMMUNE=="169 Høje Tåstrup"~"169 Høje-Taastrup",
                             TRUE~KOMMUNE))

#Estimate of personal votes
all_candidates$interval_points <- gsub(pattern =  "\\[", replacement = "",
                                       x = all_candidates$interval_points)
all_candidates$interval_points <- gsub(pattern =  "\\]", replacement = "",
                                       x = all_candidates$interval_points)

all_candidates <- all_candidates %>%
```

```
  separate(col = "interval_points", into = c("lwr", "upr"), sep = ",") %>%
  mutate(lwr = as.numeric(as.character(lwr)),
         upr = as.numeric(as.character(upr)),
         mid_pers_vote = (lwr+upr)/2)

# define vector of non-amalgated municipalities
non_amalgated_vector <- c(165, 201, 151, 153, 155, 563, 607, 147, 157, 159, 161, 253,
                          217, 163, 167, 169, 223, 183, 101, 673, 173, 825, 773, 727,
                          461, 329, 175, 741, 269, 185, 187, 751)


# define variable to sort amalgated municipalities away as outcome year in 2005
all_candidates <- all_candidates %>%
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  #make dummy equaling zero for amalgated municipalities in 2005 election but otherwise 1
  mutate(muni_to_include = ifelse(kom_nr %!in% non_amalgated_vector & election_year==2001,0,1))


av_vote_tot <- all_candidates %>%
  filter(election_year>1993) %>%
  filter(fractile_thres <= 18) %>%
  filter(muni_to_include==1) %>%
  group_by(KOMMUNE, election_year) %>%
  summarise(vote_tot = sum(mid_pers_vote, na.rm=T))

mean(av_vote_tot$vote_tot)
```

```
## [1] 2831.702
```

```
## find share of open-lists

clusters_agg <- all_candidates %>%
  filter(election_year>1993) %>%
  mutate(open = ifelse(OPSTILLINGSFORM=="Sideordnet",1,0)) %>%
  group_by(party, KOMMUNE, election_year, open, OPSTILLINGSFORM) %>%
  summarise(n = n())

mean(clusters_agg$open)
```

```
## [1] 0.7179922
```

```
# proportion of women in election-party-municipality groups
party_agg <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(female_candidates = mean(female, na.rm = T),
            party_vote = mean(total_votes, na.rm = T))

party_agg2 <- all_candidates %>%
  filter(elected == 1) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(prop_female = mean(female, na.rm = T),
            n_elect = n())


party_agg <- left_join(party_agg, party_agg2)
```

```r
# votes for women
fem_vote <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  filter(female == 1) %>%
  summarise(fem_vote = sum(mid_pers_vote, na.rm = T))

# votes for men
male_vote <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  filter(female == 0) %>%
  summarise(male_vote = sum(mid_pers_vote, na.rm = T))


party_agg <- left_join(party_agg, fem_vote)
party_agg <- left_join(party_agg, male_vote)

# create covariates from individual-level data
covariates <- all_candidates %>%
  group_by(election_year) %>%
  mutate(income_wins = Winsorize(income, probs = c(0.01,0.99), na.rm = T)) %>%
  ungroup() %>%
  mutate(immig = case_when(ethnicity %in% c(2:3) ~ 1,
                           ethnicity == 1 ~ 0,
                           TRUE ~ NA_real_),
         college = case_when(education %in% c("Bachelor",
                                              "Forskeruddannelser",
                                              "Mellemlange videregående uddannelser",
                                              "Lange videregående uddannelser") ~ 1,
                             TRUE ~ 0),
         vocational = case_when(education %in% c("Erhvervsfaglige praktik- og hovedforløb",
                                                 "Erhvervsgymnasiale uddannelser") ~ 1,
                                TRUE ~ 0),
         neg_income = case_when(income < 0 ~ 1,
                                TRUE ~ income),
         out_work = case_when(unemp %in% c(210, 321, 330) ~ 1,
                              TRUE ~ 0)) %>%

  group_by(party, KOMMUNE, election_year) %>%
  summarise(prop_immig = mean(immig, na.rm = T),
            prop_college = mean(college, na.rm = T),
            prop_vocational = mean(vocational, na.rm = T),
    av_age = mean(age, na.rm = T),
    prop_unemp = mean(out_work, na.rm = T),
    av_income = mean(income, na.rm = T),
    av_income_wins = mean(income_wins, na.rm=TRUE),
    av_competence = mean(inc_res, na.rm = T),
    prop_neg_income = mean(neg_income, na.rm = T))

party_agg <- left_join(party_agg, covariates)

# lags and leads

party_agg <- party_agg %>%
```

```
  mutate(fem_vote = case_when(is.na(fem_vote) == T ~ 0,
                              TRUE ~ fem_vote),
         male_vote = case_when(is.na(male_vote) == T ~ 0,
                               TRUE ~ male_vote)) %>%
  # change naming of three non-amalgated municipalities with changed spelling from Å to AA after the re
  mutate(KOMMUNE = ifelse(KOMMUNE=="751 Århus", "751 Aarhus", KOMMUNE)) %>%
  mutate(KOMMUNE = ifelse(KOMMUNE=="169 Høje Tåstrup", "169 Høje-Taastrup", KOMMUNE)) %>%
  mutate(KOMMUNE = ifelse(KOMMUNE=="173 Lyngby-Tårbæk", "173 Lyngby-Taarbæk", KOMMUNE)) %>%
group_by(party, KOMMUNE) %>%
  mutate(prop_female_t1 = dplyr::lead(prop_female, n = 1, order_by = election_year),
         prop_female_t2 = dplyr::lead(prop_female, n = 2, order_by = election_year),
         fem_vote_t1 = dplyr::lead(fem_vote, n = 1, order_by = election_year),
         male_vote_t1 = dplyr::lead(male_vote, n = 1, order_by = election_year),
         party_vote_t1 = dplyr::lead(party_vote, n = 1, order_by = election_year),
         n_elect_t1 = dplyr::lead(n_elect, n = 1, order_by = election_year),
         female_candidates_t1 = dplyr::lead(female_candidates, n = 1, order_by = election_year)) %>%
  ### leading creates NAs for parties that did not achieve representation in the next election   ##assi
  mutate(prop_female_lag = dplyr::lag(prop_female, 1, order_by = election_year),
         prop_immig_lag = dplyr::lag(prop_immig, 1, order_by = election_year),
         prop_college_lag = dplyr::lag(prop_college, 1, order_by = election_year),
         prop_vocational_lag = dplyr::lag(prop_vocational, 1, order_by = election_year),
         av_age_lag = dplyr::lag(av_age, 1, order_by = election_year),
         prop_unemp_lag = dplyr::lag(prop_unemp, 1, order_by = election_year),
         av_income_lag = dplyr::lag(av_income, 1, order_by = election_year),
         av_competence_lag = dplyr::lag(av_competence, 1, order_by = election_year),
         av_income_wins_lag = dplyr::lag(av_income_wins, 1, order_by = election_year),
         prop_neg_income_lag = dplyr::lag(prop_neg_income, 1, order_by = election_year)) %>%
  arrange(party, KOMMUNE, election_year)

##########################################
# this party_agg dataset is the main one! #
# export party aggregated data            #
saveRDS(party_agg, file =
          paste0(work_data, "projects/women_spearhead/data/party_agg_w_outcomes.rds"))
##########################################



##########################################
# finding treatment and control group     #
# parties where the marginal seat is between a man and woman

# create and get top two candidates within election-municipality-party
top2_cand_all <- all_candidates %>%
  filter(OPSTILLINGSFORM == "Sideordnet") %>%
  group_by(party, KOMMUNE, election_year) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  arrange(party, KOMMUNE, election_year)

# create indicators of female win -- treatment
fem_win <- top2_cand_all %>%
  mutate(female_win = case_when(female == 1 & elected == 1 ~ 1,
```

```
                                TRUE ~ 0)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(female_win = mean(female_win, na.rm = T),
            female_participant = mean(female, na.rm = T)) %>%
  mutate(female_win = case_when(female_win > 0 ~ 1,
                                TRUE ~ 0)) %>%
  # filter to exclude man/man mash-ups
  filter(female_participant == 0.5)

# define data set of parties with man/women competition of the marginal seat
top2_cand <- left_join(fem_win, top2_cand_all, by = c("party", "KOMMUNE", "election_year"))

# summarise personal votes and electoral closeness variable and add them to dataset
pers_vote <-top2_cand %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(mid_pers_vote = mean(mid_pers_vote, na.rm=T))

top2_cand <- top2_cand %>%
  group_by(party, KOMMUNE, election_year, female_win) %>%
  summarise(fractile_thres = mean(fractile_thres, na.rm=T))

top2_cand <- left_join(top2_cand, party_agg,
                       by = c("party", "KOMMUNE", "election_year"))


top2_cand <- left_join(top2_cand, pers_vote,
                       by = c("party", "KOMMUNE", "election_year"))

# create id for each party-municipality observation
top2_cand$cluster_id <- top2_cand %>%
  group_by(party, KOMMUNE) %>%
  group_indices()


top2_cand$muni_id <- top2_cand %>%
  group_by(KOMMUNE) %>%
  group_indices()


### remove amalgated municipalities in year 2005

# define variable to sort amalgated municipalities away as outcome year in 2005
top2_cand <- top2_cand %>%
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  #make dummy equaling zero for amalgated municipalities in 2005 election but otherwise 1
  mutate(muni_to_include = ifelse(kom_nr %!in% non_amalgated_vector & election_year==2001,0,1))


### leading n_elect and female_win creates NAs for parties that did not achieve
# representation in the next election these must be assigned to 0
#BUT parties that did not run at all in the next election also gets NA.
top2_cand <- top2_cand %>%
  # find the few parties that rerun but don't get any seats
```

```r
  mutate(n_elect_t1 = ifelse(is.na(n_elect_t1) & !is.na(party_vote_t1),0,n_elect_t1),
         prop_female_t1 = ifelse(is.na(prop_female_t1) & !is.na(party_vote_t1),0,prop_female_t1))


# create unique id variable for year-municipality-party
top2_cand <- top2_cand %>%
  mutate(cluster = str_c(election_year, kom_nr, party, sep = "-", collapse = NULL))

### save dataset for all BWs
top2_cand <- top2_cand %>%
  filter(muni_to_include==1 & election_year>1993 & !is.na(prop_female_t1))

##############################################################################
# top2_cand is the main treatment dataset #
# export the data set                     #
saveRDS(top2_cand, file =
          paste0(work_data, "projects/women_spearhead/data/top2_cand_data.rds"))
##############################################################################
```

Table 1

```r
fem_elect <- felm(prop_female_t1 ~ female_win |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 18))

fem_vote <- felm(log(fem_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                 data = filter(top2_cand, fractile_thres <= 18))

male_vote <- felm(log(male_vote_t1) ~ female_win |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 18))

party_vote_mod <- felm(log(party_vote_t1) ~ female_win |0|0|KOMMUNE + party,
                       data = filter(top2_cand, fractile_thres <= 18))

mandates_mod <- felm(n_elect_t1 ~ female_win |0|0|KOMMUNE + party,
                     data = filter(top2_cand, fractile_thres <= 18))


stargazer(fem_elect, fem_vote, male_vote, party_vote_mod, mandates_mod,
          covariate.labels = c("Woman win t=0"),
          dep.var.labels = c(" Share Women Elected", "Log Votes Women", "Log Votes Men",
                             "Log Votes Party", "Seats Party"),
          add.lines = list(c("Bandwidth",
                             rep(18, 5))),
          font.size = "scriptsize",
          type = "text",
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/electoral_success_optimal_bw.tex",
          title = "The Effect of Electing Women on Future Electoral Outcomes (t+1)",
          label = "tab:electoral_rewards")
```

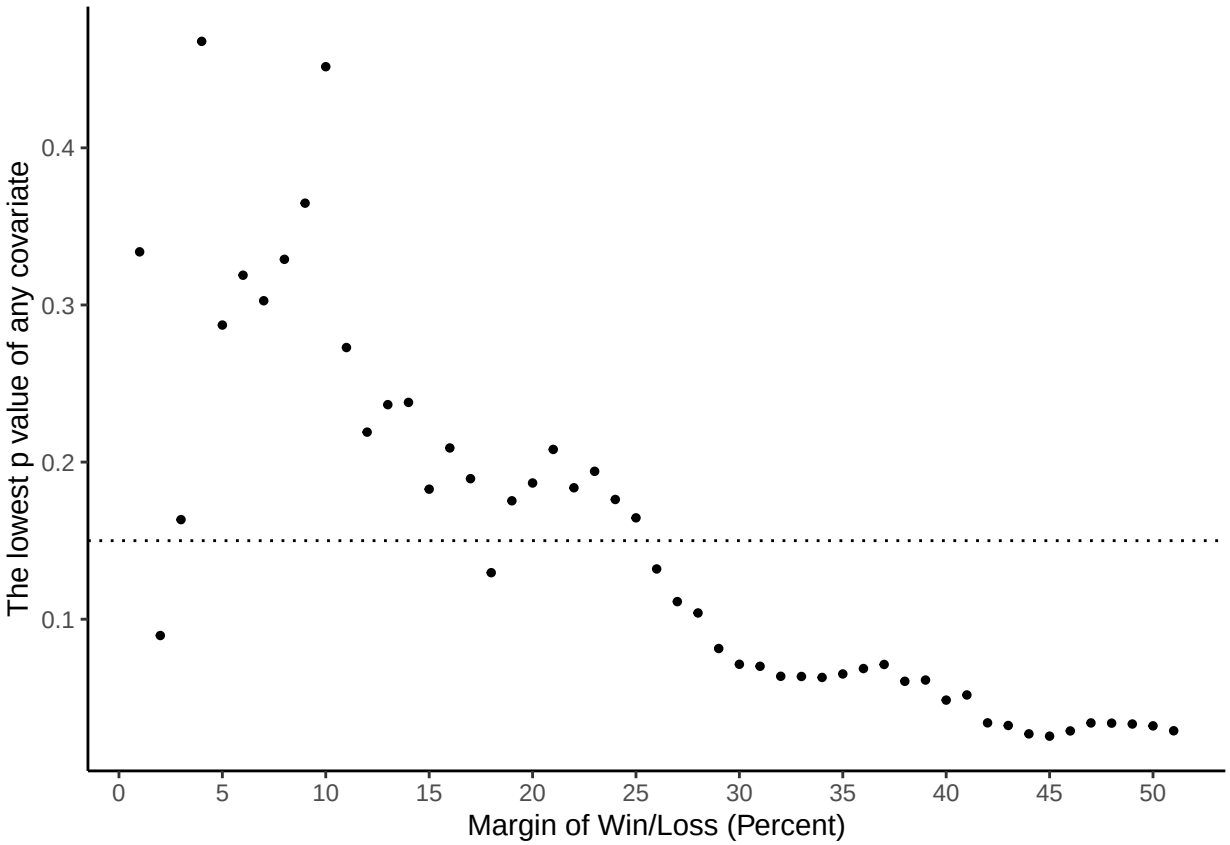
```
##
## The Effect of Electing Women on Future Electoral Outcomes (t+1)
## ==========================================================================================
##                                          Dependent variable:
##                 ------------------------------------------------------------------------
##                 Share Women Elected Log Votes Women Log Votes Men Log Votes Party Seats Party
##                         (1)               (2)            (3)            (4)           (5)
## ------------------------------------------------------------------------------------------
## Woman win t=0       0.107***          0.428***         0.031         0.134**       0.311**
##                      (0.034)           (0.076)        (0.073)        (0.058)       (0.145)
##
## Constant            0.238***          5.793***       7.323***       7.490***      4.562***
##                      (0.019)           (0.250)        (0.264)        (0.210)       (1.022)
##
## ------------------------------------------------------------------------------------------
## Bandwidth              18                18             18             18            18
## Observations           819               819            819            819           819
## ==========================================================================================
## Note:                                                          *p<0.1; **p<0.05; ***p<0.01
```

Appendix B

Figure B1

```
# load data
top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

party_agg <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/party_agg_w_outcomes.rds")

all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")


#########################################################
####### PERMUTATION TEST OF COVARIATE IMBALANCE USING LOWEST P VALUE

# create covariates and lag them
muni_vote <- party_agg %>%
  group_by(KOMMUNE, election_year) %>%
  summarise(muni_vote = sum(fem_vote+male_vote, na.rm = T),
            muni_seat = sum(n_elect, na.rm=T)) %>%
  right_join(party_agg)

muni_vote <- muni_vote %>%
  group_by(party, election_year) %>%
  arrange(election_year) %>%
  mutate(muni_vote_lag = dplyr::lag(muni_vote, n = 1, order_by = election_year),
         fem_vote_lag = dplyr::lag(fem_vote, n = 1, order_by = election_year),
         male_vote_lag = dplyr::lag(male_vote, n = 1, order_by = election_year),
         party_vote_lag = dplyr::lag(party_vote, n = 1, order_by = election_year),
         n_elect = ifelse(is.na(n_elect) == T, 0, n_elect),
         party_vote_share_lag = (fem_vote_lag+male_vote_lag)/muni_vote_lag,
         fem_vote_share = fem_vote/muni_vote_lag,
         male_vote_share = male_vote/muni_vote_lag,
         n_elect_lag = lag(n_elect, 1, order_by = election_year),
         female_candidates_lag = lag(female_candidates, 1, order_by = election_year),
         seat_share_lag = n_elect_lag/muni_seat) %>%
  mutate(national_party = case_when(party %in% c("A", "B", "V", "C", "F", "I", "O", "Ø", "Å") ~ 1,
                                    TRUE ~ 0),
         left_party = case_when(party %in% c("A", "B", "F", "Ø") ~ 1,
                                TRUE ~ 0)) %>%
  select(KOMMUNE, party, election_year,
         fem_vote_lag, seat_share_lag, party_vote_share_lag,
         male_vote_lag, female_candidates_lag,
         party_vote_lag, national_party, left_party, n_elect_lag)

# join back with main dataset
top2_cand <- left_join(top2_cand, muni_vote, by = c("KOMMUNE", "party", "election_year"))

# variable capturing the biggest cities
top2_cand <- top2_cand %>%
  mutate(big_city = ifelse(kom_nr %in% c(101,147,751,461,851),1,0))
```

```r
# function to iterate over bandwidths
balance_fstat_func <- function(bw){

  #OLS with cluster robust SEs subset within bandwidth BW
  fit <- lm_robust(female_win~
                   prop_college_lag +
                   av_age_lag +
                   prop_unemp_lag +
                   prop_immig_lag +
                   log(female_candidates_lag + 1) +
                   national_party +
                   party_vote_share_lag +
                   left_party +
                   big_city +
                   av_competence_lag,
                   clusters = cluster_id,
                   data = top2_cand, se_type = "CR2",
                   subset = fractile_thres <= bw)

  # extract minimum p value in each iteration
  fit_sum <- data.frame(tidy(fit) %>% select(term, p.value) %>%
                          filter(term != "(Intercept)") %>%
                          filter(p.value == min(p.value,na.rm=T)),
                        glance(fit) %>% select(fstat = p.value),
                        iter = bw)

  ## extract estimates for balance plot

  all_sum <- data.frame(tidy(fit) %>%
                          select(term, estimate, std.error) %>% filter(term != "(Intercept)"),
                        iter = bw)

  return(fit_sum)
}

# map over all bandwidths
fstat_bw_df <- map(seq(1, 51.1, 1),
                   ~ balance_fstat_func(bw = .))

#combine into data.frame
bw_fstats <- do.call("rbind", fstat_bw_df)

# plot
bw_fstats %>%
  ggplot(data = .,
         aes(x = iter, y = p.value)) +
  geom_point(size = 1) +
  scale_x_continuous(breaks = seq(0,50,5), labels = seq(0,50,5)) +
  scale_y_continuous(breaks = seq(0,0.8,0.1), labels = seq(0,0.8,0.1)) +
  theme_classic() +
  geom_hline(yintercept = 0.15, lty = 3) +
  labs(x = "Margin of Win/Loss (Percent)",
       y = "The lowest p value of any covariate")
```

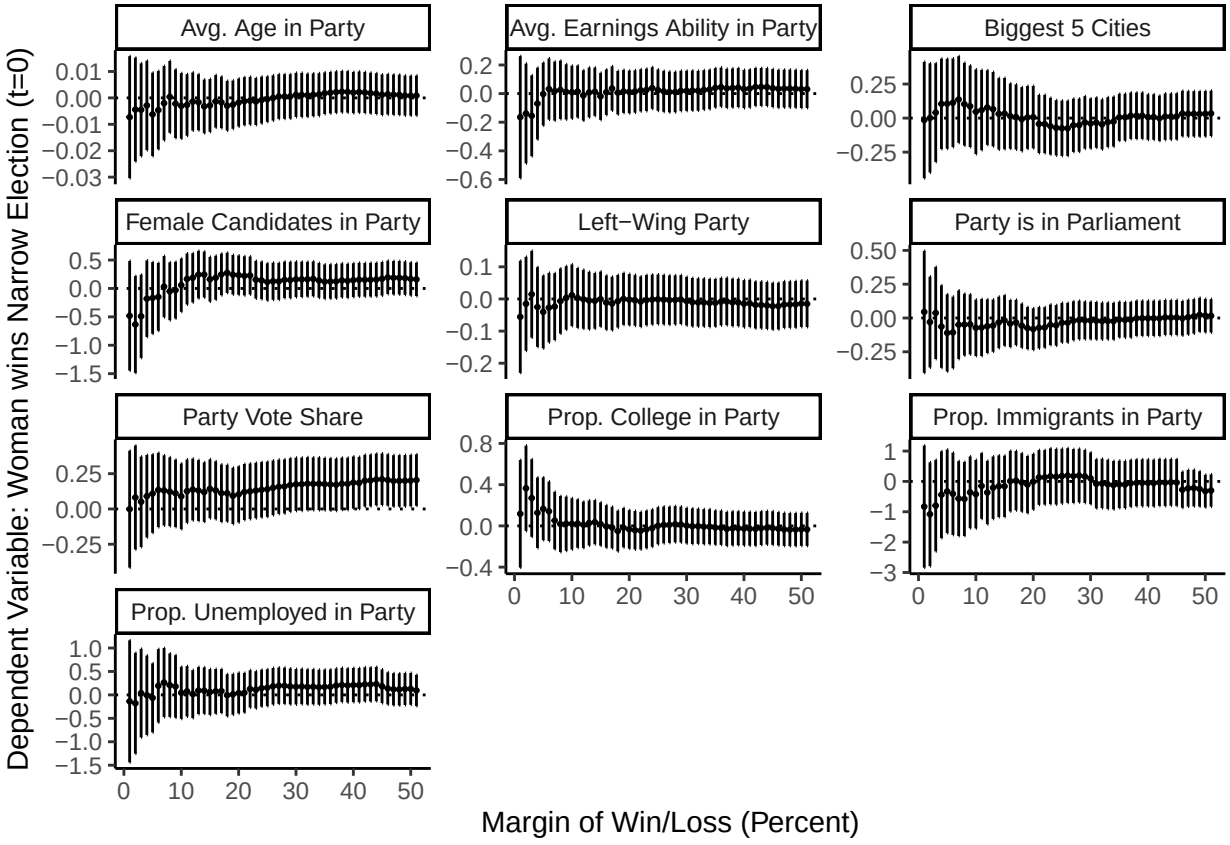

The lowest p value of any covariate
0.4
0.3
0.2
0.1
0
5
10
15
20
25
30
35
40
45
50
Margin of Win/Loss (Percent)

Appendix C

Figure C1

```
#repeat above but extract estimates for each covariate

balance_all_covar_func <- function(bw){
  fit <- lm_robust(female_win~
                   prop_college_lag +
                   av_age_lag +
                   prop_unemp_lag +
                   prop_immig_lag +
                   log(female_candidates_lag + 1) +
                   national_party +
                   party_vote_share_lag +
                   left_party +
                   big_city +
                   av_competence_lag,
                   clusters = cluster_id,
                   data = top2_cand, se_type = "CR2",
                   subset = fractile_thres <= bw)


  all_sum <- data.frame(tidy(fit) %>% select(term, estimate, std.error) %>% filter(term != "(Intercept)"
                        iter = bw)

  return(all_sum)
}


all_bw <- map(seq(1, 51.1, 1),
                   ~ balance_all_covar_func(bw = .))

all_bw_df <- do.call("rbind", all_bw)

# label covariates
all_bw_df <- all_bw_df %>%
  mutate(clean_term = case_when(term == "prop_college_lag" ~ "Prop. College in Party",
                                term == "prop_unemp_lag" ~ "Prop. Unemployed in Party",
                                term == "log(female_candidates_lag + 1)" ~ "Female Candidates in Party"
                                term == "party_vote_share_lag" ~ "Party Vote Share",
                                term == "big_city" ~ "Biggest 5 Cities",
                                term == "av_age_lag" ~ "Avg. Age in Party",
                                term == "prop_immig_lag" ~ "Prop. Immigrants in Party",
                                term == "national_party" ~ "Party is in Parliament",
                                term == "left_party" ~ "Left-Wing Party",
                                term == "av_competence_lag" ~ "Avg. Earnings Ability in Party"))

all_bw_df <- all_bw_df %>%
  mutate(optimal_bw = case_when(iter==18.01~"Selected BW",
                                TRUE~"BW")) %>%
  mutate(optimal_bw = factor(optimal_bw, levels = c("Selected BW",
                                                    "BW")))
```

```r
# plot results
p_balance<- ggplot(all_bw_df, aes(x = iter, y = estimate, color = optimal_bw)) +
  geom_point(size = 0.5) +
  geom_errorbar(aes(ymin = estimate - 1.96*std.error,
                    ymax = estimate + 1.96*std.error), width = 0) +
  facet_wrap(~clean_term, scales = "free_y", ncol = 3) +
  theme_classic() +
  geom_hline(yintercept = 0, lty = 3) +
  scale_color_manual("", values = c("black", "gray75")) +
  theme(legend.position = "none")+
  labs(x = "Margin of Win/Loss (Percent)",
       y = "Dependent Variable: Woman wins Narrow Election (t=0)")
p_balance
```



Appendix D

Table D1

```
# first stage
pool_first <- felm(prop_female ~ female_win | 0|0|KOMMUNE + party,
                   data =  top2_cand, subset = fractile_thres<=18)


stargazer(pool_first,
          covariate.labels = c("Woman Candidate Win t=0"),
          dep.var.labels = c("Share Women Elected t=0"),
          add.lines = list(c("Bandwidth", rep(18,1))),
          df = FALSE, keep.stat = "n",
          label = "tab:first_stage",
          title = "Testing the validity of the treatment. The effect of electing an additional woman on
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:28:50

Table 1: Testing the validity of the treatment. The effect of electing an additional woman on share of elected women

	Dependent variable:
	Share Women Elected t=0
Woman Candidate Win t=0	0.338***
	(0.083)
Constant	0.163***
	(0.028)
Bandwidth	18
Observations	819
Note:	*p<0.1; **p<0.05; ***p<0.01

Table D2

```
# estimate 2nd stage models
iv1 <- felm(prop_female_t1 ~ 0 |0|(prop_female ~ female_win)|KOMMUNE + party,
                 data = top2_cand, subset = fractile_thres<=18)

iv2 <- felm(log(fem_vote_t1+1) ~ 0 |0|(prop_female ~ female_win)|KOMMUNE + party,
            data = top2_cand, subset = fractile_thres<=18)

iv3 <- felm(log(male_vote_t1+1)~ 0 |0|(prop_female ~ female_win)|KOMMUNE + party,
            data = top2_cand, subset = fractile_thres<=18)

iv4 <- felm(log(party_vote_t1+1)~ 0 |0|(prop_female ~ female_win)|KOMMUNE + party,
            data = top2_cand, subset = fractile_thres<=18)

iv5 <- felm(n_elect_t1~ 0 |0|(prop_female ~ female_win)|KOMMUNE + party,
            data = top2_cand, subset = fractile_thres<=18)

# all models
stargazer(iv1, iv2, iv3, iv4, iv5, font.size = "footnotesize",
          covariate.labels = c("(IV'd) Share Women Elected"),
          dep.var.labels = c(" Share Women Elected", "Log Votes Women", "Log Votes Men",
                             "Log Votes Party", "Seats Party"),
          add.lines = list( c("Bandwidth", rep(18,5))),
          df = FALSE, keep.stat = "n",
          label = "tab:first_stage_reduced_iv",
          title = "The LATE of electing women on future party performance")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 17:54:33

Table 2: The LATE of electing women on future party performance

	Dependent variable:				
	Share Women Elected	Log Votes Women	Log Votes Men	Log Votes Party	Seats Party
	(1)	(2)	(3)	(4)	(5)
(IV'd) Share Women Elected	0.687***	12.399***	14.663***	15.196***	9.712***
	(0.064)	(1.764)	(2.242)	(2.143)	(3.013)
Bandwidth	18	18	18	18	18
Observations	819	819	819	819	819
Note:					*p<0.1; **p<0.05; ***p<0.01

Appendix E

Tabel E1

```
### finding treatment and control group: parties where the marginal seat is between a man and woman

# create and get top two candidates within election-municipality-party
top2_cand_all_semi <- all_candidates %>%
  filter(OPSTILLINGSFORM != "Sideordnet") %>%
  group_by(party, KOMMUNE, election_year) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  arrange(party, KOMMUNE, election_year)

# create indicators of female win
fem_win <- top2_cand_all_semi %>%
  mutate(female_win = case_when(female == 1 & elected == 1 ~ 1,
                                TRUE ~ 0)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(female_win = mean(female_win, na.rm = T),
            female_participant = mean(female, na.rm = T)) %>%
  mutate(female_win = case_when(female_win > 0 ~ 1,
                                TRUE ~ 0)) %>%
  # filter to exclude man/man mash-ups
  filter(female_participant == 0.5)

# define data set of parties with man/women competition of the marginal seat
top2_cand_semi <- left_join(fem_win, top2_cand_all_semi, by = c("party", "KOMMUNE", "election_year"))

top2_cand_semi <- top2_cand_semi %>%
  group_by(party, KOMMUNE, election_year, female_win) %>%
  summarise(fractile_thres = mean(fractile_thres, na.rm=T))

top2_cand_semi <- left_join(top2_cand_semi, party_agg,
                        by = c("party", "KOMMUNE", "election_year"))

top2_cand_semi <- top2_cand_semi %>%
  mutate(n_elect_t1 = ifelse(is.na(n_elect_t1) & !is.na(party_vote_t1),0,n_elect_t1),
         prop_female_t1 = ifelse(is.na(prop_female_t1) & !is.na(party_vote_t1),0,prop_female_t1))


######## estimate models
fem_elect <- felm(prop_female_t1 ~ female_win |0|0|KOMMUNE + party,
                  data = filter(top2_cand_semi, fractile_thres <= 18))

fem_vote <- felm(log(fem_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                 data = filter(top2_cand_semi, fractile_thres <= 18))

male_vote <- felm(log(male_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                  data = filter(top2_cand_semi, fractile_thres <= 18))

party_vote_mod <- felm(log(party_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                       data = filter(top2_cand_semi, fractile_thres <= 18))
```

```r
mandates_mod <- felm(n_elect_t1 ~ female_win |0|0|KOMMUNE + party,
                     data = filter(top2_cand_semi, fractile_thres <= 18))

stargazer(fem_elect, fem_vote, male_vote, party_vote_mod, mandates_mod,
          keep.stat = "n",
          covariate.labels = c("Woman win t=0"),
          dep.var.labels = c(" Share Women Elected", "Votes Women", "Votes Men",
                             "Votes Party", "Seats Party"),
          add.lines = list(c("Bandwidth",
                             rep(18, 5))),
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/semiclosed_res.tex",
          title = "The Effect of Electing Women in Parties with Semi-Closed Lists",
          label = "tab:semiopen")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:28:54

Table 3: The Effect of Electing Women in Parties with Semi-Closed Lists

	Dependent variable:				
	Share Women Elected	Votes Women	Votes Men	Votes Party	Seats Party
	(1)	(2)	(3)	(4)	(5)
Woman win t=0	0.081**	0.123	0.271	0.188*	1.097***
	(0.032)	(0.121)	(0.171)	(0.101)	(0.389)
Constant	0.256***	5.584***	6.589***	6.947***	3.366***
	(0.019)	(0.234)	(0.364)	(0.247)	(0.985)
Bandwidth	18	18	18	18	18
Observations	711	711	711	711	711
Note:					*p<0.1; **p<0.05; ***p<0.01

Table E2

```
top2_cand$close <- ifelse(top2_cand$fractile_thres <= 18, "Close Elections", "No Close Elections")

top2_cand <- top2_cand %>%
  group_by( party, KOMMUNE) %>%
  mutate(national_party = case_when(party %in% c("A", "B", "V", "C", "F", "I", "O", "Ø", "Å") ~ 1,
                                    TRUE ~ 0),
         left_party = case_when(party %in% c("A", "B", "F", "Ø") ~ 1,
                                TRUE ~ 0),
         big_city = ifelse(kom_nr %in% c(101,147,751,461,851),1,0))

top2_cand <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarize(av_inc = median(income, na.rm=T)) %>%
  right_join(top2_cand, by = c("party", "KOMMUNE", "election_year"))

desc_data <- top2_cand %>% ungroup %>%
  select(close,
         "Total Votes for Party"= party_vote,
         "N Elected for Party" = n_elect,
         "Represented in Parliament" = national_party,
         "Left Party" = left_party,
         "Big City" = big_city,
         "Proportion Women" = prop_female,
         "Estimated Personal Votes" = mid_pers_vote,
         "Median Income" = av_inc,
         "Average Age" = av_age)

datasummary_balance(~close,
                    data = desc_data,
                    output = 'gt')
```

	Close Elections (N=819)		No Close Elections (N=581)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Total Votes for Party	3598.6	5734.1	681.6	955.8	-2916.9	204.3
N Elected for Party	4.8	3.4	1.5	1.0	-3.4	0.1
Represented in Parliament	0.9	0.3	0.7	0.5	-0.2	0.0
Left Party	0.4	0.5	0.4	0.5	0.0	0.0
Big City	0.0	0.2	0.1	0.3	0.0	0.0
Proportion Women	0.3	0.3	0.3	0.4	0.0	0.0
Estimated Personal Votes	175.8	148.3	215.0	706.6	39.2	29.8
Median Income	342630.7	101442.2	296916.1	103351.5	-45714.6	5586.0
Average Age	48.5	4.1	48.5	6.7	0.0	0.3

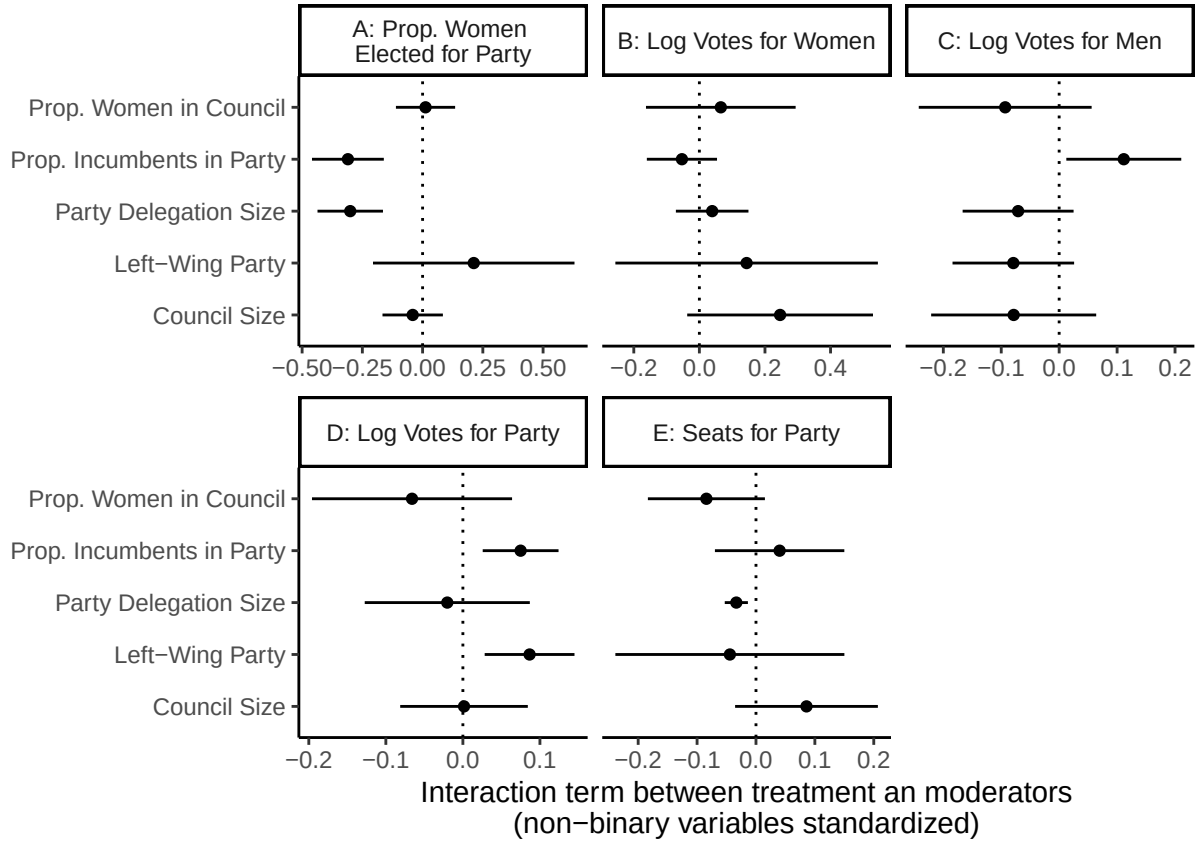
Appendix F

Figure F1

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")


top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

## define left/right wing
left_parties <- c("A", "B", "F", "Ø")
right_parties <- c("C", "O", "V")

all_candidates <- all_candidates %>%
  mutate(lr = case_when(party %in% left_parties ~ 1,
                        PARTI == "Å Alternativet" ~ 1,
                        party %in% right_parties ~ 0,
                        PARTI == "I Liberal Alliance" ~ 0,
                        PARTI == "K Kristendemokraterne" ~ 0,
                        TRUE ~ NA_real_))

# proportion of women in election-party-municipality groups
party_agg <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(prop_incumbents = mean(incumbent, na.rm = T),
            lr = mean(lr, na.rm = TRUE))

# covariates at the municipality level
muni_cov <- all_candidates %>%
  group_by(KOMMUNE, election_year) %>%
  filter(elected == 1) %>%
  summarise(council_size = n(),
            fem_in_muni = mean(female,na.rm = TRUE))

top2_cand <- left_join(top2_cand, muni_cov,
                       by = c("KOMMUNE", "election_year"))

top2_cand <- left_join(top2_cand, party_agg,
                       by = c("party", "KOMMUNE", "election_year"))


# standardizing variables
top2_cand$prop_female_t1 <- top2_cand$prop_female_t1/sd(top2_cand$prop_female_t1, na.rm=T)
top2_cand$fem_vote_t1 <- top2_cand$fem_vote_t1/sd(top2_cand$fem_vote_t1, na.rm=T)
top2_cand$male_vote_t1 <- top2_cand$male_vote_t1/sd(top2_cand$male_vote_t1, na.rm=T)
top2_cand$party_vote_t1 <- top2_cand$party_vote_t1/sd(top2_cand$party_vote_t1, na.rm=T)
top2_cand$n_elect_t1 <- top2_cand$n_elect_t1/sd(top2_cand$n_elect_t1, na.rm=T)

top2_cand$prop_female <- top2_cand$prop_female/sd(top2_cand$prop_female, na.rm=T)
top2_cand$fem_in_muni <- top2_cand$fem_in_muni/sd(top2_cand$fem_in_muni, na.rm=T)
top2_cand$council_size <- top2_cand$council_size/sd(top2_cand$council_size, na.rm=T)
top2_cand$prop_incumbents <- top2_cand$prop_incumbents/sd(top2_cand$prop_incumbents, na.rm=T)
```

```r
top2_cand$n_elect <- top2_cand$n_elect/sd(top2_cand$n_elect, na.rm=T)

eqs <- expand.grid(c("prop_female_t1", "fem_vote_t1", "male_vote_t1", "party_vote_t1", "n_elect_t1"),
                   c("~ female_win*lr","~ female_win*fem_in_muni",
                     "~ female_win*council_size",
                     "~ female_win*n_elect", "~ female_win*prop_incumbents"))

full_eqs <- paste0(eqs[,1], eqs[,2], "|0|0|KOMMUNE + party")

res <- list()

for(i in 1:length(full_eqs)){

  crnt_mod <- felm(as.formula(full_eqs[i]),
                   data = filter(top2_cand, fractile_thres <= 18))
  crnt_est <- tidy(crnt_mod)[4,]
  res[[i]] <- data.frame(crnt_est,
                         DV = eqs[i, 1])

}

res <- do.call("rbind", res)

res <- res %>%
  separate(col = term, into = c("treat", "moderator"), sep = "\\:") %>%
  select(-c(treat)) %>%
  mutate(lwr95 = estimate - 1.96*std.error,
         upr95 = estimate + 1.96*std.error,
         moder_clean = case_when(moderator == "lr" ~ "Left-Wing Party",
                                 moderator == "fem_in_muni" ~ "Prop. Women in Council",
                                 moderator == "council_size" ~ "Council Size",
                                 moderator == "n_elect" ~ "Party Delegation Size",
                                 moderator == "prop_incumbents" ~ "Prop. Incumbents in Party"),
         DV_clean = case_when(DV == "prop_female_t1" ~ "A: Prop. Women \nElected for Party",
                              DV == "fem_vote_t1" ~ "B: Log Votes for Women",
                              DV == "male_vote_t1" ~ "C: Log Votes for Men",
                              DV == "party_vote_t1" ~ "D: Log Votes for Party",
                              DV == "n_elect_t1" ~ "E: Seats for Party"))



# plot with the other moderators for the appendix
p <- ggplot(res, aes(y = moder_clean, x = estimate)) +
  geom_point() +
  geom_linerange(aes(xmin = lwr95,
                     xmax = upr95)) +
  #geom_errorbarh(aes(xmin = lwr95, xmax = upr95), height = 0)+
  theme_classic() +
  facet_wrap(~DV_clean, scale = "free_x") +
  geom_vline(xintercept = 0, lty = 3) +
  labs(x = "Interaction term between treatment an moderators\n(non-binary variables standardized)",
       y = "")
```

```
p
```


A: Prop. Women Elected for Party
B: Log Votes for Women
C: Log Votes for Men
D: Log Votes for Party
E: Seats for Party
Prop. Women in Council
Prop. Incumbents in Party
Party Delegation Size
Left−Wing Party
Council Size
−0.50 −0.25 0.00 0.25 0.50
−0.2 0.0 0.2 0.4
−0.2 −0.1 0.0 0.1 0.2
−0.2 −0.1 0.0 0.1
−0.2 −0.1 0.0 0.1 0.2
Interaction term between treatment an moderators
(non−binary variables standardized)

Appendix G

Figure G2

```
top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")


bw_map <- function(bw){

  # select candidate pools for every bandwidth

  top2_cand_bw <- top2_cand %>%
    filter(fractile_thres < bw)

  # estimate models

  fem_elect <- felm(prop_female_t1 ~ female_win |0|0|KOMMUNE + party,
                    data = filter(top2_cand_bw))

  fem_vote <- felm(log(fem_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                   data = filter(top2_cand_bw))

  male_vote <- felm(log(male_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                    data = filter(top2_cand_bw))

  party_vote_mod <- felm(log(party_vote_t1) ~ female_win |0|0|KOMMUNE + party,
                         data = filter(top2_cand_bw))

  mandates_mod <- felm(n_elect_t1 ~ female_win |0|0|KOMMUNE + party,
                       data = filter(top2_cand_bw))


  # extract estimates
  fem_elect_estimates <- data.frame(PE = coef(fem_elect)[2],
                                    SE = sqrt(diag(vcov(fem_elect)))[2],
                                    spec = "A: Share Women Elected t+1",
                                    FE = "No FE",
                                    df = fem_elect$df,
                                    f_stat = summary(fem_elect)$F.fstat[1],
                                    bw = bw,
                                    n = fem_elect$N)


  fem_vote_estimates <- data.frame(PE = coef(fem_vote)[2],
                                   SE = sqrt(diag(vcov(fem_vote)))[2],
                                   spec = "B: Log Votes Women t+1",
                                   FE = "No FE",
                                   df = fem_vote$df,
                                   f_stat = NA,
                                   bw = bw,
                                   n = fem_elect$N)

  male_vote_estimates <- data.frame(PE = coef(male_vote)[2],
```

```
                                    SE = sqrt(diag(vcov(male_vote)))[2],
                                    spec = "C: Log Votes Men t+1",
                                    FE = "No FE",
                                    df = male_vote$df,
                                    f_stat = NA,
                                    bw = bw,
                                    n = male_vote$N)

  party_vote_estimates <- data.frame(PE = coef(party_vote_mod)[2],
                                     SE = sqrt(diag(vcov(party_vote_mod)))[2],
                                     spec = "D: Log Votes Party t+1",
                                     FE = "No FE",
                                     df = party_vote_mod$df,
                                     f_stat = NA,
                                     bw = bw,
                                     n = party_vote_mod$N)

  mandates_estimates <- data.frame(PE = coef(mandates_mod)[2],
                                   SE = sqrt(diag(vcov(mandates_mod)))[2],
                                   spec = "E: Seats Party t+1",
                                   FE = "No FE",
                                   df = mandates_mod$df,
                                   f_stat = NA,
                                   bw = bw,
                                   n = mandates_mod$N)

  return(rbind(fem_elect_estimates,
               fem_vote_estimates,
               male_vote_estimates,
               party_vote_estimates,
               mandates_estimates))

}


## extract estimates into df of estimates
all_bw <- map(seq(0.01, 50, 1),
              ~ bw_map(bw = .))

all_bw_plot_df <- do.call("rbind", all_bw)

## add optimal bandwidths (manually, as they have to be rounded):
all_bw_plot_df <- all_bw_plot_df %>%
  mutate(optimal_bw = case_when(bw==18.01~"Selected BW",
                                TRUE~"BW")) %>%
  mutate(optimal_bw = factor(optimal_bw, levels = c("Selected BW",
                                                    "BW")))

## plot estimates over bandwidths
all_bw_plot_df %>%
  ggplot(data=., aes(x = bw, y = PE, color = optimal_bw, size = optimal_bw)) +
  geom_point() +
  geom_linerange(aes(ymin = PE - 1.96*SE,
```

```
                    ymax = PE + 1.96*SE)) +
theme_classic() +
facet_wrap(~ spec, scales = "free_y") +
geom_hline(yintercept = 0, lty = 3) +
ylab("Effect of an Additional Woman in Office") +
xlab("Margin of Win/Loss (Percent)") +
scale_size_manual("", values = c(0.4, 0.4)) +
scale_color_manual("", values = c("black", "gray75")) +
theme(legend.position = "none")
```

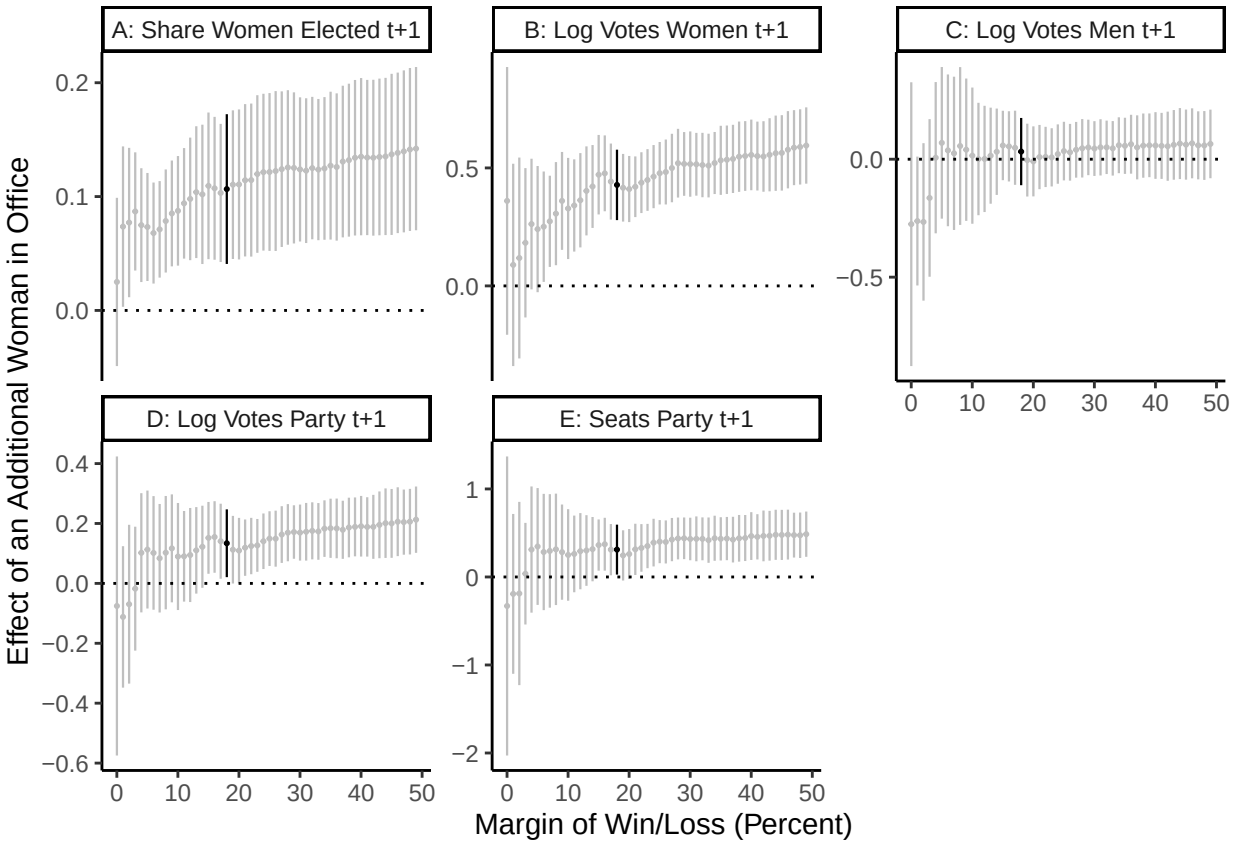


Figure G3

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")

top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

party_agg <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/party_agg_w_outcomes.rds")


## create individual-level covariates
all_candidates <- all_candidates %>%
  mutate(college = case_when(education %in% c("Bachelor",
                                              "Forskeruddannelser",
                                              "Mellemlange videregående uddannelser",
                                              "Lange videregående uddannelser") ~ 1,
                             TRUE ~ 0),
         rich = case_when(income > quantile(income, probs = 0.5, na.rm=T) ~ 1,
                          TRUE ~ 0),
         old = case_when(age > quantile(age, probs = 0.5, na.rm=T) ~ 1,
                         TRUE ~ 0),
         competent = case_when(inc_res > quantile(inc_res, probs = 0.5, na.rm = T)~1,
                               TRUE ~ 0),
         college = ifelse(is.na(education), NA, college),
         rich    = ifelse(is.na(income), NA, rich),
         old     = ifelse(is.na(age), NA, old),
         competent = ifelse(is.na(competent), NA, competent))



# create individual level data for involved parties
bw_data <-
  filter(top2_cand,
         fractile_thres <= 18) %>%
  select(c("party", "KOMMUNE", "election_year")) %>%
  left_join(., all_candidates) %>%
  filter(OPSTILLINGSFORM == "Sideordnet") %>%
  group_by(party, KOMMUNE, election_year) %>%
  arrange(fractile_thres) %>%
  slice(1:2)

### find difference in age, education, competence and income in treatment group
# estimate gender gap with OLS
#these regression tables are not in the paper manuscript or appendix
# # USD/DKK exchange rate 2022-08-24 (Danish National Bank):
stargazer(lm(income / 7.45  ~ female, data =  bw_data))
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: to, aug 15, 2024 - 16:40:31

```
stargazer(lm(college ~ female, data = bw_data))
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at

Table 5

	Dependent variable:
	income/7.45
female	−9,113.724***
	(2,700.804)
Constant	53,621.750***
	(1,909.757)
Observations	1,630
R^2	0.007
Adjusted R^2	0.006
Residual Std. Error	54,520.130 (df = 1628)
F Statistic	11.387*** (df = 1; 1628)
Note:	*p<0.1; **p<0.05; ***p<0.01

gmail.com % Date and time: to, aug 15, 2024 - 16:40:31

Table 6

	Dependent variable:
	college
female	0.076***
	(0.024)
Constant	0.357***
	(0.017)
Observations	1,609
R^2	0.006
Adjusted R^2	0.005
Residual Std. Error	0.488 (df = 1607)
F Statistic	9.662*** (df = 1; 1607)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
stargazer(lm(age ~ female, data = bw_data))
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: to, aug 15, 2024 - 16:40:31

```
stargazer(lm(inc_res~female, data = bw_data))
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: to, aug 15, 2024 - 16:40:31

```
###################### create placebo treatments
### create separate data objects for each placebo

top2_cand <- all_candidates %>%
```

Table 7

	Dependent variable:
	age
female	-1.548^{***}
	(0.552)
Constant	48.659^{***}
	(0.390)
Observations	1,630
R^2	0.005
Adjusted R^2	0.004
Residual Std. Error	11.148 (df = 1628)
F Statistic	7.863^{***} (df = 1; 1628)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 8

	Dependent variable:
	inc_res
female	0.201^{***}
	(0.032)
Constant	0.076^{***}
	(0.023)
Observations	1,630
R^2	0.023
Adjusted R^2	0.023
Residual Std. Error	0.649 (df = 1628)
F Statistic	38.985^{***} (df = 1; 1628)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
  filter(OPSTILLINGSFORM == "Sideordnet" & !is.na(education)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  mutate(any_fem = sum(female, na.rm=T)) %>%
  filter(any_fem > 0) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  filter(female == 0) %>%
  arrange(party, KOMMUNE, election_year)

top2_cand <- top2_cand %>%
  group_by(party, KOMMUNE, election_year) %>%
  mutate(any_col = sum(college, na.rm=T)) %>%
  filter(any_col > 0) %>%
  arrange(fractile_thres) %>%
  slice(1:2)

# make placebo treatment variables
plac_win <- top2_cand %>%
  mutate(college_win = case_when(college == 1 & elected == 1 ~ 1,
                                 TRUE ~ 0)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(college_win = mean(college_win, na.rm = T),
            college_participant = mean(college, na.rm = T))%>%
  mutate(college_participant = case_when(college_participant == .50 ~ 1,
                                         TRUE ~ 0),
         college_win = case_when(college_win > 0 ~ 1,
                                 TRUE ~ 0))

top2_cand <- left_join(top2_cand, plac_win, by = c("party", "KOMMUNE", "election_year"))

##########
# college

college_df <- top2_cand %>%
  group_by(party, KOMMUNE, election_year, college_win) %>%
  summarise(fractile_thres = mean(fractile_thres, na.rm=T))

college_df <- college_df %>%
  mutate(frac_bw = case_when(college_win == 0 ~ fractile_thres*-1,
                             TRUE ~ fractile_thres))


# merge with party-list-election data

college_df <- left_join(college_df, party_agg,
                        by = c("party", "KOMMUNE", "election_year"))

college_df$cluster_id <- college_df %>%
  group_by(party, KOMMUNE) %>%
  group_indices()
```

```
################
# repeat for old

top2_cand <- all_candidates %>%
  filter(OPSTILLINGSFORM == "Sideordnet" & !is.na(age)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  mutate(any_fem = sum(female, na.rm=T)) %>%
  filter(any_fem > 0) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  filter(female == 0) %>%
  arrange(party, KOMMUNE, election_year)


# make placebo treatment variables
plac_win <- top2_cand %>%
  mutate(old_win = case_when(old == 1 & elected == 1 ~ 1,
                             TRUE ~ 0)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(old_win = mean(old_win, na.rm = T),
            old_participant = mean(old, na.rm = T))%>%
  mutate(old_participant = case_when(old_participant == .50 ~ 1,
                                     TRUE ~ 0),
         old_win = case_when(old_win > 0 ~ 1,
                             TRUE ~ 0))

top2_cand <- left_join(top2_cand, plac_win, by = c("party", "KOMMUNE", "election_year"))

# create separate objects for each placebo
old_df <- top2_cand %>%
  group_by(party, KOMMUNE, election_year, old_win) %>%
  summarise(fractile_thres = mean(fractile_thres, na.rm=T))

old_df <- old_df %>%
  mutate(frac_bw = case_when(old_win == 0 ~ fractile_thres*-1,
                             TRUE ~ fractile_thres))


# create and merge with party-list-election data
old_df <- left_join(old_df, party_agg,
                    by = c("party", "KOMMUNE", "election_year"))

old_df$cluster_id <- old_df %>%
  group_by(party, KOMMUNE) %>%
  group_indices()

##################
# repeat for rich

top2_cand <- all_candidates %>%
  filter(OPSTILLINGSFORM == "Sideordnet" & !is.na(income)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  mutate(any_fem = sum(female, na.rm=T)) %>%
```

```
  filter(any_fem > 0) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  filter(female == 0) %>%
  arrange(party, KOMMUNE, election_year)


# make placebo treatment variables
plac_win <- top2_cand %>%
  mutate(rich_win = case_when(rich == 1 & elected == 1 ~ 1,
                              TRUE ~ 0)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(rich_win = mean(rich_win, na.rm = T),
            rich_participant = mean(rich, na.rm = T))%>%
  mutate(rich_participant = case_when(rich_participant == .50 ~ 1,
                                      TRUE ~ 0),
         rich_win = case_when(rich_win > 0 ~ 1,
                              TRUE ~ 0))

top2_cand <- left_join(top2_cand, plac_win, by = c("party", "KOMMUNE", "election_year"))

# create separate objects for each placebo
rich_df <- top2_cand %>%
  group_by(party, KOMMUNE, election_year, rich_win) %>%
  summarise(fractile_thres = mean(fractile_thres, na.rm=T))

rich_df <- rich_df %>%
  mutate(frac_bw = case_when(rich_win == 0 ~ fractile_thres*-1,
                             TRUE ~ fractile_thres))


# create and merge with party-list-election data
rich_df <- left_join(rich_df, party_agg,
                     by = c("party", "KOMMUNE", "election_year"))

rich_df$cluster_id <- rich_df %>%
  group_by(party, KOMMUNE) %>%
  group_indices()


#################
# repeat for earning score (competence)

top2_cand <- all_candidates %>%
  filter(OPSTILLINGSFORM == "Sideordnet" & !is.na(income)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  mutate(any_fem = sum(female, na.rm=T)) %>%
  filter(any_fem > 0) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  filter(female == 0) %>%
  arrange(party, KOMMUNE, election_year)
```

```r
# make placebo treatment variables
plac_win <- top2_cand %>%
  mutate(competent_win = case_when(competent == 1 & elected == 1 ~ 1,
                                   TRUE ~ 0)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(competent_win = mean(competent_win, na.rm = T),
            competent_participant = mean(competent, na.rm = T))%>%
  mutate(competent_participant = case_when(competent_participant == .50 ~ 1,
                                           TRUE ~ 0),
         competent_win = case_when(competent_win > 0 ~ 1,
                                   TRUE ~ 0))

top2_cand <- left_join(top2_cand, plac_win, by = c("party", "KOMMUNE", "election_year"))

# create separate objects for each placebo
competent_df <- top2_cand %>%
  group_by(party, KOMMUNE, election_year, competent_win) %>%
  summarise(fractile_thres = mean(fractile_thres, na.rm=T))

competent_df <- competent_df %>%
  mutate(frac_bw = case_when(competent_win == 0 ~ fractile_thres*-1,
                             TRUE ~ fractile_thres))


# create and merge with party-list-election data
competent_df <- left_join(competent_df, party_agg,
                          by = c("party", "KOMMUNE", "election_year"))

competent_df$cluster_id <- competent_df %>%
  group_by(party, KOMMUNE) %>%
  group_indices()

## create one big df of the three placebo dfs
# ad label for placebo
rich_df$placebo <- "Rich"
college_df$placebo <- "College"
old_df$placebo <- "Old"
competent_df$placebo <- "Competent"

# make the treatment have the same name: placebo_win
rich_df$placebo_win <- rich_df$rich_win
college_df$placebo_win <- college_df$college_win
old_df$placebo_win <- old_df$old_win
competent_df$placebo_win <- competent_df$competent_win

placebo_df <- bind_rows(rich_df, college_df, old_df, competent_df)

################################################################################
##### EXTRACT PLACEBO ESTIMATES (RICH, OLD, COLLEGE) FOR THE FIVE OUTCOMES #####
################################################################################

# select candidate pools within bandwidths
```

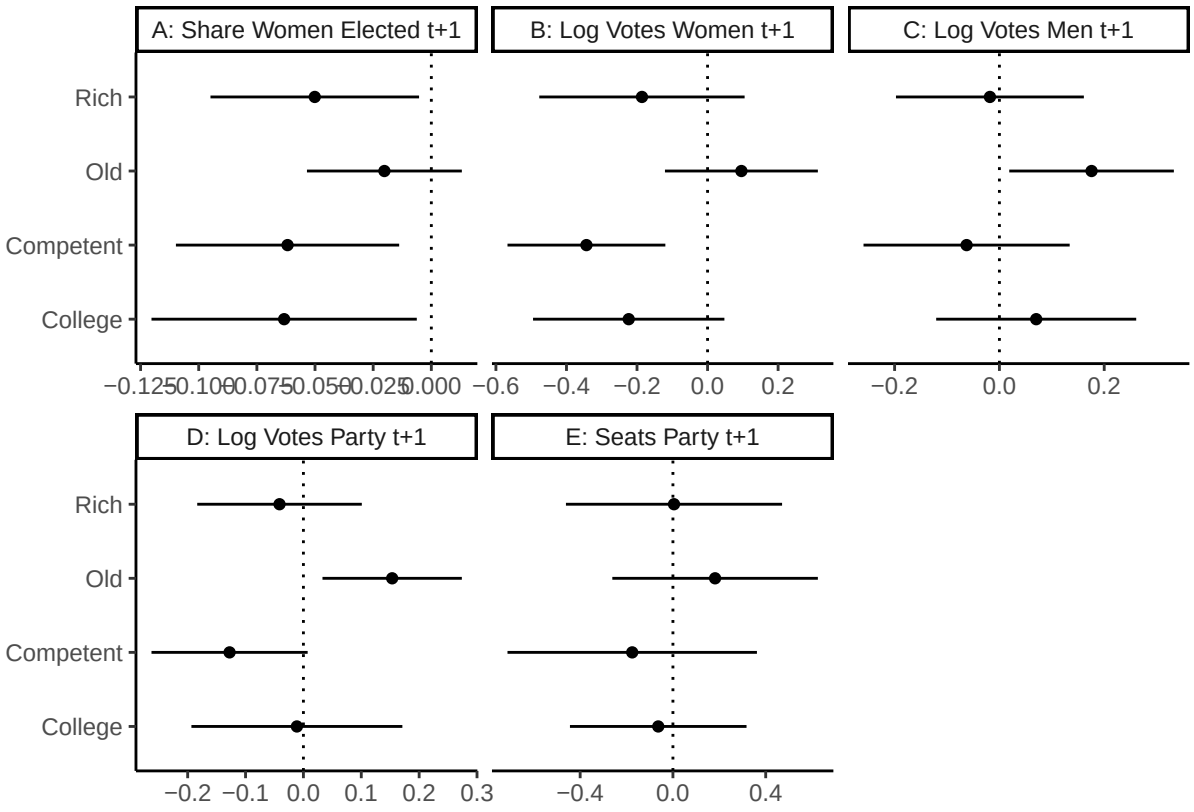
```r
top2_cand_bw <- placebo_df

# estimate models for rich

fem_elect_rich <- felm(prop_female_t1 ~ placebo_win |0|0|KOMMUNE + party,
                       data = top2_cand_bw, subset = placebo=="Rich" & fractile_thres <= 18)

fem_vote_rich <- felm(log(fem_vote_t1+1) ~ placebo_win |0|0|KOMMUNE + party,
                      data = top2_cand_bw, subset = placebo=="Rich" & fractile_thres <= 18)


male_vote_rich <- felm(log(male_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                       data = top2_cand_bw, subset = placebo=="Rich" & fractile_thres <= 18)

party_vote_mod_rich <- felm(log(party_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                            data = top2_cand_bw, subset = placebo=="Rich" & fractile_thres <= 18)

mandates_mod_rich <- felm(n_elect_t1 ~ placebo_win |0|0|KOMMUNE + party,
                          data = top2_cand_bw, subset = placebo=="Rich" & fractile_thres <= 18)


# extract estimates
fem_elect_estimates_df_rich <- data.frame(PE = fem_elect_rich$coefficients[2,1],
                                          SE = fem_elect_rich$cse[2],
                                          placebo = "Rich",
                                          spec = "A: Share Women Elected t+1",
                                          bw = 18,
                                          n = fem_elect_rich$N)

fem_vote_estimates_df_rich <- data.frame(PE = fem_vote_rich$coefficients[2,1],
                                         SE = fem_vote_rich$cse[2],
                                         placebo = "Rich",
                                         spec = "B: Log Votes Women t+1",
                                         bw = 18,
                                         n = fem_vote_rich$N)

male_vote_estimates_df_rich <- data.frame(PE = male_vote_rich$coefficients[2,1],
                                          SE = male_vote_rich$cse[2],
                                          placebo = "Rich",
                                          spec = "C: Log Votes Men t+1",
                                          bw = 18,
                                          n = male_vote_rich$N)

party_vote_estimates_df_rich <- data.frame(PE = party_vote_mod_rich$coefficients[2,1],
                                           SE = party_vote_mod_rich$cse[2],
                                           placebo = "Rich",
                                           spec = "D: Log Votes Party t+1",
                                           bw = 18,
                                           n = party_vote_mod_rich$N)

mandates_estimates_df_rich <- data.frame(PE = mandates_mod_rich$coefficients[2,1],
                                         SE = mandates_mod_rich$cse[2],
                                         placebo = "Rich",
```

```
                                        spec = "E: Seats Party t+1",
                                        bw = 18,
                                        n = mandates_mod_rich$N)

## estimate for old

fem_elect_old <- felm(prop_female_t1 ~ placebo_win |0|0|KOMMUNE + party,
                      data = top2_cand_bw, subset = placebo=="Old" & fractile_thres <= 18)

fem_vote_old <- felm(log(fem_vote_t1+1) ~ placebo_win |0|0|KOMMUNE + party,
                     data = top2_cand_bw, subset = placebo=="Old" & fractile_thres <= 18)

male_vote_old <- felm(log(male_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                      data = top2_cand_bw, subset = placebo=="Old" & fractile_thres <= 18)

party_vote_mod_old <- felm(log(party_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                           data = top2_cand_bw, subset = placebo=="Old" & fractile_thres <= 18)

mandates_mod_old <- felm(n_elect_t1 ~ placebo_win |0|0|KOMMUNE + party, data = top2_cand_bw,
                         subset = placebo=="Old" & fractile_thres <= 18)

# extract estimates
fem_elect_estimates_df_old <- data.frame(PE = fem_elect_old$coefficients[2,1],
                                         SE = fem_elect_old$cse[2],
                                         placebo = "Old",
                                         spec = "A: Share Women Elected t+1",
                                         bw = 18,
                                         n = fem_elect_old$N)

fem_vote_estimates_df_old <- data.frame(PE = fem_vote_old$coefficients[2,1],
                                        SE = fem_vote_old$cse[2],
                                        placebo = "Old",
                                        spec = "B: Log Votes Women t+1",
                                        bw = 18,
                                        n = fem_vote_old$N)

male_vote_estimates_df_old <- data.frame(PE = male_vote_old$coefficients[2,1],
                                         SE = male_vote_old$cse[2],
                                         placebo = "Old",
                                         spec = "C: Log Votes Men t+1",
                                         bw = 18,
                                         n = male_vote_old$N)

party_vote_estimates_df_old <- data.frame(PE = party_vote_mod_old$coefficients[2,1],
                                          SE = party_vote_mod_old$cse[2],
                                          placebo = "Old",
                                          spec = "D: Log Votes Party t+1",
                                          bw = 18,
                                          n = party_vote_mod_old$N)

mandates_estimates_df_old <- data.frame(PE = mandates_mod_old$coefficients[2,1],
                                        SE = mandates_mod_old$cse[2],
                                        placebo = "Old",
```

```
                                        spec = "E: Seats Party t+1",
                                        bw = 18,
                                        n = mandates_mod_old$N)

# estimate models for college

fem_elect_college <- felm(prop_female_t1 ~ placebo_win |0|0|KOMMUNE + party,
                          data = top2_cand_bw, subset = placebo=="College" & fractile_thres <= 18)

fem_vote_college <- felm(log(fem_vote_t1+1) ~ placebo_win |0|0|KOMMUNE + party,
                         data = top2_cand_bw, subset = placebo=="College" & fractile_thres <= 18)

male_vote_college <- felm(log(male_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                          data = top2_cand_bw, subset = placebo=="College" & fractile_thres <= 18)

party_vote_mod_college <- felm(log(party_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                               data = top2_cand_bw, subset = placebo=="College" & fractile_thres <= 18)

mandates_mod_college <- felm(n_elect_t1 ~ placebo_win |0|0|KOMMUNE + party,
                             data = top2_cand_bw, subset = placebo=="College" & fractile_thres <= 18)


# extract estimates
fem_elect_estimates_df_college <- data.frame(PE = fem_elect_college$coefficients[2,1],
                                             SE = fem_elect_college$cse[2],
                                             placebo = "College",
                                             spec = "A: Share Women Elected t+1",
                                             bw = 18,
                                             n = fem_elect_college$N)

fem_vote_estimates_df_college <- data.frame(PE = fem_vote_college$coefficients[2,1],
                                            SE = fem_vote_college$cse[2],
                                            placebo = "College",
                                            spec = "B: Log Votes Women t+1",
                                            bw = 18,
                                            n = fem_vote_college$N)

male_vote_estimates_df_college <- data.frame(PE = male_vote_college$coefficients[2,1],
                                             SE = male_vote_college$cse[2],
                                             placebo = "College",
                                             spec = "C: Log Votes Men t+1",
                                             bw = 18,
                                             n = male_vote_college$N)

party_vote_estimates_df_college <- data.frame(PE = party_vote_mod_college$coefficients[2,1],
                                              SE = party_vote_mod_college$cse[2],
                                              placebo = "College",
                                              spec = "D: Log Votes Party t+1",
                                              bw = 18,
                                              n = party_vote_mod_college$N)

mandates_estimates_df_college <- data.frame(PE = mandates_mod_college$coefficients[2,1],
                                            SE = mandates_mod_college$cse[2],
```

```
                                          placebo = "College",
                                          spec = "E: Seats Party t+1",
                                          bw = 18,
                                          n = mandates_mod_college$N)

## estimate for competence
fem_elect_comp <- felm(prop_female_t1 ~ placebo_win |0|0|KOMMUNE + party,
                       data = top2_cand_bw, subset = placebo=="Competent" & fractile_thres <= 18)

fem_vote_comp <- felm(log(fem_vote_t1+1) ~ placebo_win |0|0|KOMMUNE + party,
                      data = top2_cand_bw, subset = placebo=="Competent" & fractile_thres <=18)

male_vote_comp <- felm(log(male_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                       data = top2_cand_bw, subset = placebo=="Competent" & fractile_thres <= 18)

party_vote_mod_comp <- felm(log(party_vote_t1) ~ placebo_win |0|0|KOMMUNE + party,
                            data = top2_cand_bw, subset = placebo=="Competent" & fractile_thres <= 18)

mandates_mod_comp <- felm(n_elect_t1 ~ placebo_win |0|0|KOMMUNE + party, data = top2_cand_bw,
                          subset = placebo=="Competent" & fractile_thres <= 18)

# extract estimates
fem_elect_estimates_df_comp <- data.frame(PE = fem_elect_comp$coefficients[2,1],
                                          SE = fem_elect_comp$cse[2],
                                          placebo = "Competent",
                                          spec = "A: Share Women Elected t+1",
                                          bw = 18,
                                          n = fem_elect_comp$N)

fem_vote_estimates_df_comp <- data.frame(PE = fem_vote_comp$coefficients[2,1],
                                         SE = fem_vote_comp$cse[2],
                                         placebo = "Competent",
                                         spec = "B: Log Votes Women t+1",
                                         bw = 18,
                                         n = fem_vote_comp$N)

male_vote_estimates_df_comp <- data.frame(PE = male_vote_comp$coefficients[2,1],
                                          SE = male_vote_comp$cse[2],
                                          placebo = "Competent",
                                          spec = "C: Log Votes Men t+1",
                                          bw = 18,
                                          n = male_vote_comp$N)

party_vote_estimates_df_comp <- data.frame(PE = party_vote_mod_comp$coefficients[2,1],
                                           SE = party_vote_mod_comp$cse[2],
                                           placebo = "Competent",
                                           spec = "D: Log Votes Party t+1",
                                           bw = 18,
                                           n = party_vote_mod_comp$N)

mandates_estimates_df_comp <- data.frame(PE = mandates_mod_comp$coefficients[2,1],
                                         SE = mandates_mod_comp$cse[2],
                                         placebo = "Competent",
```

```
                                        spec = "E: Seats Party t+1",
                                        bw = 18,
                                        n = mandates_mod_comp$N)

# gather dfs

placebo_plot_df <- bind_rows(fem_elect_estimates_df_rich,
                             fem_vote_estimates_df_rich,
                             male_vote_estimates_df_rich,
                             party_vote_estimates_df_rich,
                             mandates_estimates_df_rich,

                             fem_elect_estimates_df_old,
                             fem_vote_estimates_df_old,
                             male_vote_estimates_df_old,
                             party_vote_estimates_df_old,
                             mandates_estimates_df_old,

                             fem_elect_estimates_df_comp,
                             fem_vote_estimates_df_comp,
                             male_vote_estimates_df_comp,
                             party_vote_estimates_df_comp,
                             mandates_estimates_df_comp,

                             fem_elect_estimates_df_college,
                             fem_vote_estimates_df_college,
                             male_vote_estimates_df_college,
                             party_vote_estimates_df_college,
                             mandates_estimates_df_college)

placebo_plot_df <- placebo_plot_df %>%
  mutate(bw = factor(bw))

placebo_plot_df %>%
  ggplot(data=., aes(x=PE, y=placebo)) +
  geom_point(position = position_dodge2(width = 0.5)) +
  geom_linerange(aes(xmin = PE - 1.96*SE,
                     xmax = PE + 1.96*SE),
                 position = position_dodge2(width = 0.5)) +
  theme_classic() +
  geom_vline(xintercept = 0, lty = 3, linetype = "dashed") +
  facet_wrap(~spec, scales = "free_x") +
  ylab("") +
  xlab("") +
  scale_shape_discrete("") +
  scale_color_manual("", values = c("black", "gray45", "gray75")) +
  theme(legend.position = c(0.85,0.3))
```


A: Share Women Elected t+1
B: Log Votes Women t+1
C: Log Votes Men t+1
D: Log Votes Party t+1
E: Seats Party t+1
Rich
Old
Competent
College
−0.125
−0.100
−0.075
−0.050
−0.025
0.000
−0.6
−0.4
−0.2
0.0
0.2
−0.2
0.0
0.2
−0.2
−0.1
0.0
0.1
0.2
0.3
−0.4
0.0
0.4

Appendix H

Figure H1

```
top2_cand <-
 readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

all_sims <- list()

for(bw in 1:22){


  df_in_bw <- top2_cand %>% ungroup %>% filter(fractile_thres <=bw)
  df_in_bw$obs_id <- 1:nrow(df_in_bw)

  n_treat <- sum(df_in_bw$female_win)

  this_bw <- list()
    set.seed(888)
    for(i in 1:1000){

    treated_munis <- sample(df_in_bw$obs_id, n_treat)

    sampled_munis <- df_in_bw %>%
      mutate(treated = case_when(obs_id %in% treated_munis ~ 1,
                                 TRUE ~ 0))



      sampled_munis <- sampled_munis %>%
        mutate(sim_prop_fem = prop_female_t1 + 0.114*treated,
               sim_fem_vote = fem_vote_t1 + (0.438*mean(fem_vote_t1, na.rm=T))*treated,
               sim_male_vote = male_vote_t1 + (0.009*mean(male_vote_t1, na.rm=T))*treated,
               sim_party_vote = party_vote_t1 + (0.126*mean(party_vote_t1, na.rm=T))*treated,
               sim_seats = n_elect_t1 + 0.328*treated)


    fem_elect <- felm(sim_prop_fem ~ female_win |0|0|KOMMUNE + party,
                      data = sampled_munis)

    fem_vote <- felm(log(sim_fem_vote+1) ~ female_win |0|0|KOMMUNE + party,
                     data = sampled_munis)

    male_vote <- felm(log(sim_male_vote+1) ~ female_win |0|0|KOMMUNE + party,
                      data = sampled_munis)

    party_vote_mod <- felm(log(sim_party_vote+1) ~ female_win |0|0|KOMMUNE + party,
                           data = sampled_munis)

    mandates_mod <- felm(sim_seats ~ female_win |0|0|KOMMUNE + party,
                         data = sampled_munis)


    this_bw[[i]] <- data.frame(est = c(coef(fem_elect)[2],
```

```
                        coef(fem_vote)[2],
                        coef(male_vote)[2],
                        coef(party_vote_mod)[2],
                        coef(mandates_mod)[2]),
               se = c(sqrt(diag(vcov(fem_elect)))[2],
                      sqrt(diag(vcov(fem_vote)))[2],
                     sqrt(diag(vcov(male_vote)))[2],
                      sqrt(diag(vcov(party_vote_mod)))[2],
                      sqrt(diag(vcov(mandates_mod)))[2]),
               DV = c("fem_elect", "fem_vote", "male_vote", "party_vote", "party_seat")) %>%
      mutate(sig = ifelse(abs(est/se) > 1.96, 1, 0 ))


    }

  all_sims[[bw]] <- this_bw


}


for(i in 1:22){

all_sims[[i]] <- do.call("rbind.data.frame", all_sims[[i]])


}



all_sims_df <- list()

for(i in 1:22){

  all_sims_df[[i]] <- all_sims[[i]] %>%
    group_by(DV) %>%
    summarize(power = mean(sig),
              bw = i)
}

all_sims_df <- do.call("rbind.data.frame", all_sims_df)

n_in_bw <- data.frame(n_obs = rep(NA, 22),
                      bw = 1:22)

for(bw in 1:22){

  df_in_bw <- filter(top2_cand, fractile_thres <= bw)
n_in_bw[bw,1] <- nrow(df_in_bw)

}
```

```
p_power <- ggplot(subset(all_sims_df, bw < 17), aes(x = bw, y = power, color = DV)) +
  geom_line() +
  geom_point() +
  theme_bw() +
  geom_hline(yintercept = .8, lty = 2) +
  labs(x = "Percentile of Election-Closeness\n(Bandwidth)",
       y = "Proportion Statistically Significant Estimates\n(Power)") +
  scale_x_continuous(breaks = seq(1,18,2)) +
  geom_text(data = subset(n_in_bw, bw < 18),
            aes(x = bw, y = 1.1, label = n_obs), inherit.aes = FALSE)

p_power
```

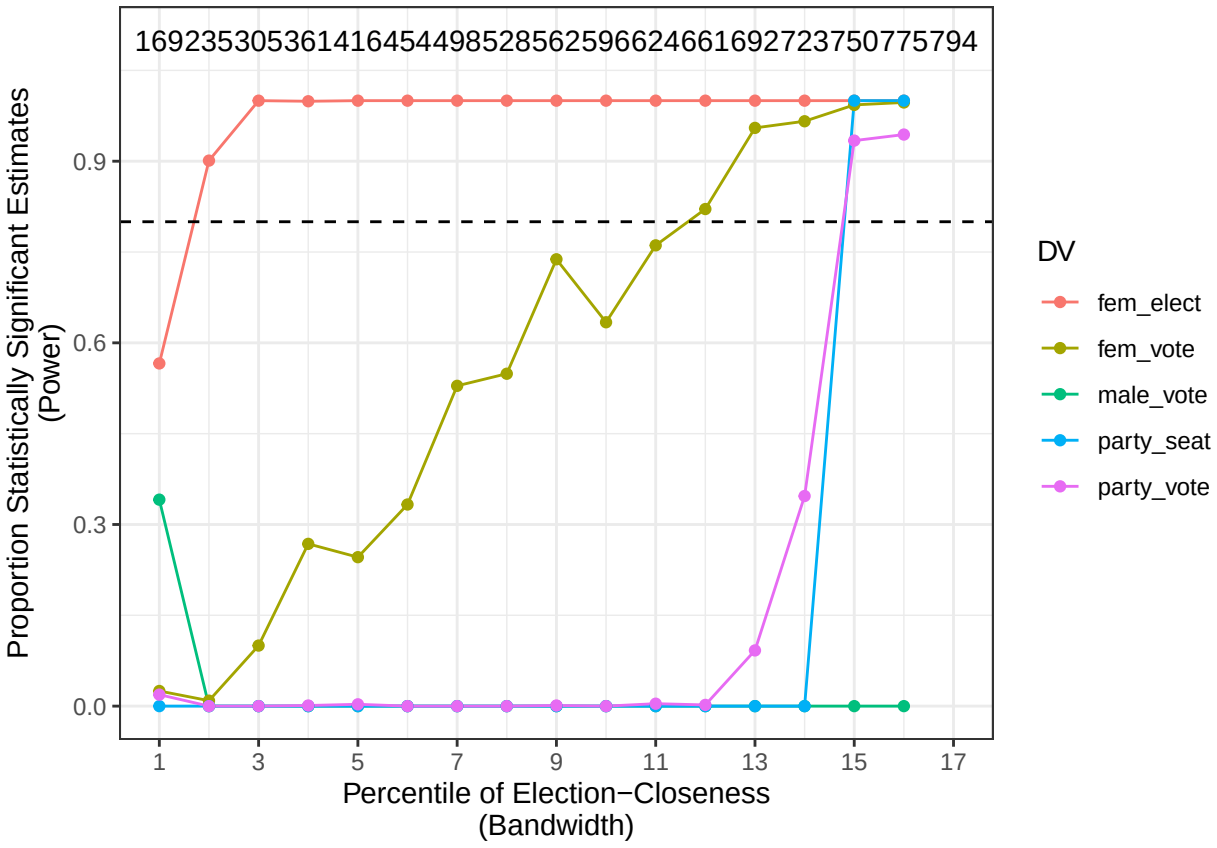
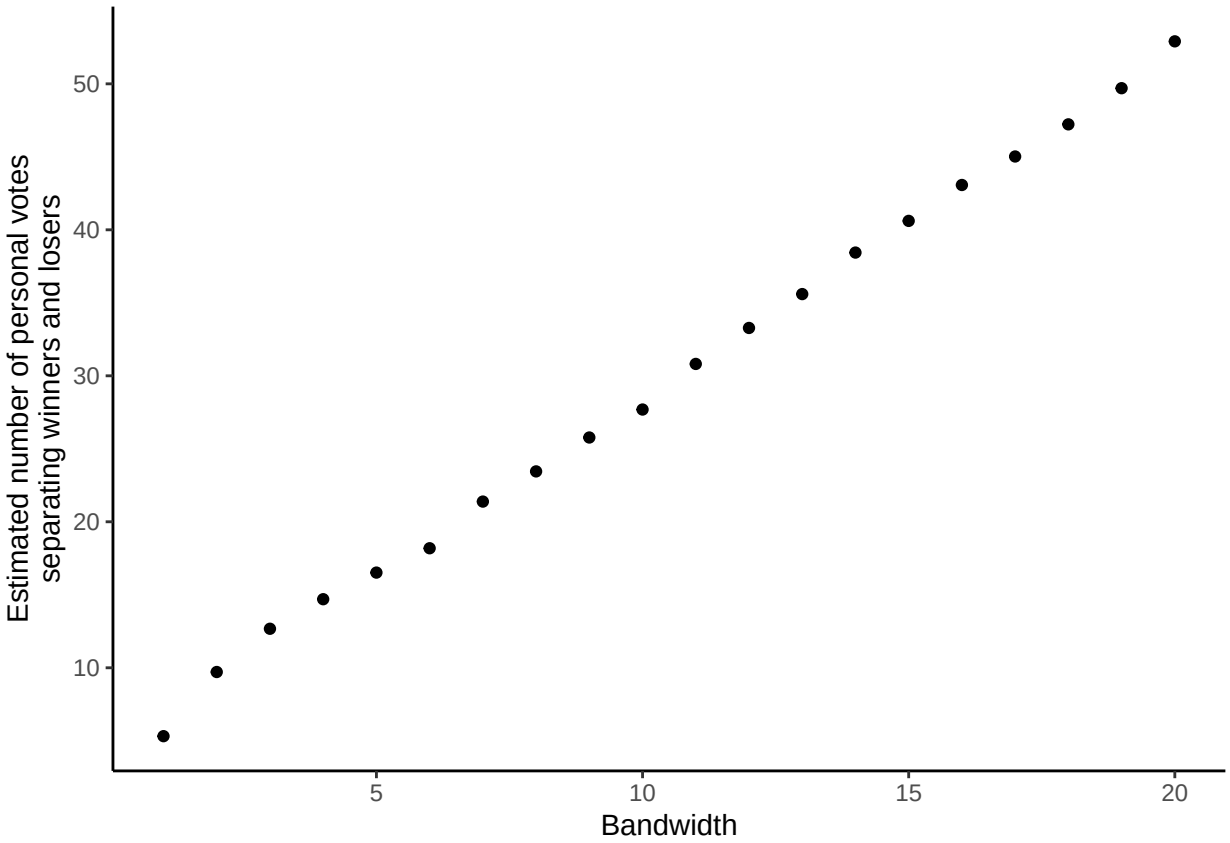


Figure H2

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")

#Estimate of personal votes
all_candidates$interval_points <- gsub(pattern =  "\\[", replacement = "",
                                       x = all_candidates$interval_points)
all_candidates$interval_points <- gsub(pattern =  "\\]", replacement = "",
                                       x = all_candidates$interval_points)

all_candidates <- all_candidates %>%
  separate(col = "interval_points", into = c("lwr", "upr"), sep = ",") %>%
  mutate(lwr = as.numeric(as.character(lwr)),
         upr = as.numeric(as.character(upr)),
         mid_pers_vote = (lwr+upr)/2)

# define variable to sort amalgated municipalities away as outcome year in 2005
all_candidates <- all_candidates %>%
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  #make dummy equaling zero for amalgated municipalities in 2005 election but otherwise 1
  mutate(muni_to_include = ifelse(kom_nr %!in% non_amalgated_vector & election_year==2001,0,1))




votes <- rep(NA, 20)

for(i in 1:20){
  this_bw <- all_candidates %>%
    filter(OPSTILLINGSFORM == "Sideordnet" & muni_to_include==1) %>%
    group_by(party, KOMMUNE, election_year) %>%
    arrange(fractile_thres) %>%
    slice(1:2) %>%
    filter(fractile_thres <= i)

  this_mod <- lm(mid_pers_vote ~ elected,
                 data = this_bw)

  votes[i] <- coef(this_mod)[2]


}

votes <- data.frame(votes = votes, bw = 1:20)

p <- ggplot(votes, aes(x = bw, y = votes)) +
  geom_point() +
  theme_classic() +
  labs(x = "Bandwidth", y = "Estimated number of personal votes\nseparating winners and losers")

p
```


Estimated number of personal votes separating winners and losers
50
40
30
20
10
5
10
15
20
Bandwidth

Appendix I

Table I1

```
top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds") %>%
  filter(fractile_thres<=18)

all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds") %>%
  # create variable for year-muni-party to select only candidates from our sample
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  mutate(cluster = str_c(election_year, kom_nr, party, sep = "-", collapse = NULL)) %>%
  filter(cluster %in% top2_cand$cluster)


#-------------------------
# incumbency advantage


# both with interaction
all_rerun <-felm(running_next_elec ~ elected*female |0|0|cluster_id,
                   data = filter(all_candidates, fractile_thres <= 18))

all_elec <-felm(incumbent_future~ elected*female|0|0|cluster_id,
                  data = filter(all_candidates, fractile_thres <= 18))

## interaction table
stargazer(all_rerun, all_elec,
          keep = c("elected","Constant"),
          covariate.labels = c("Candidate Wins t=0",
                               "Candidate Wins t=0 X Female"),
          dep.var.labels = c("Rerunning t+1", "Rerun and Win t+1"),
          add.lines = list(c("Bandwidth", rep(18
                                              ,2))),
          df = FALSE, omit.stat = c( "adj.rsq", "rsq", "ser", "f"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/incumbency.tex",
          title = "No Gender Difference in Incumbency Advantage",
          label = "tab:incumbency")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:32:18

Table 9: No Gender Difference in Incumbency Advantage

	Dependent variable:	
	Rerunning t+1	Rerun and Win t+1
	(1)	(2)
Candidate Wins t=0	0.270***	0.327***
	(0.021)	(0.020)
Candidate Wins t=0 X Female	−0.039	−0.053*
	(0.036)	(0.032)
Constant	0.471***	0.156***
	(0.015)	(0.010)
Bandwidth	18	18
Observations	3,268	3,268
Note:		*p<0.1; **p<0.05; ***p<0.01

Table I2

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")

top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds") %>%
  mutate(cluster_id = as.factor(cluster_id))

top2_cand <- top2_cand %>%
  group_by(party, KOMMUNE, election_year, female_win) %>%
  summarise(fractile_thres_close = mean(fractile_thres, na.rm=T))

# find candidates who are elected beyond 18 bw -- safe candidates
safe_cand <-
  all_candidates %>%
  filter(fractile_thres > 18 & elected == 1) %>%
  left_join(., top2_cand) %>%
  mutate(frac_bw = case_when(female_win == 0 ~ fractile_thres_close*-1,
                             TRUE ~ fractile_thres))

safe_cand$cluster_id <- safe_cand %>%
  group_by(party, KOMMUNE, election_year) %>%
  group_indices()

safe_cand <-
  left_join(safe_cand, filter(safe_cand, female == 1) %>%
              group_by(cluster_id) %>%
              summarise(women_share_rerun = mean(running_next_elec))) %>%
  left_join(., filter(safe_cand, female == 0) %>%
              group_by(cluster_id) %>%
              summarise(men_share_rerun = mean(running_next_elec)))
```

```
safe_cand_agg <-
  safe_cand %>%
  group_by(cluster_id) %>%
  sample_n(1)


#### regressions


safe_rerun_fem <- felm(running_next_elec ~ female_win | 0|0|cluster_id,
                       data = filter(safe_cand, fractile_thres_close  <= 18 &
                                     female == 1))

safe_rerun_male <- felm(running_next_elec ~ female_win | 0|0|cluster_id,
                        data = filter(safe_cand, fractile_thres_close  <= 18 &
                                      female == 0))

safe_rerun_fem_agg <- felm(women_share_rerun ~ female_win | 0|0|KOMMUNE + party,
                           data = filter(safe_cand_agg, fractile_thres_close  <= 18))

safe_rerun_male_agg <- felm(men_share_rerun ~ female_win | 0|0|KOMMUNE + party,
                            data = filter(safe_cand_agg, fractile_thres_close  <= 18))

stargazer(safe_rerun_fem, safe_rerun_fem_agg, safe_rerun_male, safe_rerun_male_agg,
          covariate.labels = c("Woman Win t=0"),
          dep.var.caption = "Incumbents Rerunning t+1",
          dep.var.labels = c("Share Women",
                             "Share Women",
                             "Share Men",
                             "Share Men"),
          column.labels = rep(c("Individual level", "Candidate pool"), 2),
          add.lines = list(c("Bandwidth",
                             rep(18, 4))),
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          label = "tab:incumbents_rerun",
          #out = "E:/workdata/706687/projects/women_spearhead/plots/incumbents_rerunning.tex",
          title = "Electing women does not make other incumbents rerun more")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:32:25

Table 10: Electing women does not make other incumbents rerun more

	Incumbents Rerunning t+1			
	Share Women Individual level	Share Women Candidate pool	Share Men Individual level	Share Men Candidate pool
	(1)	(2)	(3)	(4)
Woman Win t=0	0.016	−0.002	−0.002	−0.011
	(0.037)	(0.026)	(0.019)	(0.016)
Constant	0.717***	0.719***	0.772***	0.785***
	(0.027)	(0.016)	(0.013)	(0.015)
Bandwidth	18	18	18	18
Observations	672	403	1,877	641
Note:				*p<0.1; **p<0.05; ***p<0.01

Appendix J

Table J1

```
top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")

new_fem <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  filter(first_run == 1 & female == 1) %>%
  summarise(tot_new_fem =  n())


new_men <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  filter(first_run == 1 & female == 0) %>%
  summarise(tot_new_men = n())

new <- full_join(new_men, new_fem, by = c("party", "KOMMUNE", "election_year"))  %>%
  # add 0 to NA, as NA occurs when there is no new men or women
  mutate(tot_new_men = ifelse(is.na(tot_new_men), 0, tot_new_men),
         tot_new_fem = ifelse(is.na(tot_new_fem), 0, tot_new_fem))

party_agg <- new %>%
  group_by(party, KOMMUNE) %>%
  mutate(tot_new = tot_new_fem+tot_new_men) %>%
  mutate(tot_new_fem_t1 = dplyr::lead(tot_new_fem, n = 1, order_by = election_year),
         tot_new_men_t1 = dplyr::lead(tot_new_men, n = 1, order_by = election_year),
         tot_new_t1 = dplyr::lead(tot_new, n = 1, order_by = election_year))

top2_cand <- left_join(top2_cand, party_agg,
```

```
                      by = c("party", "KOMMUNE", "election_year"))


top2_cand$cluster_id <- top2_cand %>%
  group_by(party, KOMMUNE) %>%
  group_indices()


top2_cand$muni_id <- top2_cand %>%
  group_by(KOMMUNE) %>%
  group_indices()

#----------------------
# Use bandwidths from main specification

mod_new <- felm(tot_new_t1 ~ female_win | 0 |0|KOMMUNE + party,
                    data =  top2_cand, subset = fractile_thres <= 18)

mod_new_fem <- felm(tot_new_fem_t1 ~ female_win | 0 |0|KOMMUNE + party,
                    data =  top2_cand, subset = fractile_thres <= 18)

mod_new_male <- felm(tot_new_men_t1 ~ female_win | 0 |0|KOMMUNE + party,
                     data =  top2_cand, subset = fractile_thres <= 18)


stargazer(mod_new, mod_new_fem, mod_new_male,
          covariate.labels = c("Woman Win t=0"),
          dep.var.caption = "New Candidtes t+1",
          dep.var.labels = c("In total", "Women", "Men"),
          add.lines = list(c("Bandwidth",18, 18, 18)),
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/new_candidates.tex",
          title = "The Effect of Electing Women on Future Candidate Emergence",
          label = "tab:new_candidates")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:32:29

Table 11: The Effect of Electing Women on Future Candidate Emergence

	New Candidtes t+1		
	In total	Women	Men
	(1)	(2)	(3)
Woman Win t=0	−0.012	−0.035	0.023
	(0.148)	(0.051)	(0.129)
Constant	5.657***	1.770***	3.887***
	(0.504)	(0.125)	(0.403)
Bandwidth	18	18	18
Observations	806	806	806
Note:		*p<0.1; **p<0.05; ***p<0.01	

Appendix K

Figure K1 and Table K1

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")


top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

party_agg <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/party_agg_w_outcomes.rds")


## define left/right wing


left_parties <- c("A", "B", "F", "Ø")
right_parties <- c("C", "O", "V")

all_candidates <- all_candidates %>%
  mutate(lr = case_when(party %in% left_parties ~ 1,
                        PARTI == "Å Alternativet" ~ 1,
                        party %in% right_parties ~ 0,
                        PARTI == "I Liberal Alliance" ~ 0,
                        PARTI == "K Kristendemokraterne" ~ 0,
                        TRUE ~ NA_real_))

# proportion of women in election-party-municipality groups
party_agg <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(prop_incumbents = mean(incumbent, na.rm = T),
            lr = mean(lr, na.rm = TRUE))
```

```r
# covariates at the municipality level
muni_cov <- all_candidates %>%
  group_by(KOMMUNE, election_year) %>%
  filter(elected == 1) %>%
  summarise(council_size = n(),
            fem_in_muni = mean(female,na.rm = TRUE))

top2_cand <- left_join(top2_cand, muni_cov,
                       by = c("KOMMUNE", "election_year"))

top2_cand <- left_join(top2_cand, party_agg,
                       by = c("party", "KOMMUNE", "election_year"))


# gap btwn party and ideological camp
top2_cand <- top2_cand %>%
  group_by(KOMMUNE, lr, election_year) %>%
  summarise(ideo_gender = mean(prop_female, na.rm=T)) %>%
  right_join(top2_cand, by = c("KOMMUNE", "lr", "election_year"))

top2_cand$party_ideo_gap <- top2_cand$prop_female - top2_cand$ideo_gender

prop_mod <- felm(prop_female_t1 ~ female_win*I(party_ideo_gap*100)|0|0|KOMMUNE + party,
                 data = filter(top2_cand, fractile_thres <= 18))

fem_mod <- felm(log(fem_vote_t1+1) ~ female_win*I(party_ideo_gap*100)|0|0|KOMMUNE + party,
                data = filter(top2_cand, fractile_thres <= 18))

male_mod <- felm(log(male_vote_t1) ~ female_win*I(party_ideo_gap*100)|0|0|KOMMUNE + party,
                 data = filter(top2_cand, fractile_thres <= 18))

party_mod <- felm(log(party_vote_t1)~ female_win*I(party_ideo_gap*100)|0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 18))

seat_mod <- felm(n_elect_t1 ~ female_win*I(party_ideo_gap*100)|0|0|KOMMUNE + party,
                 data = filter(top2_cand, fractile_thres <= 18))


stargazer(prop_mod, fem_mod, male_mod, party_mod, seat_mod,
          covariate.labels = c("Woman win t0", "Women relative to bloc t0",
                               "Woman win t0 X Women relative to bloc t0"),
          dep.var.labels = c(" Share Women Elected", "Log Votes Women", "Log Votes Men",
                             "Log Votes Party", "Seats Party"),
          add.lines = list(c("Bandwidth",
                             rep(18, 5))),
          df = FALSE, keep.stat = "n",
          font.size = "scriptsize",
          #out = "E:/workdata/706687/projects/women_spearhead/plots/PercWomen_V_IdeoCampInteraction.tex
          title = "The Effect of Electing Women is Moderated by the Prior Representation of Women in th
          label = "tab:perc_women_bloc_interaction")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:32:58

Table 12: The Effect of Electing Women is Moderated by the Prior Representation of Women in the party's ideological bloc.

	Dependent variable:				
	Share Women Elected	Log Votes Women	Log Votes Men	Log Votes Party	Seats Party
	(1)	(2)	(3)	(4)	(5)
Woman win t0	0.088***	0.407***	0.094	0.162***	0.473**
	(0.028)	(0.078)	(0.078)	(0.061)	(0.194)
Women relative to bloc t0	0.002**	0.012***	0.010***	0.009***	0.032**
	(0.001)	(0.003)	(0.002)	(0.003)	(0.015)
Woman win t0 X Women relative to bloc t0	0.001	−0.016***	−0.032***	−0.022***	−0.094***
	(0.001)	(0.005)	(0.006)	(0.005)	(0.015)
Constant	0.245***	5.834***	7.358***	7.519***	4.673***
	(0.016)	(0.244)	(0.257)	(0.199)	(0.973)
Bandwidth	18	18	18	18	18
Observations	819	819	819	819	819
Note:					*p<0.1; **p<0.05; ***p<0.01

```
# create graphs of the models
top2_cand$ln_male_vote <- log(top2_cand$male_vote_t1+1)
top2_cand$ln_fem_vote <- log(top2_cand$fem_vote_t1+1)
top2_cand$ln_party_vote <- log(top2_cand$party_vote_t1+1)


# subset to close elections
close_sub <- filter(top2_cand, fractile_thres <= 18) %>%
  as.data.frame() %>%
  drop_na(female_win)


# estimate interflex models
flex_prop <- interflex(estimator = "binning", data = close_sub, cl = "cluster_id",nbins=5,
                       full.moderate = FALSE,
                       Y = "prop_female_t1", D = "female_win", X = "party_ideo_gap", na.rm=T)
```

Baseline group not specified; choose treat = 0 as the baseline group.

```
flex_fem <- interflex(estimator = "binning", data = close_sub, cl = "cluster_id",nbins=5,
                      Y = "ln_fem_vote", D = "female_win", X = "party_ideo_gap", na.rm=T)
```

Baseline group not specified; choose treat = 0 as the baseline group.

```
flex_male <- interflex(estimator = "binning", data = close_sub, cl = "cluster_id",nbins=5,
                       Y = "ln_male_vote", D = "female_win", X = "party_ideo_gap", na.rm=T)
```

Baseline group not specified; choose treat = 0 as the baseline group.

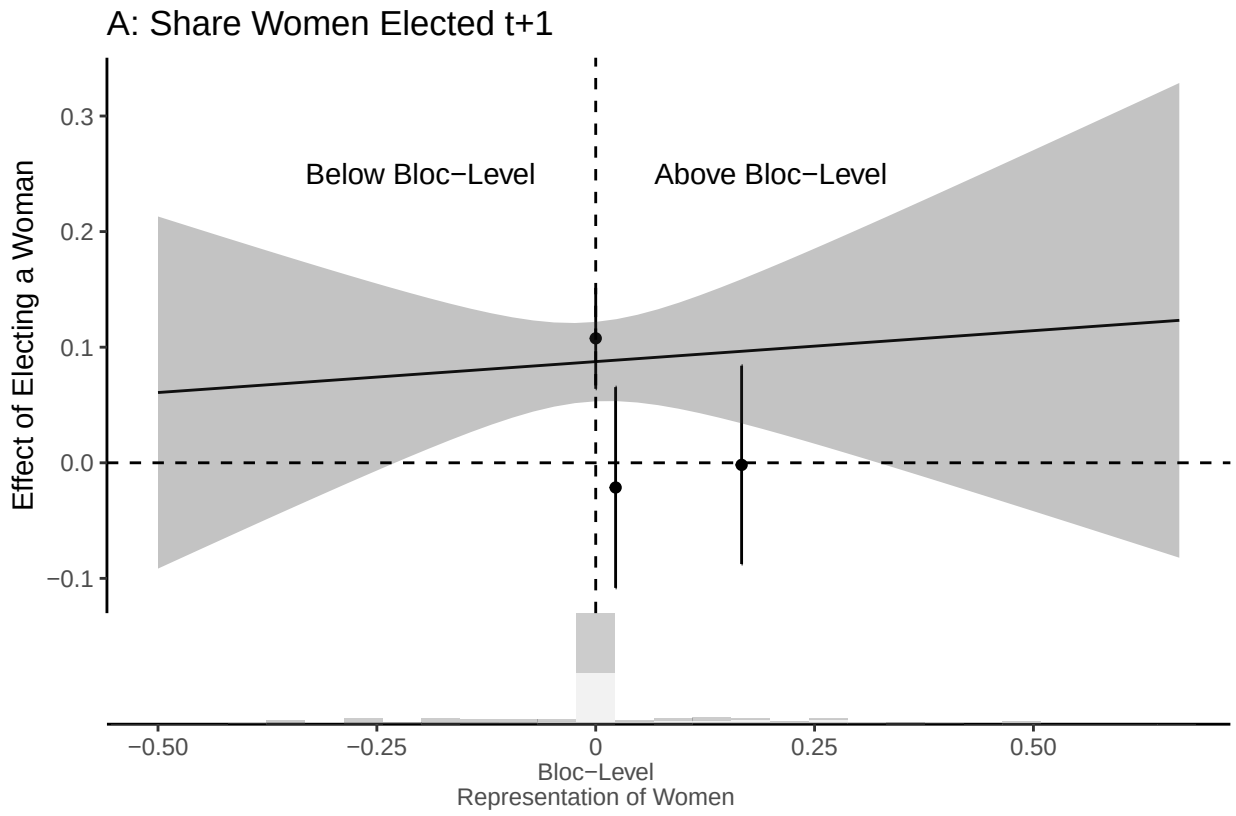
```
flex_party <- interflex(estimator = "binning", data = close_sub, cl = "cluster_id",nbins=5,
                        Y = "ln_party_vote", D = "female_win", X = "party_ideo_gap", na.rm=T)
```

Baseline group not specified; choose treat = 0 as the baseline group.

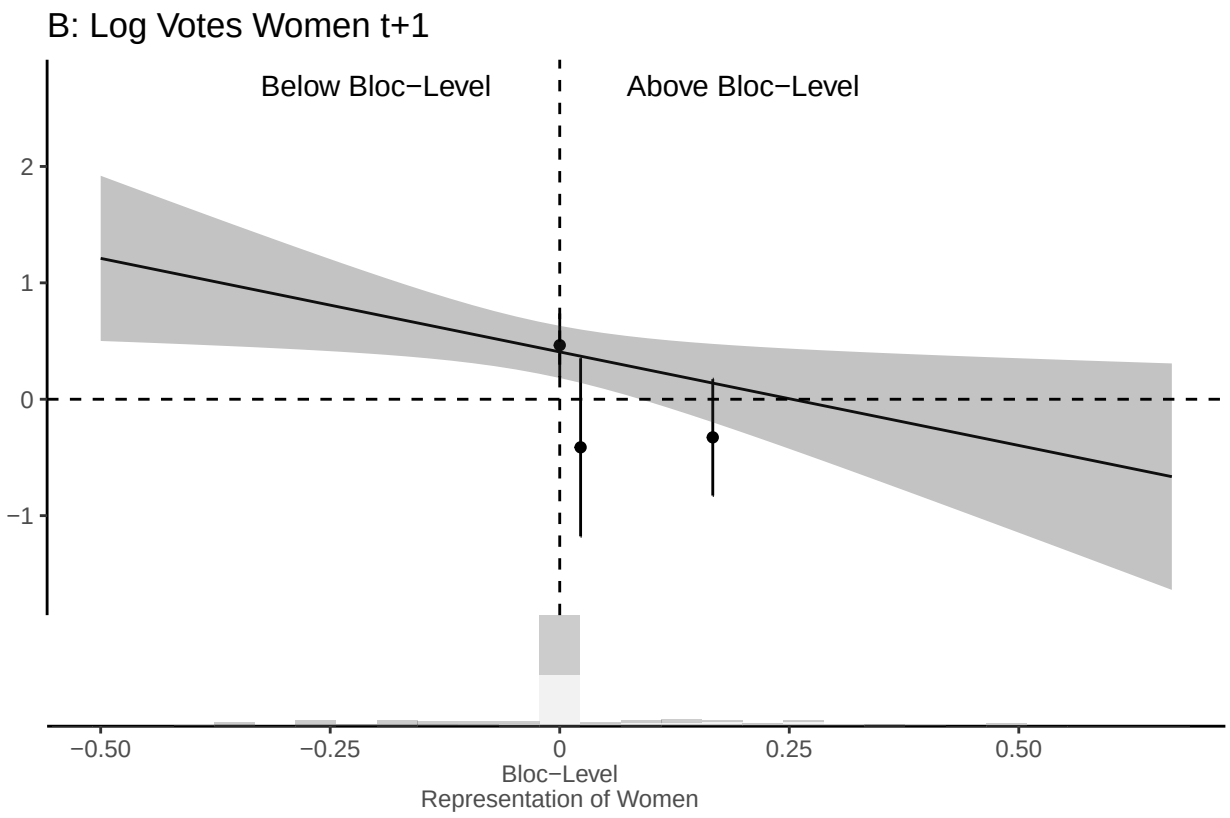
```
flex_seat <- interflex(estimator = "binning", data = close_sub, cl = "cluster_id",nbins=5,
                       Y = "n_elect_t1", D = "female_win", X = "party_ideo_gap", na.rm=T)
```

Baseline group not specified; choose treat = 0 as the baseline group.

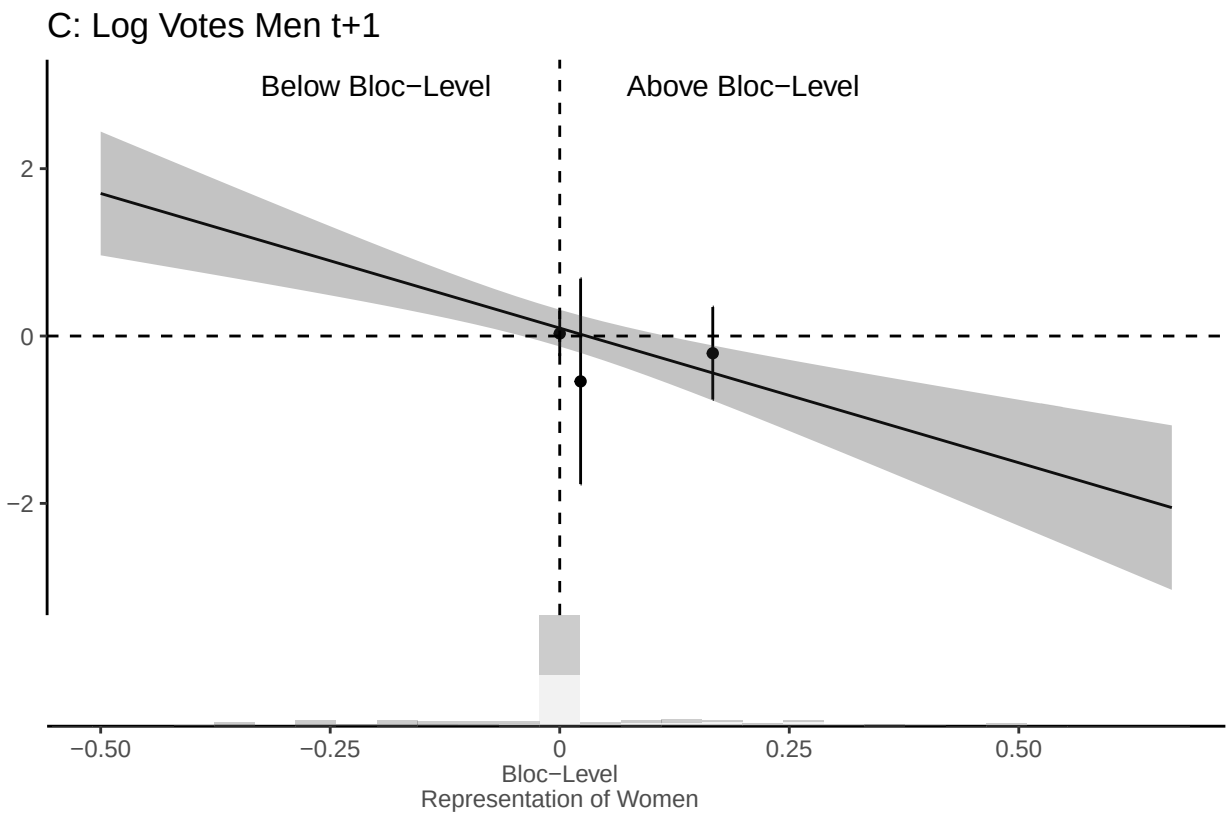
```r
# proportion

flex_plot <- function(flex_est, plot_ylab, plot_title, xcoord1, xcoord2,
                      ycoord1, ycoord2){

  prop_est <- as.data.frame(flex_est$est.lin)
  names(prop_est) <- c("mod", "TE", "sd", "lower", "upper")
  prop_est_bin <- as.data.frame(flex_est$est.bin)
  names(prop_est_bin) <- c("bin", "TE", "sd", "lower", "upper")


  p_prop <- ggplot(prop_est, aes(x = mod, y = TE)) +
    geom_line() +
    geom_point(data = prop_est_bin,
               aes(x = bin, y = TE), inherit.aes = FALSE ) +
    geom_errorbar(data = prop_est_bin,
                  aes(x = bin, y = TE,
                      ymin = lower, ymax = upper), inherit.aes = FALSE, width = 0) +
    theme_classic() +
    geom_ribbon(aes(ymin = lower, ymax = upper), alpha = .3) +
    geom_hline(yintercept = 0, lty = 2) +
    geom_vline(xintercept = 0, lty = 2) +
    labs(x = "",
         y = plot_ylab,
         title = plot_title) +
    annotate(geom = "text", x = xcoord1, y = ycoord1,
             label = "Below Bloc-Level")+
    annotate(geom = "text", x = xcoord2, y = ycoord2,
             label = "Above Bloc-Level") +
    scale_x_continuous(breaks = seq(-.5,.5,by = 0.25),
                       labels = c("-0.50", "-0.25", "0\nBloc-Level\nRepresentation of Women", "0.25", "


  x_hist_prop <- axis_canvas(p_prop, axis = "x") +
    geom_histogram(data = close_sub,
                   aes(x = party_ideo_gap, fill = factor(female_win)), alpha = .2) +
    scale_fill_manual(values = c("black", "grey"))


  p_prop_comb <- insert_xaxis_grob(p_prop, x_hist_prop, position = "bottom")

  return(p_prop_comb)

}


prop_plot <- flex_plot(flex_est = flex_prop, plot_ylab = "Effect of Electing a Woman",
                       plot_title = "A: Share Women Elected t+1",
                       xcoord1 = -0.2, ycoord1 = 0.25,
                       xcoord2 = 0.2, ycoord2 = 0.25)

ggdraw(prop_plot)
```



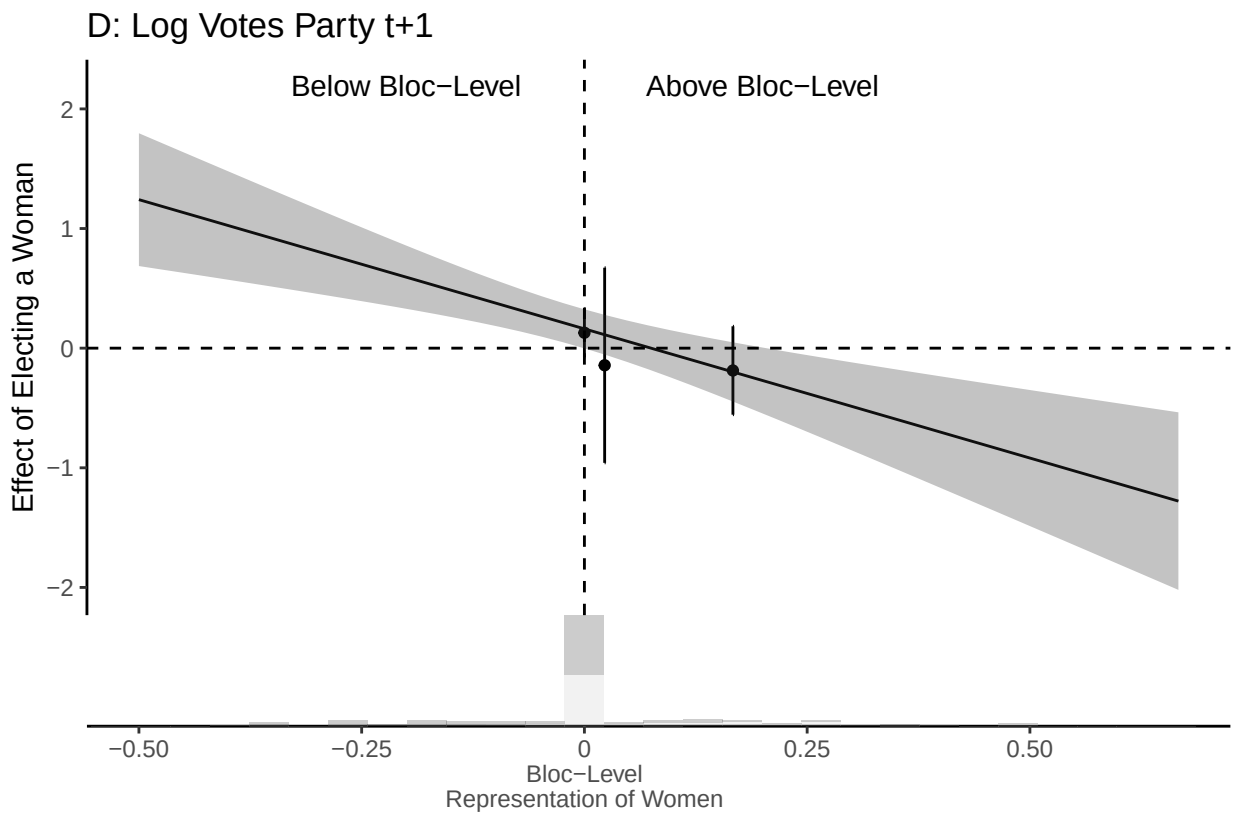

```
fem_vote_plot <- flex_plot(flex_est = flex_fem, plot_ylab = NULL,
                           plot_title = "B: Log Votes Women t+1",
                           xcoord1 = -0.2, ycoord1 = 2.7,
                           xcoord2 = 0.2, ycoord2 = 2.7)

ggdraw(fem_vote_plot)
```



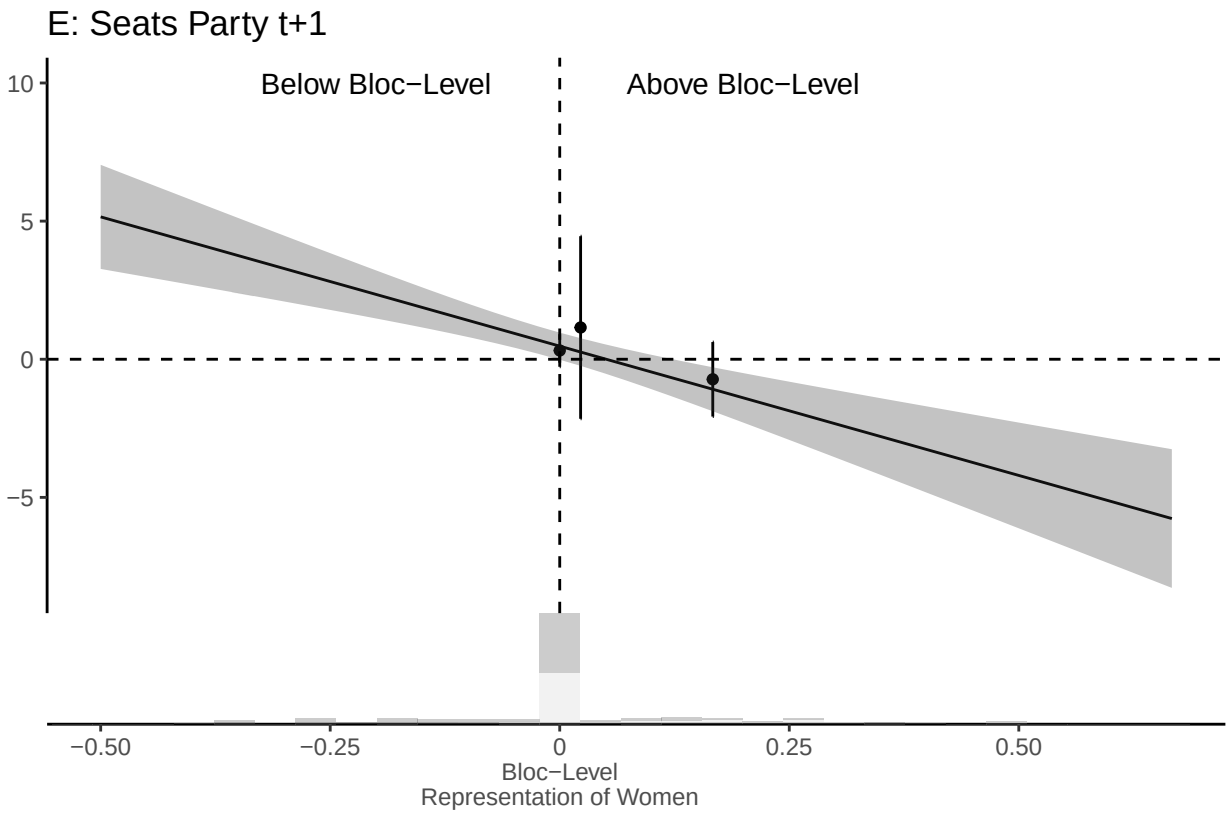

```
male_vote_plot <- flex_plot(flex_est = flex_male, plot_ylab = NULL,
                            plot_title = "C: Log Votes Men t+1",
                            xcoord1 = -0.2, ycoord1 = 3,
                            xcoord2 = 0.2, ycoord2 = 3)

ggdraw(male_vote_plot)
```




```
party_vote_plot <- flex_plot(flex_est = flex_party, plot_ylab = "Effect of Electing a Woman",
                             plot_title = "D: Log Votes Party t+1",
                             xcoord1 = -0.2, ycoord1 = 2.2,
                             xcoord2 = 0.2, ycoord2 = 2.2)

ggdraw(party_vote_plot)
```




```
seat_vote_plot <- flex_plot(flex_est = flex_seat, plot_ylab = NULL,
                            plot_title = "E: Seats Party t+1",
                            xcoord1 = -0.2, ycoord1 = 10,
                            xcoord2 = 0.2, ycoord2 = 10)

ggdraw(seat_vote_plot)
```


E: Seats Party t+1
Below Bloc−Level
Above Bloc−Level
10
5
0
−5
−0.50
−0.25
0
0.25
0.50
Bloc−Level
Representation of Women

Appendix L

Table L1

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")


top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

party_agg <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/party_agg_w_outcomes.rds")

## municipality size moderation


top2_cand <- top2_cand %>%
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  # # dummy for big cities: copenhagen, frederiksberg, aarhus, odense and aalborg
  mutate(big_city = ifelse(kom_nr %in% c(101,147,751,461,851),1,0))


######## estimate models - copenhagen, aarhus, odense and aalborg

fem_elect <- felm(prop_female_t1 ~ female_win*big_city |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 18))

fem_vote <- felm(log(fem_vote_t1+1) ~ female_win*big_city |0|0|KOMMUNE + party,
                 data = filter(top2_cand, fractile_thres <= 18))

male_vote <- felm(log(male_vote_t1) ~ female_win*big_city |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 18))

party_vote_mod <- felm(log(party_vote_t1) ~ female_win*big_city |0|0|KOMMUNE + party,
                       data = filter(top2_cand, fractile_thres <= 18))

mandates_mod <- felm(n_elect_t1 ~ female_win*big_city |0|0|KOMMUNE + party,
                     data = filter(top2_cand, fractile_thres <= 18))

stargazer(fem_elect, fem_vote, male_vote, party_vote_mod, mandates_mod,
          covariate.labels = c("Woman win t=0"),
          dep.var.labels = c(" Share Women Elected", "Log Votes Women", "Log Votes Men",
                             "Log Votes Party", "Seats Party"),
          add.lines = list(c("Bandwidth",
              rep(18, 5))),
          font.size = "scriptsize",
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/estimates_big_cities.tex",
          title = "Estimates when interacting for big cities",
          label = "tab:big_cities")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: to, aug 15, 2024 - 17:09:31

Table 13: Estimates when interacting for big cities

	Dependent variable:				
	Share Women Elected	Log Votes Women	Log Votes Men	Log Votes Party	Seats Party
	(1)	(2)	(3)	(4)	(5)
Woman win t=0	0.105***	0.402***	0.030	0.121**	0.245
	(0.033)	(0.075)	(0.089)	(0.060)	(0.151)
big_city	0.002	1.813***	1.558***	1.454***	−0.066
	(0.050)	(0.230)	(0.486)	(0.246)	(0.852)
female_win:big_city	0.032***	0.804***	0.232	0.475***	1.421
	(0.007)	(0.122)	(0.345)	(0.170)	(0.872)
Constant	0.238***	5.699***	7.242***	7.414***	4.566***
	(0.019)	(0.268)	(0.286)	(0.223)	(1.025)
Bandwidth	18	18	18	18	18
Observations	819	819	819	819	819
Note:					*p<0.1; **p<0.05; ***p<0.01

Appendix M

Table M1

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")


top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

party_agg <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/party_agg_w_outcomes.rds")



### aggregate party votes pr. municipality year


 muni_vote <- party_agg %>%
   group_by(KOMMUNE, election_year) %>%
   summarise(muni_vote = sum(fem_vote+male_vote, na.rm = T),
             muni_seat = sum(n_elect, na.rm=T)) %>%
   right_join(party_agg)

 muni_vote <- muni_vote %>%
   group_by(party, election_year) %>%
   arrange(election_year) %>%
   mutate(muni_vote_lag = dplyr::lag(muni_vote, n = 1, order_by = election_year),
          fem_vote_lag = dplyr::lag(fem_vote, n = 1, order_by = election_year),
          male_vote_lag = dplyr::lag(male_vote, n = 1, order_by = election_year),
          party_vote_lag = dplyr::lag(party_vote, n = 1, order_by = election_year),
          n_elect = ifelse(is.na(n_elect) == T, 0, n_elect),
          party_vote_share = party_vote/muni_vote_lag,
          fem_vote_share = fem_vote/muni_vote_lag,
          male_vote_share = male_vote/muni_vote_lag,
```

```
          seat_share = n_elect/muni_seat) %>%
   mutate(party_vote_share_t1 = dplyr::lead(party_vote_share, n = 1, order_by = election_year),
          fem_vote_share_t1 = dplyr::lead(fem_vote_share, n = 1, order_by = election_year),
          male_vote_share_t1 = dplyr::lead(male_vote_share, n = 1, order_by = election_year),
          seat_share_t1 = dplyr::lead(seat_share, n = 1, order_by = election_year)) %>%
   select(KOMMUNE, party, election_year,
          muni_vote, muni_seat,
          muni_vote_lag,
          fem_vote_lag,
          male_vote_lag,
          party_vote_lag,
          party_vote_share,
          fem_vote_share,
          male_vote_share,
          seat_share = n_elect/muni_seat,
          party_vote_share_t1,
          fem_vote_share_t1,
          male_vote_share_t1,
          seat_share_t1)

 top2_cand <- left_join(top2_cand, muni_vote, by = c("KOMMUNE", "party", "election_year"))

######## estimate models

 fem_vote <- felm(I(fem_vote_t1/muni_vote) ~ female_win |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 22))

 male_vote <- felm(I(male_vote_t1/muni_vote) ~ female_win |0|0|KOMMUNE + party,
                   data = filter(top2_cand, fractile_thres <= 22))

 party_vote_mod <- felm(I(party_vote_t1/muni_vote) ~ female_win |0|0|KOMMUNE + party,
                        data = filter(top2_cand, fractile_thres <= 22 ))

 party_seat_mod <- felm(I(n_elect_t1/muni_seat) ~ female_win |0|0|KOMMUNE + party,
                        data = filter(top2_cand, fractile_thres <= 22))

stargazer(fem_vote, male_vote, party_vote_mod, party_seat_mod,
          covariate.labels = c("Woman win t=0"),
          dep.var.labels = c("Vote Share Women", "Vote Share Men",
                             "Vote Share Party", "Seat Share Party"),
          add.lines = list(c("Bandwidth", rep(18, 5))),
          font.size = "scriptsize",
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/estimates_voteshares.tex",
          title = "Effects on Vote and Seat Shares in Municipality",
          label = "tab:votes_shares")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:33:10

Table 14: Effects on Vote and Seat Shares in Municipality

	Dependent variable:			
	Vote Share Women	Vote Share Men	Vote Share Party	Seat Share Party
	(1)	(2)	(3)	(4)
Woman win t=0	0.014**	0.024**	0.007	0.010
	(0.007)	(0.012)	(0.008)	(0.007)
Constant	0.045***	0.179***	0.156***	0.203***
	(0.009)	(0.056)	(0.038)	(0.049)
Bandwidth	18	18	18	18
Observations	889	889	889	889
Note:				*p<0.1; **p<0.05; ***p<0.01

Table M2

```
fem_vote <- felm(log(fem_vote_t1+1) ~ female_win + log(muni_vote) |0|0|KOMMUNE + party,
                 data = filter(top2_cand, fractile_thres <= 22))

male_vote <- felm(log(male_vote_t1 +1 )~ female_win + log(muni_vote) |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 22))

party_vote_mod <- felm(log(party_vote_t1) ~ female_win + log(muni_vote) |0|0|KOMMUNE + party,
                       data = filter(top2_cand, fractile_thres <= 22 ))

party_seat_mod <- felm(n_elect_t1 ~ female_win +muni_seat |0|0|KOMMUNE + party,
                       data = filter(top2_cand, fractile_thres <= 22))

stargazer(fem_vote, male_vote, party_vote_mod, party_seat_mod,
          covariate.labels = c("Woman win t=0", "log(Total Votes in Municipality)",
                               "Total Seats in Municipality"),
          dep.var.labels = c("log(Vote Women+1)", "log(Vote Share Men+1)",
                             "log(Vote Party+1)", "Party Seats"),
          add.lines = list(c("Bandwidth", rep(18, 5))),
          font.size = "scriptsize",
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/estimates_voteshares_controlling.te
          title = "Controlling for the Total Numbers of Votes and Seats in Municipality",
          label = "tab:votes_controlling")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: on, aug 14, 2024 - 20:33:12

Table 15: Controlling for the Total Numbers of Votes and Seats in Municipality

	Dependent variable:			
	log(Vote Women+1)	log(Vote Share Men+1)	log(Vote Party+1)	Party Seats
	(1)	(2)	(3)	(4)
Woman win t=0	0.459^{***}	0.029	0.142^{**}	0.344^{**}
	(0.088)	(0.056)	(0.061)	(0.143)
log(Total Votes in Municipality)	0.666^{***}	0.606^{***}	0.552^{***}	
	(0.053)	(0.075)	(0.054)	
Total Seats in Municipality				0.088
				(0.056)
Constant	-0.789^{*}	1.286^{**}	2.003^{***}	2.326^{***}
	(0.438)	(0.605)	(0.435)	(0.567)
Bandwidth	18	18	18	18
Observations	889	889	889	889
Note:				*p<0.1; **p<0.05; ***p<0.01

Appendix N

Figure N1

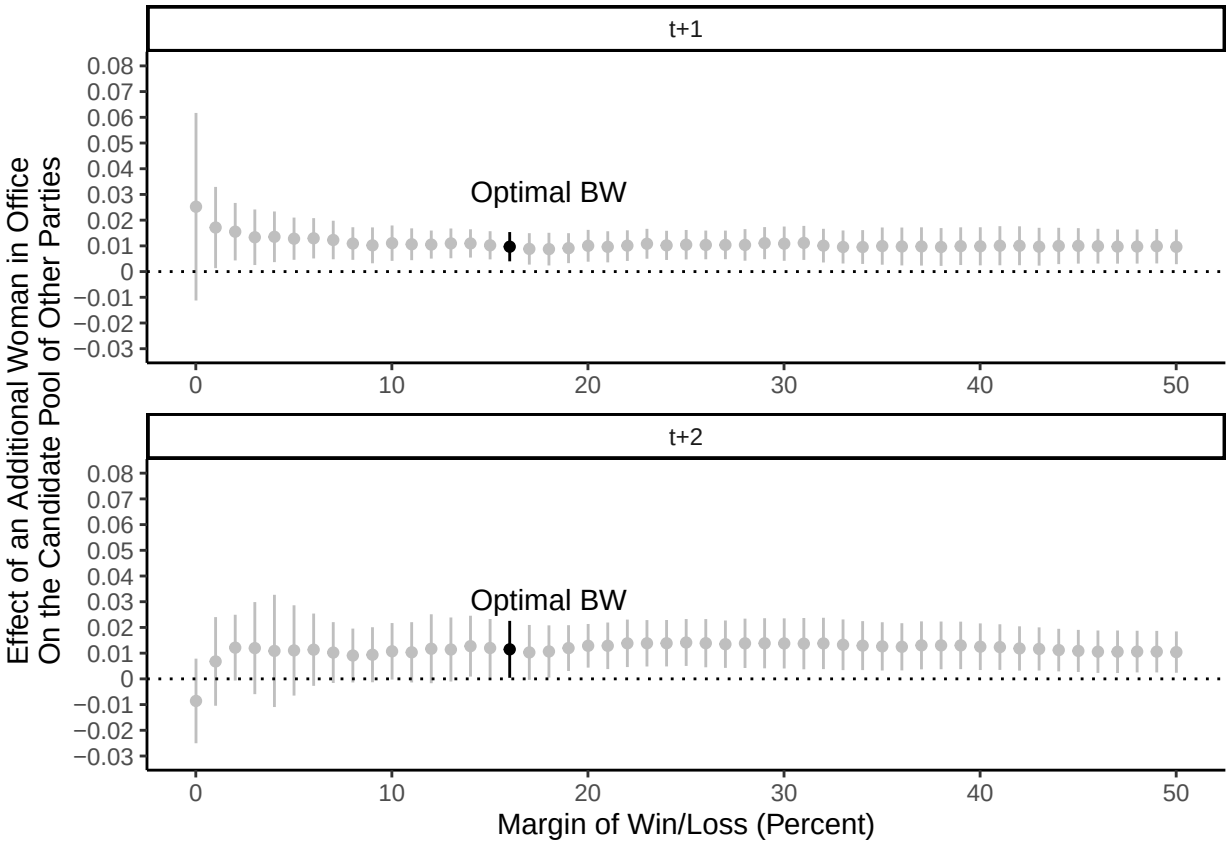
```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")

# function for mapping over all bandwidths

# proportion of women in election-party-municipality groups
party_agg <- all_candidates %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(female_candidates = mean(female, na.rm = T))

party_agg2 <- all_candidates %>%
  filter(elected == 1) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(prop_female = mean(female, na.rm = T))


party_agg <- left_join(party_agg, party_agg2)

party_agg <- party_agg %>%
  group_by(party, KOMMUNE) %>%
  mutate(prop_female_t1 = dplyr::lead(prop_female, n = 1, order_by = election_year),
         prop_female_t2 = dplyr::lead(prop_female, n = 2, order_by = election_year),
         female_candidates_t1 = dplyr::lead(female_candidates, n = 1, order_by = election_year),
         female_candidates_t2 = dplyr::lead(female_candidates, n = 2, order_by = election_year)) %>%
  arrange(party, KOMMUNE, election_year)

#-------------------------------------------------------------
# Analyzing spillovers to other parties

muni_other_party_func <- function(this_muni){

  crnt_muni <- filter(party_agg, KOMMUNE == this_muni)
```

```
  other_party_func <- function(this_party){

    other_party <- filter(crnt_muni, party != this_party)

    other_party <- other_party %>%
      group_by(election_year) %>%
      summarise(spillover_t1 = mean(female_candidates_t1, na.rm = T),
                spillover_t2 = mean(female_candidates_t2, na.rm = T))

    other_party <- data.frame(party = this_party, other_party)

    return(other_party)

  }

  other_party_ave <- map(unique(crnt_muni$party),
                         ~ other_party_func(.))

  other_party_ave <- do.call("rbind", other_party_ave)

  crnt_muni <- left_join(crnt_muni, other_party_ave, by = c("party", "election_year"))

  return(crnt_muni)

}


spillover_est <- map(unique(party_agg$KOMMUNE),
                     ~ muni_other_party_func(.))
spillover_est <- do.call("rbind", spillover_est)



# estimate models on other parties t1
spillover_all_bw_t1 <- function(bw, other_parties){

  top2_cand <- all_candidates %>%
    filter(OPSTILLINGSFORM == "Sideordnet") %>%
    group_by(party, KOMMUNE, election_year) %>%
    arrange(fractile_thres) %>%
    slice(1:2) %>%
    arrange(party, KOMMUNE, election_year) %>%
    filter(fractile_thres < bw)

  # create indicators of female win
  top2_cand <- top2_cand %>%
    mutate(female_win = case_when(female == 1 & elected == 1 ~ 1,
                                  TRUE ~ 0)) %>%
    group_by(party, KOMMUNE, election_year) %>%
    summarise(female_candidate = mean(female, na.rm = T),
              female_win = mean(female_win, na.rm = T))%>%
    mutate(female_candidate = case_when(female_candidate == .50 ~ 1,
                                        TRUE ~ 0),
```

```
          female_win = case_when(female_win > 0 ~ 1,
                                 TRUE ~ 0))

  # merge in female proportions
  top2_cand <- left_join(top2_cand, other_parties,
                         by = c("party", "KOMMUNE", "election_year"))

  # party-municipality id
  top2_cand$muni_party <- top2_cand %>%
    group_by(party, KOMMUNE) %>%
    group_indices()

  # period_id
  top2_cand <- top2_cand %>%
    mutate(period = case_when(election_year == 1993 ~ 1,
                              election_year == 1997 ~ 2,
                              election_year == 2001 ~ 3,
                              election_year == 2005 ~ 4,
                              election_year == 2009 ~ 5,
                              election_year == 2013 ~ 6,
                              election_year == 2017 ~ 7))


  pool_red_t1 <- felm(spillover_t1 ~ female_win |0|0|KOMMUNE + party, data = top2_cand)

  reduced_form_t1 <- data.frame(PE = coef(pool_red_t1)[2],
                                SE = sqrt(diag(vcov(pool_red_t1)))[2],
                                df = pool_red_t1$df,
                                t = 1,
                                f_stat = NA,
                                bw = bw)

  return(rbind(reduced_form_t1))

}

## spill over estimation for t2
spillover_all_bw_t2 <- function(bw, other_parties){

  top2_cand <- all_candidates %>%
    filter(OPSTILLINGSFORM == "Sideordnet") %>%
    group_by(party, KOMMUNE, election_year) %>%
    arrange(fractile_thres) %>%
    slice(1:2) %>%
    arrange(party, KOMMUNE, election_year) %>%
    filter(fractile_thres < bw)

  # create indicators of female win
  top2_cand <- top2_cand %>%
    mutate(female_win = case_when(female == 1 & elected == 1 ~ 1,
                                  TRUE ~ 0)) %>%
    group_by(party, KOMMUNE, election_year) %>%
    summarise(female_candidate = mean(female, na.rm = T),
```

```
              female_win = mean(female_win, na.rm = T))%>%
    mutate(female_candidate = case_when(female_candidate == .50 ~ 1,
                                        TRUE ~ 0),
           female_win = case_when(female_win > 0 ~ 1,
                                  TRUE ~ 0))

  # merge in female proportions
  top2_cand <- left_join(top2_cand, other_parties,
                         by = c("party", "KOMMUNE", "election_year"))

  # party-municipality id
  top2_cand$muni_party <- top2_cand %>%
    group_by(party, KOMMUNE) %>%
    group_indices()

  # period_id
  top2_cand <- top2_cand %>%
    mutate(period = case_when(election_year == 1993 ~ 1,
                              election_year == 1997 ~ 2,
                              election_year == 2001 ~ 3,
                              election_year == 2005 ~ 4,
                              election_year == 2009 ~ 5,
                              election_year == 2013 ~ 6,
                              election_year == 2017 ~ 7))


  pool_red_t2 <- felm(spillover_t2 ~ female_win |0|0|KOMMUNE + party, data = top2_cand)


  reduced_form_t2 <- data.frame(PE = coef(pool_red_t2)[2],
                                SE = sqrt(diag(vcov(pool_red_t2)))[2],
                                df = pool_red_t2$df,
                                t = 2,
                                f_stat = NA,
                                bw = bw)

  return(rbind(reduced_form_t2))

}




spillover_res_t1 <- map(seq(0.01, 50.1, 1),
                        ~ spillover_all_bw_t1(bw = ., other_parties = spillover_est))

spillover_df_t1 <- do.call("rbind", spillover_res_t1)

spillover_res_t2 <- map(seq(0.01, 50.1, 1),
                        ~ spillover_all_bw_t2(bw = ., other_parties = spillover_est))

spillover_df_t2 <- do.call("rbind", spillover_res_t2)
```

```r
spillover_df <- bind_rows(spillover_df_t1, spillover_df_t2)

spillover_df <- spillover_df %>%
  mutate(optimal = factor(ifelse(bw==16.01,1,0))) %>%
  mutate(t = ifelse(t==1,"t+1", "t+2"))

ggplot(spillover_df, aes(x = bw, y = PE, color = optimal)) +
  geom_point() +
  geom_linerange(aes(ymin = PE - 1.96*SE,
                     ymax = PE + 1.96*SE), width = 0) +
  theme_classic() +
  geom_hline(yintercept = 0, lty = 3) +
  scale_y_continuous(breaks = seq(-0.03,0.08,0.01), labels = seq(-0.03,0.08,0.01)) +
  coord_cartesian(ylim=c(-0.03,0.08)) +
  labs(x = "Margin of Win/Loss (Percent)",
       y = "Effect of an Additional Woman in Office\nOn the Candidate Pool of Other Parties") +
  scale_color_manual("", values = c("gray75", "black")) +
  theme(legend.position = "none") +
  facet_wrap(~t, nrow = 2, scales = "free") +
  annotate(geom = "text", label = "Optimal BW", x = 18, y = 0.031)
```

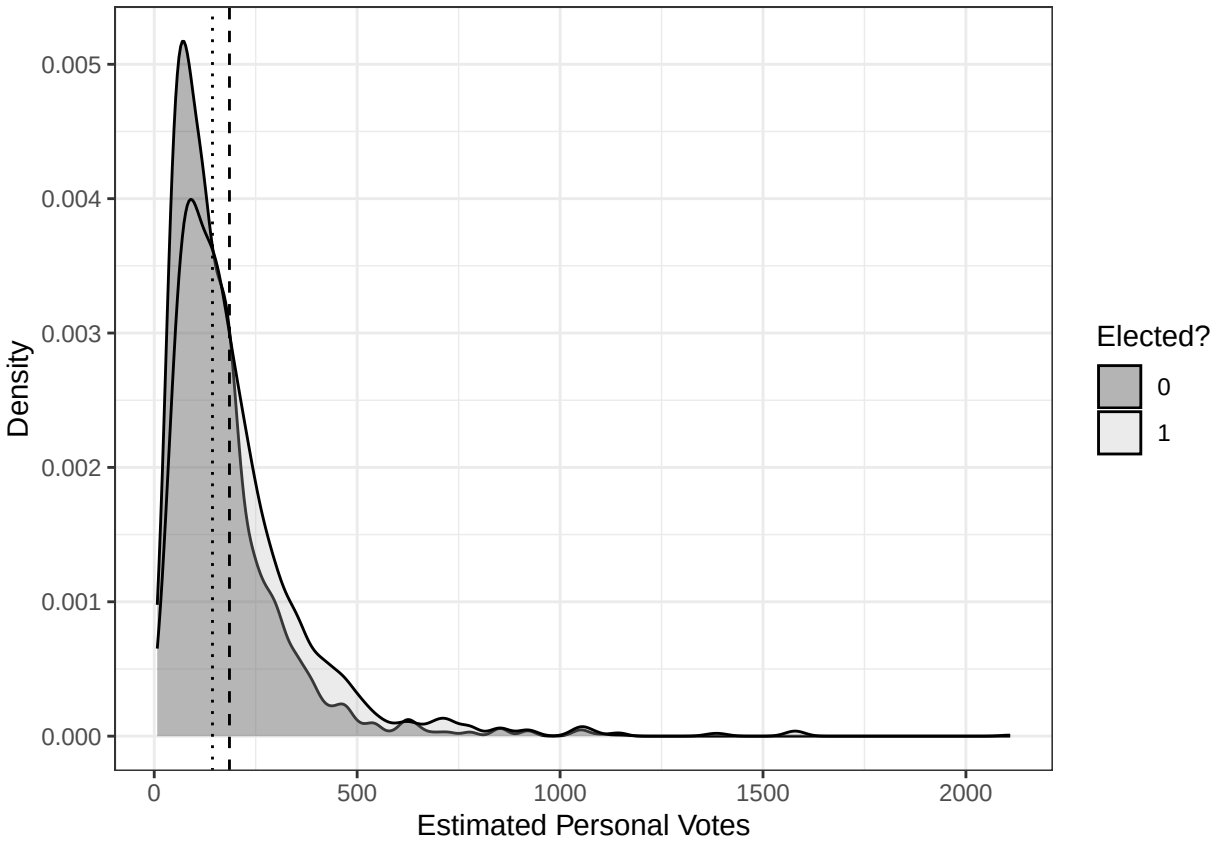



```r
#ggsave(paste0(work_data, "projects/women_spearhead/plots/spillover_optimal_bw.pdf"), width = 6, height
```

Appendix O

Figure O1

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")

#Estimate of personal votes
all_candidates$interval_points <- gsub(pattern =  "\\[", replacement = "",
                                       x = all_candidates$interval_points)
all_candidates$interval_points <- gsub(pattern =  "\\]", replacement = "",
                                       x = all_candidates$interval_points)

all_candidates <- all_candidates %>%
  separate(col = "interval_points", into = c("lwr", "upr"), sep = ",") %>%
  mutate(lwr = as.numeric(as.character(lwr)),
         upr = as.numeric(as.character(upr)),
         mid_pers_vote = (lwr+upr)/2)

# define vector of non-amalgated municipalities
non_amalgated_vector <- c(165, 201, 151, 153, 155, 563, 607, 147, 157, 159, 161, 253, 217, 163, 167, 169
                          673, 173, 825, 773, 727, 461, 329, 175, 741, 269, 185, 187, 751)

'%!in%' <- function(x,y)!('%in%'(x,y))

# define variable to sort amalgated municipalities away as outcome year in 2005
all_candidates <- all_candidates %>%
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  #make dummy equaling zero for amalgated municipalities in 2005 election but otherwise 1
  mutate(muni_to_include = ifelse(kom_nr %!in% non_amalgated_vector & election_year==2001,0,1))


close_votes <- all_candidates %>%
  filter(OPSTILLINGSFORM == "Sideordnet" & muni_to_include==1) %>%
  group_by(party, KOMMUNE, election_year) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  filter(fractile_thres <= 18)


# the distribution of votes
p_votes <- ggplot(close_votes, aes(x = mid_pers_vote, fill = factor(elected))) +
  geom_density(alpha = .3) +
  theme_bw() +
  geom_vline(xintercept = mean(close_votes$mid_pers_vote), lty = 2)+
  geom_vline(xintercept = median(close_votes$mid_pers_vote), lty = 3) +
  labs(x = "Estimated Personal Votes", y = "Density") +
  scale_fill_manual(name = "Elected?", values = c("black", "grey"))

p_votes
```


0.005
0.004
0.003
0.002
0.001
0.000
Density
0
500
1000
1500
2000
Estimated Personal Votes
Elected?
0
1

Appendix P

Figure P1

```
all_candidates <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")


top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")


#Estimate of personal votes
all_candidates$interval_points <- gsub(pattern =  "\\[", replacement = "",
                                       x = all_candidates$interval_points)
all_candidates$interval_points <- gsub(pattern =  "\\]", replacement = "",
                                       x = all_candidates$interval_points)

all_candidates <- all_candidates %>%
  separate(col = "interval_points", into = c("lwr", "upr"), sep = ",") %>%
  mutate(lwr = as.numeric(as.character(lwr)),
         upr = as.numeric(as.character(upr)),
         mid_pers_vote = (lwr+upr)/2)


### aggregate party votes pr. municipality year

top_woman <- all_candidates %>%
  group_by(KOMMUNE, party, election_year) %>%
  mutate(top_woman = ifelse(mid_pers_vote == max(mid_pers_vote, na.rm=T) & female == 1, 1, 0)) %>%
  group_by(KOMMUNE, party) %>%
  mutate(top_woman_t1 = dplyr::lead(top_woman, n = 1, order_by = election_year)) %>%
  select(KOMMUNE, party, election_year, top_woman, top_woman_t1)

top_woman <- top_woman %>%
  group_by(KOMMUNE, party, election_year) %>%
  summarize(top_woman = ifelse(any(top_woman==1), 1, 0),
          top_woman_t1 = ifelse(any(top_woman_t1==1), 1, 0))


top2_cand <- left_join(top2_cand, top_woman, by = c("KOMMUNE", "party", "election_year"))

######## models
top_woman_mod <- felm(top_woman_t1 ~ female_win |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <=18 & election_year >1993))

stargazer(top_woman_mod,
          covariate.labels = c("Woman win t=0"),
          dep.var.labels = c("Woman as Party Leader"),
          add.lines = list(c("Bandwidth", rep(18, 5))),
          font.size = "scriptsize",
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/women_leaders.tex",
          title = "Effects on Parties Gaining Women as Leaders",
```

```
          label = "tab:top_woman")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: to, aug 15, 2024 - 17:09:52

Table 16: Effects on Parties Gaining Women as Leaders

	Dependent variable:
	Woman as Party Leader
Woman win t=0	0.104
	(0.070)
Constant	0.154^{***}
	(0.027)
Bandwidth	18
Observations	815
Note:	*p<0.1; **p<0.05; ***p<0.01

Appendix Q

```
all_candidates_rd <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/candidates_with_covariates.rds")

#Estimate of personal votes
all_candidates_rd$interval_points <- gsub(pattern =  "\\[", replacement = "",
                                          x = all_candidates_rd$interval_points)
all_candidates_rd$interval_points <- gsub(pattern =  "\\]", replacement = "",
                                          x = all_candidates_rd$interval_points)

all_candidates_rd <- all_candidates_rd %>%
  separate(col = "interval_points", into = c("lwr", "upr"), sep = ",") %>%
  mutate(lwr = as.numeric(as.character(lwr)),
         upr = as.numeric(as.character(upr)),
         mid_pers_vote = (lwr+upr)/2)

# define variable to sort amalgated municipalities away as outcome year in 2005
all_candidates_rd <- all_candidates_rd %>%
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  #make dummy equaling zero for amalgated municipalities in 2005 election but otherwise 1
  mutate(muni_to_include = ifelse(kom_nr %!in% non_amalgated_vector & election_year==2001,0,1)) %>%
  # change naming of three non-amalgated municipalities with changed spelling from Å to AA
  # after the reform
  mutate(KOMMUNE = ifelse(KOMMUNE=="751 Århus", "751 Aarhus", KOMMUNE)) %>%
  mutate(KOMMUNE = ifelse(KOMMUNE=="169 Høje Tåstrup", "169 Høje-Taastrup", KOMMUNE)) %>%
  mutate(KOMMUNE = ifelse(KOMMUNE=="173 Lyngby-Tårbæk", "173 Lyngby-Taarbæk", KOMMUNE))



# proportion of women in election-party-municipality groups
party_agg_rd <- all_candidates_rd %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(female_candidates = mean(female, na.rm = T),
            party_vote = mean(total_votes, na.rm = T))

party_agg_rd2 <- all_candidates_rd %>%
  filter(elected == 1) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(prop_female = mean(female, na.rm = T),
            n_elect = n())


party_agg_rd <- left_join(party_agg_rd, party_agg_rd2)

# votes for women
fem_vote <- all_candidates_rd %>%
  group_by(party, KOMMUNE, election_year) %>%
  filter(female == 1) %>%
  summarise(fem_vote = sum(mid_pers_vote, na.rm = T))

# votes for men
male_vote <- all_candidates_rd %>%
  group_by(party, KOMMUNE, election_year) %>%
```

```
  filter(female == 0) %>%
  summarise(male_vote = sum(mid_pers_vote, na.rm = T))


party_agg_rd <- left_join(party_agg_rd, fem_vote)
party_agg_rd <- left_join(party_agg_rd, male_vote)


# create covariates from individual-level data
covariates <- all_candidates_rd %>%
  group_by(election_year) %>%
  mutate(income_wins = Winsorize(income, probs = c(0.01,0.99), na.rm = T)) %>%
  ungroup() %>%
  mutate(immig = case_when(ethnicity %in% c(2:3) ~ 1,
                           ethnicity == 1 ~ 0,
                           TRUE ~ NA_real_),
         college = case_when(education %in% c("Bachelor",
                                              "Forskeruddannelser",
                                              "Mellemlange videregående uddannelser",
                                              "Lange videregående uddannelser") ~ 1,
                             TRUE ~ 0),
         vocational = case_when(education %in% c("Erhvervsfaglige praktik- og hovedforløb",
                                                 "Erhvervsgymnasiale uddannelser") ~ 1,
                                TRUE ~ 0),
         neg_income = case_when(income < 0 ~ 1,
                                TRUE ~ income),
         out_work = case_when(unemp %in% c(210, 321, 330) ~ 1,
                              TRUE ~ 0)) %>%

  group_by(party, KOMMUNE, election_year) %>%
  summarise(prop_immig = mean(immig, na.rm = T),
            prop_college = mean(college, na.rm = T),
            prop_vocational = mean(vocational, na.rm = T),
            av_age = mean(age, na.rm = T),
            prop_unemp = mean(out_work, na.rm = T),
            av_income = mean(income, na.rm = T),
            av_income_wins = mean(income_wins, na.rm=TRUE),
            av_competence = mean(inc_res, na.rm = T),
            prop_neg_income = mean(neg_income, na.rm = T))

party_agg_rd <- left_join(party_agg_rd, covariates)

# lags and leads

party_agg_rd <- party_agg_rd %>%
  mutate(fem_vote = case_when(is.na(fem_vote) == T ~ 0,
                              TRUE ~ fem_vote),
         male_vote = case_when(is.na(male_vote) == T ~ 0,
                               TRUE ~ male_vote)) %>%
  # change naming of three non-amalgated municipalities with changed spelling from Å to AA
  # after the reform
  mutate(KOMMUNE = ifelse(KOMMUNE=="751 Århus", "751 Aarhus", KOMMUNE)) %>%
  mutate(KOMMUNE = ifelse(KOMMUNE=="169 Høje Tåstrup", "169 Høje-Taastrup", KOMMUNE)) %>%
```

```
  mutate(KOMMUNE = ifelse(KOMMUNE=="173 Lyngby-Tårbæk", "173 Lyngby-Taarbæk", KOMMUNE)) %>%
  group_by(party, KOMMUNE) %>%
  mutate(prop_female_t1 = dplyr::lead(prop_female, n = 1, order_by = election_year),
         prop_female_t2 = dplyr::lead(prop_female, n = 2, order_by = election_year),
         fem_vote_t1 = dplyr::lead(fem_vote, n = 1, order_by = election_year),
         male_vote_t1 = dplyr::lead(male_vote, n = 1, order_by = election_year),
         party_vote_t1 = dplyr::lead(party_vote, n = 1, order_by = election_year),
         n_elect_t1 = dplyr::lead(n_elect, n = 1, order_by = election_year),
         female_candidates_t1 = dplyr::lead(female_candidates, n = 1, order_by = election_year)) %>%
  mutate(prop_female_lag = dplyr::lag(prop_female, 1, order_by = election_year),
         prop_immig_lag = dplyr::lag(prop_immig, 1, order_by = election_year),
         prop_college_lag = dplyr::lag(prop_college, 1, order_by = election_year),
         prop_vocational_lag = dplyr::lag(prop_vocational, 1, order_by = election_year),
         av_age_lag = dplyr::lag(av_age, 1, order_by = election_year),
         prop_unemp_lag = dplyr::lag(prop_unemp, 1, order_by = election_year),
         av_income_lag = dplyr::lag(av_income, 1, order_by = election_year),
         av_competence_lag = dplyr::lag(av_competence, 1, order_by = election_year),
         av_income_wins_lag = dplyr::lag(av_income_wins, 1, order_by = election_year),
         prop_neg_income_lag = dplyr::lag(prop_neg_income, 1, order_by = election_year)) %>%
  arrange(party, KOMMUNE, election_year)



#------------------------------
# estimates at optimal BW

### finding treatment and control group: parties where the marginal seat is between a man and woman
# create and get top two candidates within election-municipality-party
top2_cand_rd_all_rd <- all_candidates_rd %>%
  filter(OPSTILLINGSFORM == "Sideordnet") %>%
  group_by(party, KOMMUNE, election_year) %>%
  arrange(fractile_thres) %>%
  slice(1:2) %>%
  mutate(vote_diff = ifelse(elected == 1, mid_pers_vote - min(mid_pers_vote),
                            mid_pers_vote - max(mid_pers_vote))) %>%
  arrange(party, KOMMUNE, election_year)

# create indicators of female win
fem_win <- top2_cand_rd_all_rd %>%
  mutate(female_win = case_when(female == 1 & elected == 1 ~ 1,
                                TRUE ~ 0)) %>%
  group_by(party, KOMMUNE, election_year) %>%
  summarise(female_win = mean(female_win, na.rm = T),
            female_participant = mean(female, na.rm = T)) %>%
  mutate(female_win = case_when(female_win > 0 ~ 1,
                                TRUE ~ 0)) %>%
  # filter to exclude man/man mash-ups
  filter(female_participant == 0.5)

# define data set of parties with man/women competition of the marginal seat
top2_cand_rd <- left_join(fem_win, top2_cand_rd_all_rd, by = c("party", "KOMMUNE", "election_year"))

top2_cand_rd <- top2_cand_rd %>%
```

```
  group_by(party, KOMMUNE, election_year, female_win) %>%
  summarise(vote_diff = vote_diff[1])

top2_cand_rd <- top2_cand_rd %>%
  mutate(vote_diff_fem = ifelse(female_win == 1, abs(vote_diff), -abs(vote_diff) ))

top2_cand_rd <- left_join(top2_cand_rd, party_agg_rd,
                      by = c("party", "KOMMUNE", "election_year"))


top2_cand_rd$cluster_id <- top2_cand_rd %>%
  group_by(party, KOMMUNE) %>%
  group_indices()


top2_cand_rd$muni_id <- top2_cand_rd %>%
  group_by(KOMMUNE) %>%
  group_indices()


### remove amalgamated municipalities
# define variable to sort amalgated municipalities away as outcome year in 2005
top2_cand_rd <- top2_cand_rd %>%
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  #make dummy equaling zero for amalgated municipalities in 2005 election but otherwise 1
  mutate(muni_to_include = ifelse(kom_nr %!in% non_amalgated_vector & election_year==2001,0,1))


top2_cand_rd <- top2_cand_rd %>%
  filter(muni_to_include==1)

### leading n_elect and female_win creates NAs for parties that did not achieve representation
# in the next election
# these must be assigned to 0 - BUT parties that did not run at all in the next election also gets NA.
top2_cand_rd <- top2_cand_rd %>%
  # find the few parties that rerun but don't get any seats
  mutate(n_elect_t1 = ifelse(is.na(n_elect_t1) & !is.na(party_vote_t1),0,n_elect_t1),
         prop_female_t1 = ifelse(is.na(prop_female_t1) & !is.na(party_vote_t1),0,prop_female_t1))



# create unique id variable for year-municipality-party
top2_cand_rd <- top2_cand_rd %>%
  mutate(cluster = str_c(election_year, kom_nr, party, sep = "-", collapse = NULL))


top2_cand_rd <- top2_cand_rd %>%
  filter(muni_to_include==1 & election_year>1993 & election_year<2017 & !is.na(prop_female_t1))

# remove places with zero votes as vote difference
no_zero <- filter(top2_cand_rd, vote_diff_fem %in% -500:500 & vote_diff_fem != 0)

no_zero$vote_diff_fem_2 <- ifelse(no_zero$vote_diff_fem == 0 & no_zero$female_win == 0, -1,
```

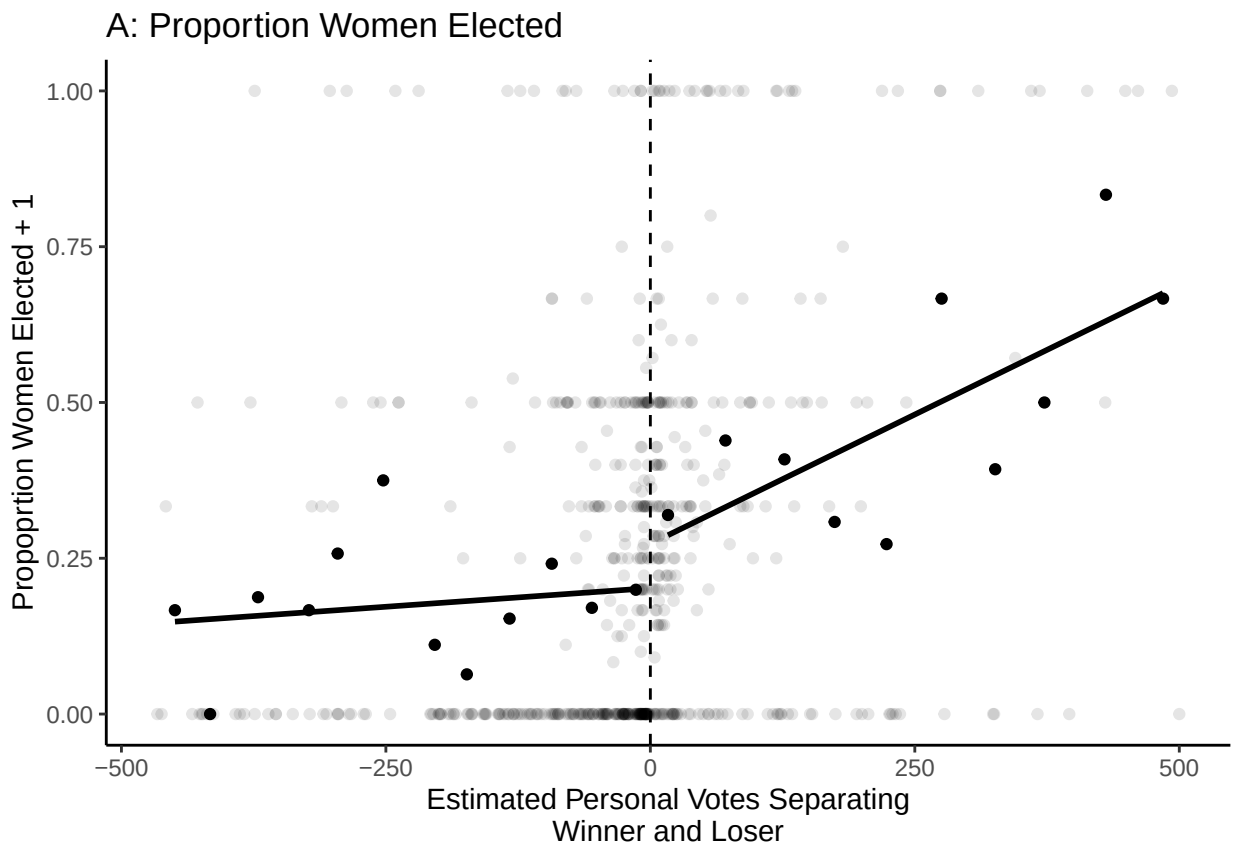
```
                                    no_zero$vote_diff_fem)

## find party-municipality-year observations with less than 3 candidates
# we cannot include these in the plot following Statistics Denmark's rules
# about aggregation of micro-level data

klynger <- unique(no_zero$cluster)

cluster_n2 <- all_candidates_rd %>%
  # create cluster variable for each candidate
  mutate(kom_nr = str_sub(KOMMUNE,1,3)) %>%
  mutate(cluster = str_c(election_year, kom_nr, party, sep = "-", collapse = NULL)) %>%
  filter(cluster %in% no_zero$cluster) %>%
  group_by(cluster) %>%
  count() %>%
  filter(n<3)

nrow(cluster_n2) # this concern 13 party-municipality-year observations
```

[1] 13

```
# remove these 13 instances of parties w. less than 3 candidates
no_zero <- no_zero %>% filter(cluster %!in% cluster_n2$cluster)



fem_prop_plot<- rdplot(y = no_zero$prop_female_t1,
                       x = no_zero$vote_diff_fem, p = 1,
                       hide = T)
```

[1] "Mass points detected in the running variable."

```
p1<-ggplot(fem_prop_plot$vars_bins, aes(x = rdplot_mean_x,
                                        y = rdplot_mean_y)) +
  geom_point() +
  geom_smooth(data = filter(fem_prop_plot$vars_bins,
                            rdplot_mean_x < 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black")+
  geom_smooth(data = filter(fem_prop_plot$vars_bins,
                            rdplot_mean_x >= 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black") +
  geom_point(data = no_zero,
             aes(x = vote_diff_fem, y= no_zero$prop_female_t1),
             inherit.aes = F, alpha = .1) +
  theme_classic() +
  labs(x = "Estimated Personal Votes Separating\nWinner and Loser",
       y = "Propoprtion Women Elected + 1",
       title = "A: Proportion Women Elected") +
  geom_vline(xintercept = 0, lty = 2)

p1
```




```
fem_vote_plot<- rdplot(y = log(no_zero$fem_vote_t1+1),
       x = no_zero$vote_diff_fem, p = 1,
       hide = T)
```

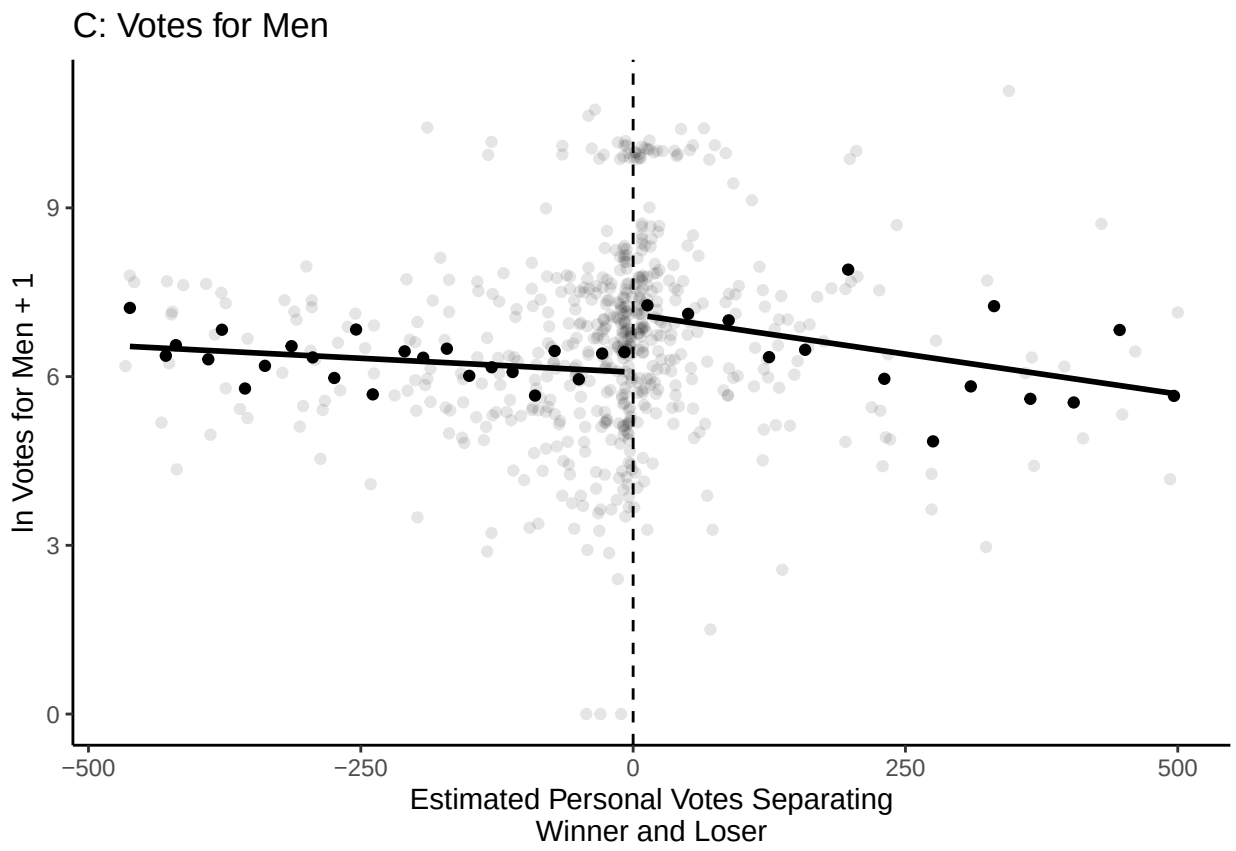
[1] "Mass points detected in the running variable."

```
p2<-ggplot(fem_vote_plot$vars_bins, aes(x = rdplot_mean_x,
                                        y = rdplot_mean_y)) +
  geom_point() +
  geom_smooth(data = filter(fem_vote_plot$vars_bins,
                            rdplot_mean_x < 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black")+
  geom_smooth(data = filter(fem_vote_plot$vars_bins,
                            rdplot_mean_x >= 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black") +
  geom_point(data = no_zero,
             aes(x = vote_diff_fem, y= log(no_zero$fem_vote_t1+1)),
             inherit.aes = F, alpha = .1) +
  theme_classic() +
  labs(x = "Estimated Personal Votes Separating\nWinner and Loser",
       y = "ln Votes for Women + 1",
       title = "B: Votes for Women") +
  geom_vline(xintercept = 0, lty = 2)
```

```
male_vote_plot<- rdplot(y = log(no_zero$male_vote_t1+1),
                        x = no_zero$vote_diff_fem, p = 1,
                        hide = T)
```

[1] "Mass points detected in the running variable."

```
p3<-ggplot(male_vote_plot$vars_bins, aes(x = rdplot_mean_x,
                                         y = rdplot_mean_y)) +
  geom_point() +
  geom_smooth(data = filter(male_vote_plot$vars_bins,
                            rdplot_mean_x < 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black")+
  geom_smooth(data = filter(male_vote_plot$vars_bins,
                            rdplot_mean_x >= 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black") +
  geom_point(data = no_zero,
             aes(x = vote_diff_fem, y= log(no_zero$male_vote_t1+1)),
             inherit.aes = F, alpha = .1) +
  theme_classic() +
  labs(x = "Estimated Personal Votes Separating\nWinner and Loser",
       y = "ln Votes for Men + 1",
       title = "C: Votes for Men") +
  geom_vline(xintercept = 0, lty = 2)

p3
```

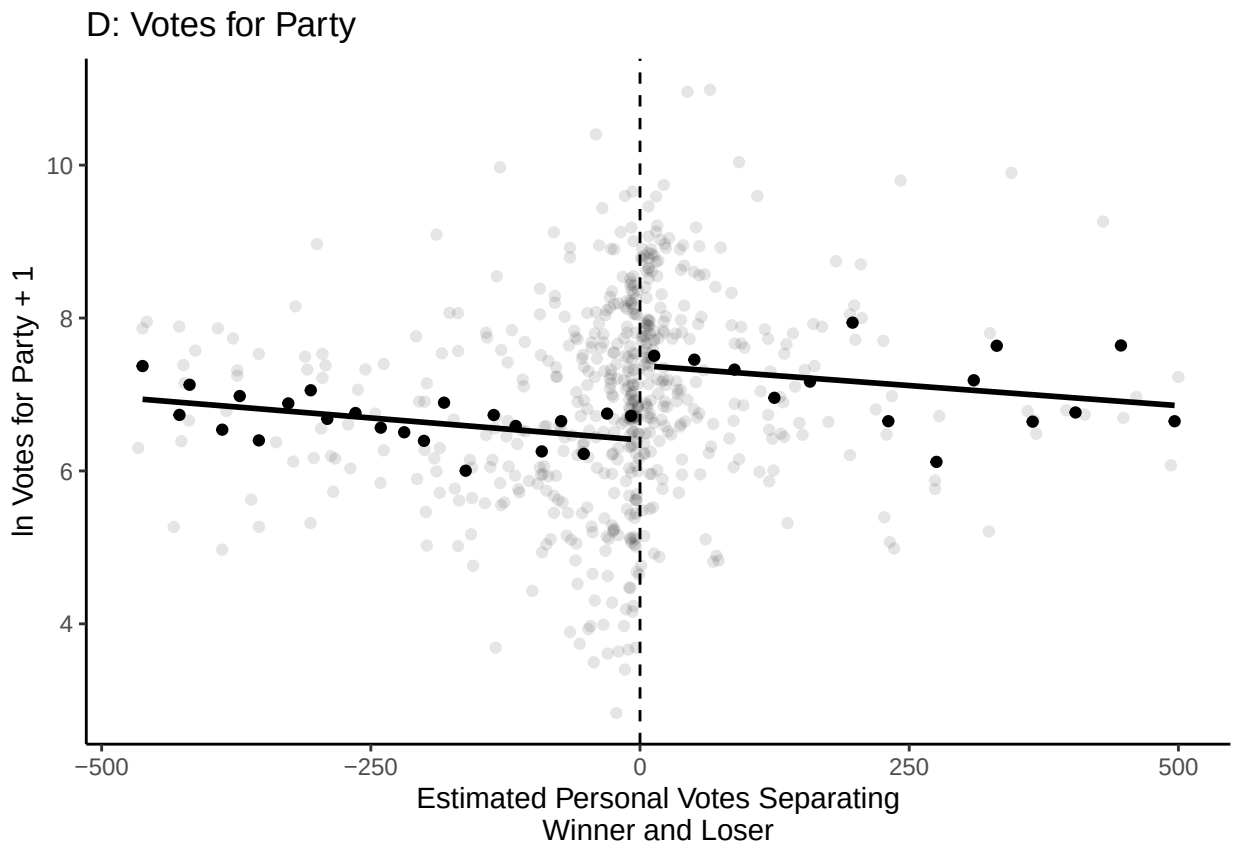



```
party_vote_plot<- rdplot(y = log(no_zero$party_vote_t1+1),
                         x = no_zero$vote_diff_fem, p = 1,
                         hide = T)
```

[1] "Mass points detected in the running variable."

```
p4<-ggplot(party_vote_plot$vars_bins, aes(x = rdplot_mean_x,
                                          y = rdplot_mean_y)) +
  geom_point() +
  geom_smooth(data = filter(party_vote_plot$vars_bins,
                            rdplot_mean_x < 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black")+
  geom_smooth(data = filter(party_vote_plot$vars_bins,
                            rdplot_mean_x >= 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black") +
  geom_point(data = no_zero,
             aes(x = vote_diff_fem, y= log(no_zero$party_vote_t1+1)),
             inherit.aes = F, alpha = .1) +
  theme_classic() +
  labs(x = "Estimated Personal Votes Separating\nWinner and Loser",
       y = "ln Votes for Party + 1",
       title = "D: Votes for Party") +
  geom_vline(xintercept = 0, lty = 2)
```

```
p4
```

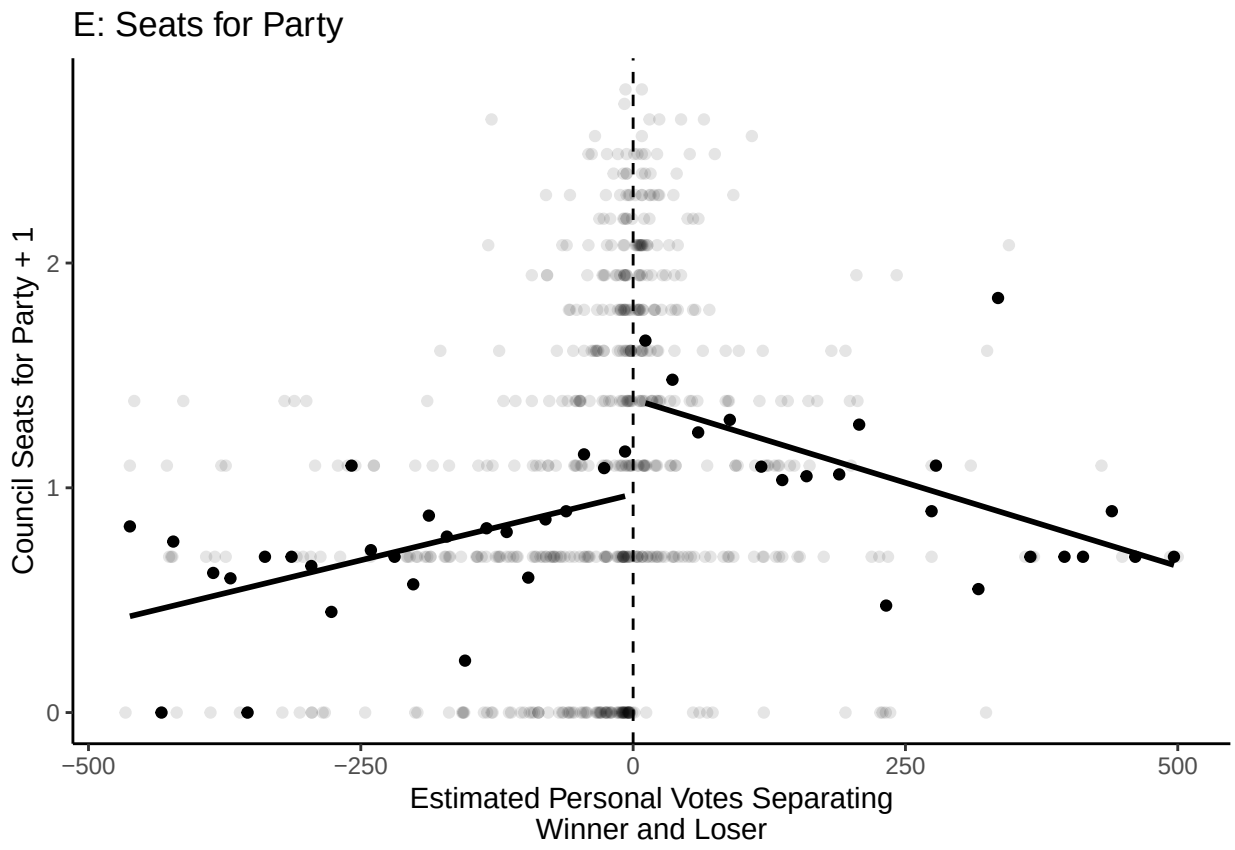



```
party_seats_plot<- rdplot(y = log(no_zero$n_elect_t1+1),
                          x = no_zero$vote_diff_fem, p = 1,
                          hide = T)
```

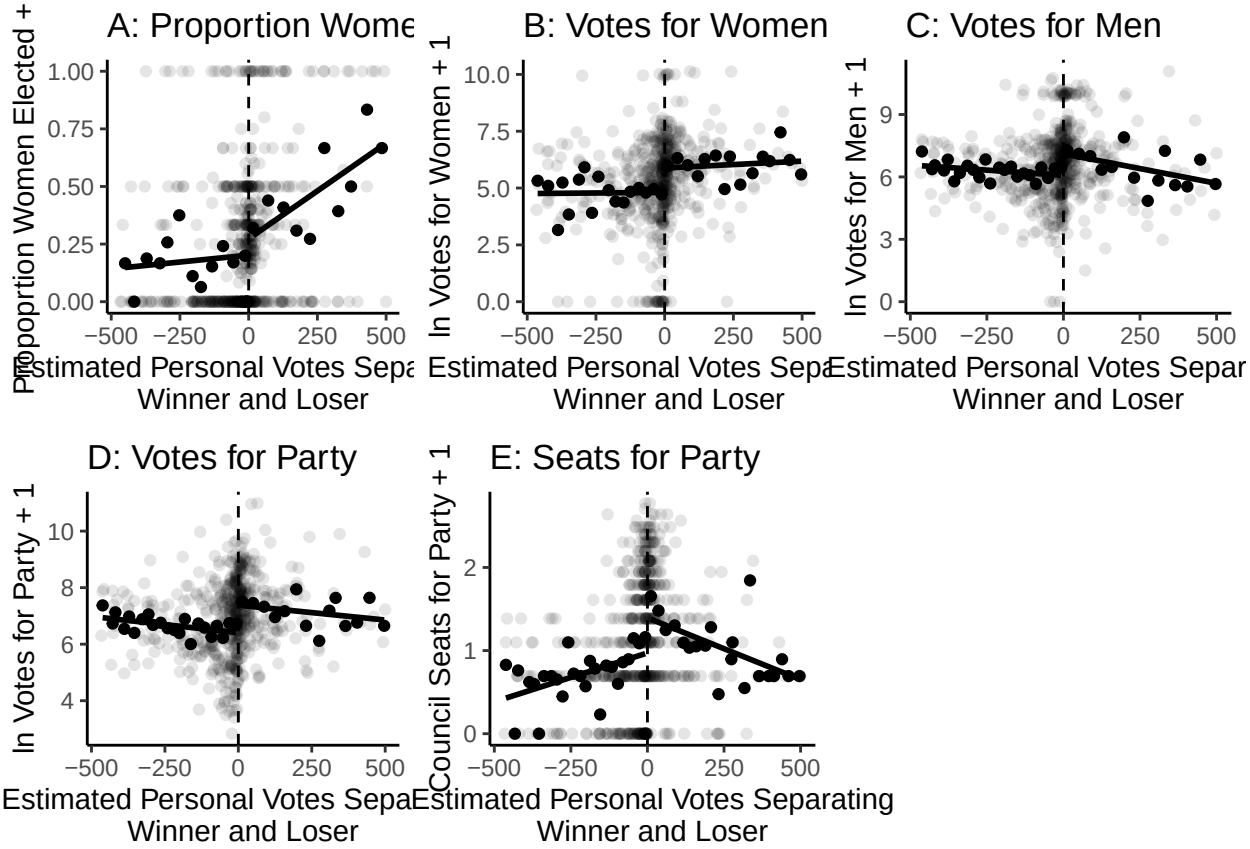
[1] "Mass points detected in the running variable."

```
p5<-ggplot(party_seats_plot$vars_bins, aes(x = rdplot_mean_x,
                                           y = rdplot_mean_y)) +
  geom_point() +
  geom_smooth(data = filter(party_seats_plot$vars_bins,
                            rdplot_mean_x < 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black")+
  geom_smooth(data = filter(party_seats_plot$vars_bins,
                            rdplot_mean_x >= 0),
              aes(x = rdplot_mean_x,
                  y = rdplot_mean_y), inherit.aes = F,
              method = "lm", se = F, color = "black") +
  geom_point(data = no_zero,
             aes(x = vote_diff_fem, y= log(no_zero$n_elect_t1+1)),
             inherit.aes = F, alpha = .1) +
  theme_classic() +
  labs(x = "Estimated Personal Votes Separating\nWinner and Loser",
```

```
       y = "Council Seats for Party + 1",
       title = "E: Seats for Party") +
  geom_vline(xintercept = 0, lty = 2)

p5
```

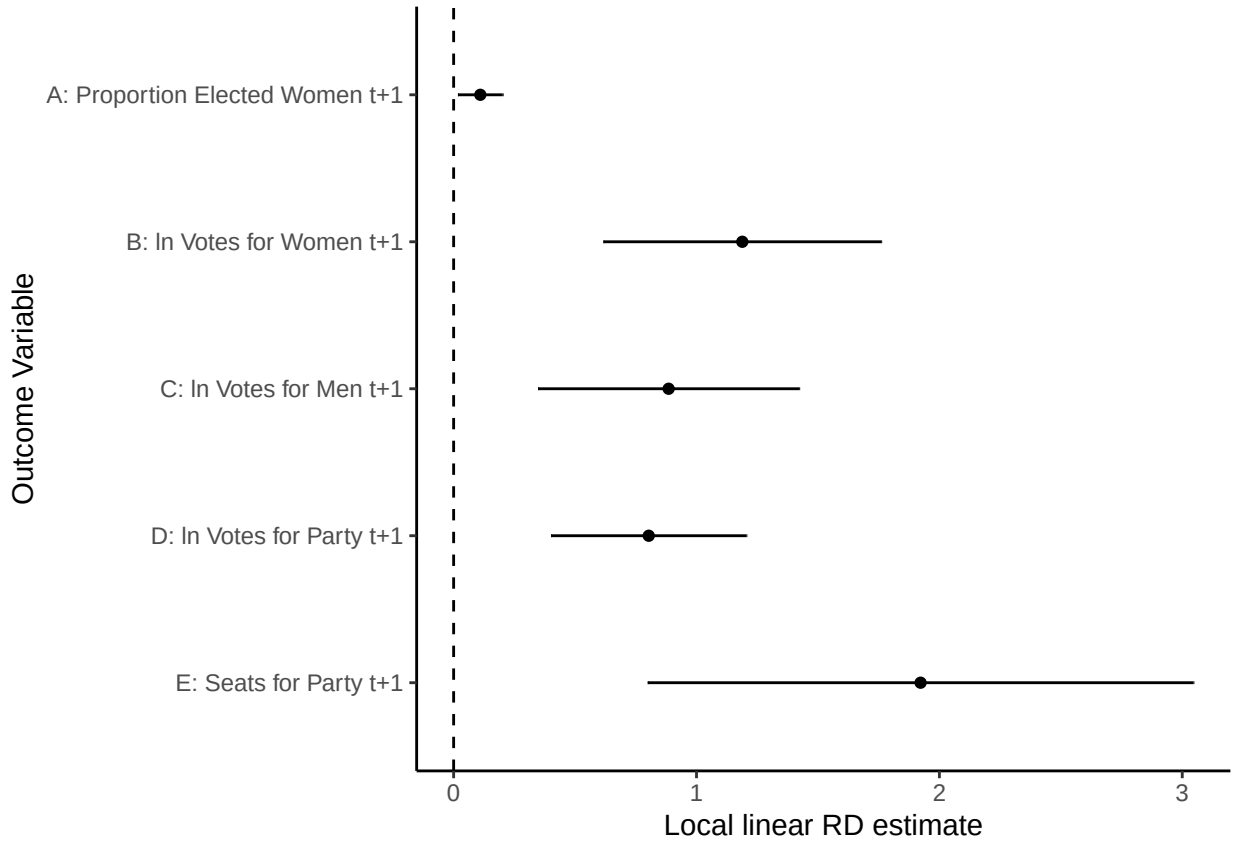



```
# combine plots

rd_plot <- plot_grid(p1, p2, p3,p4, p5)
rd_plot
```



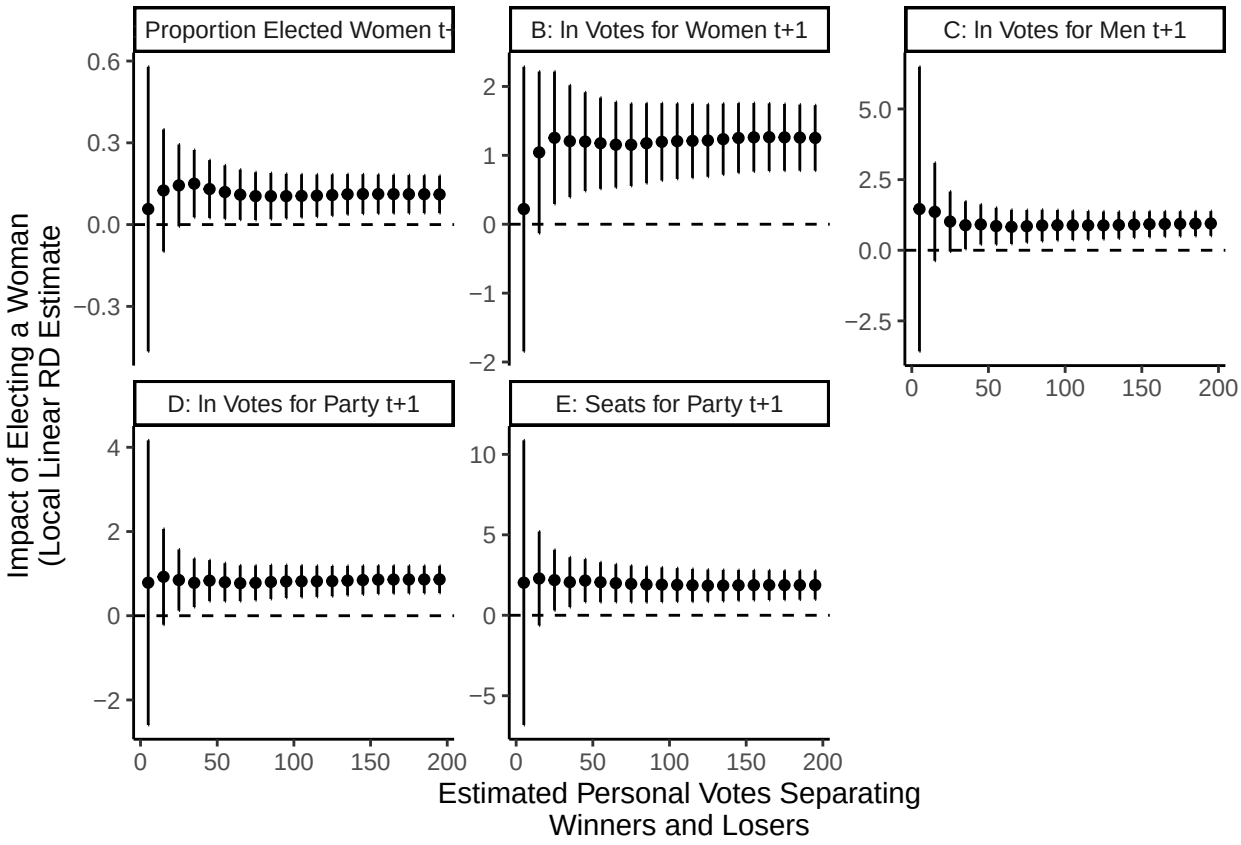

```
# rd models

rd_fem_prop <- rdrobust(y = no_zero$prop_female_t1,
                        x = no_zero$vote_diff_fem, masspoints = F,
                        cluster = no_zero$cluster_id)


rd_fem_vote <- rdrobust(y = log(no_zero$fem_vote_t1+1),
                    x = no_zero$vote_diff_fem, masspoints = F,
                    cluster = no_zero$cluster_id)


rd_male_vote <- rdrobust(y = log(no_zero$male_vote_t1+1),
                    x = no_zero$vote_diff_fem, masspoints = F,
                    cluster = no_zero$cluster_id)

rd_party_vote <- rdrobust(y = log(no_zero$party_vote_t1+1),
                    x = no_zero$vote_diff_fem, masspoints = F,
                    cluster = no_zero$cluster_id)


rd_seats <- rdrobust(y = no_zero$n_elect_t1,
                    x = no_zero$vote_diff_fem, masspoints = F,
                    cluster = no_zero$cluster_id)
```

```r
all_rd <- data.frame(est = c(rd_fem_prop$coef[1],
                             rd_fem_vote$coef[1],
                             rd_male_vote$coef[1],
                             rd_party_vote$coef[1],
                             rd_seats$coef[1]),
                     se = c(rd_fem_prop$se[1],
                            rd_fem_vote$se[1],
                            rd_male_vote$se[1],
                            rd_party_vote$se[1],
                            rd_seats$se[1]),
                     DV = c("A: Proportion Elected Women t+1",
                            "B: ln Votes for Women t+1",
                            "C: ln Votes for Men t+1",
                            "D: ln Votes for Party t+1",
                            "E: Seats for Party t+1"))

p_rd <- ggplot(all_rd, aes(x = est, y = reorder(DV, c(5,4,3,2,1)))) +
  geom_point() +
  geom_errorbarh(aes(xmin = est-1.96*se,
                     xmax = est+1.96*se), height = 0) +
  geom_vline(xintercept = 0, lty = 2) +
  theme_classic() +
  labs(y = "Outcome Variable",
       x = "Local linear RD estimate")

p_rd
```




```
rd_all_bw <- list()

for(i in seq(5, 200, 10)){


  rd_fem_prop <- rdrobust(y = no_zero$prop_female_t1,
                          x = no_zero$vote_diff_fem, masspoints = F, h = i,
                          cluster = no_zero$cluster_id)

  rd_fem_vote <- rdrobust(y = log(no_zero$fem_vote_t1+1),
                          x = no_zero$vote_diff_fem, masspoints = F, h = i,
                          cluster = no_zero$cluster_id)

  rd_male_vote <- rdrobust(y = log(no_zero$male_vote_t1+1),
                           x = no_zero$vote_diff_fem, masspoints = F, h = i,
                           cluster = no_zero$cluster_id)

  rd_party_vote <- rdrobust(y = log(no_zero$party_vote_t1+1),
                            x = no_zero$vote_diff_fem, masspoints = F, h = i,
                            cluster = no_zero$cluster_id)

  rd_seats <- rdrobust(y = no_zero$n_elect_t1,
                       x = no_zero$vote_diff_fem, masspoints = F, h = i,
                       cluster = no_zero$cluster_id)
```

```
  rd_all_bw[[i]] <- data.frame(est = c(rd_fem_prop$coef[1],
                               rd_fem_vote$coef[1],
                               rd_male_vote$coef[1],
                               rd_party_vote$coef[1],
                               rd_seats$coef[1]),
                       se = c(rd_fem_prop$se[1],
                              rd_fem_vote$se[1],
                              rd_male_vote$se[1],
                              rd_party_vote$se[1],
                              rd_seats$se[1]),
                       DV = c("A: Proportion Elected Women t+1",
                              "B: ln Votes for Women t+1",
                              "C: ln Votes for Men t+1",
                              "D: ln Votes for Party t+1",
                              "E: Seats for Party t+1"),
                       bw = i)


}

rd_all_bw <- do.call("rbind.data.frame", rd_all_bw)



p_rd_bw <- ggplot(rd_all_bw, aes(x = bw, y = est)) +
  geom_point() +
  geom_errorbar(aes(ymin = est - 1.96*se,
                    ymax = est + 1.96*se), width = 0)+
  facet_wrap( ~ DV, scales = "free_y") +
  theme_classic() +
  geom_hline(yintercept = 0, lty = 2) +
  labs(x = "Estimated Personal Votes Separating\nWinners and Losers",
       y = "Impact of Electing a Woman\n(Local Linear RD Estimate")

p_rd_bw
```


Proportion Elected Women t-
B: ln Votes for Women t+1
C: ln Votes for Men t+1
D: ln Votes for Party t+1
E: Seats for Party t+1
Impact of Electing a Woman
(Local Linear RD Estimate
Estimated Personal Votes Separating
Winners and Losers

Appendix R

Figure R1

```
top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")

close <- top2_cand %>%filter(election_year >1993 & fractile_thres <= 22 & muni_to_include==1)

# leave municipalities out
leave_muni_res <- list()

for(i in unique(close$muni_id)){

  this_close <- close %>% filter(muni_id != i)

fem_elect <- felm(prop_female_t1 ~ female_win |0|0|KOMMUNE + party,
                  data = this_close)

fem_vote <- felm(log(fem_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                 data = this_close)

male_vote <- felm(log(male_vote_t1) ~ female_win |0|0|KOMMUNE + party,
                  data = this_close)

party_vote_mod <- felm(log(party_vote_t1) ~ female_win |0|0|KOMMUNE + party,
                       data = this_close)

mandates_mod <- felm(n_elect_t1 ~ female_win |0|0|KOMMUNE + party,
                     data = this_close)

leave_muni_res[[i]] <- data.frame(
  est = c(coef(fem_elect)[2], coef(fem_vote)[2],
          coef(male_vote)[2], coef(party_vote_mod)[2],
          coef(mandates_mod)[2]),
  se = c(sqrt(diag(vcov(fem_elect)))[2], sqrt(diag(vcov(fem_vote)))[2],
         sqrt(diag(vcov(male_vote)))[2], sqrt(diag(vcov(party_vote_mod)))[2],
         sqrt(diag(vcov(mandates_mod)))[2]),
  iter = i,
  DV = c("Proportion Women", "Votes for Women (log)",
         "Votes for Men (log)", "Votes for Party (log)",
         "Seats for Party")
)


}

leave_muni_res <- do.call("rbind.data.frame",
                          leave_muni_res)


p_muni <- ggplot(leave_muni_res, aes(x = est, y = iter) )+
  geom_point() +
  geom_errorbarh(aes(xmin = est - 1.96*se,
```

```
                    xmax = est + 1.96*se), height = 0) +
  geom_vline(xintercept = 0, lty = 2) +
  facet_wrap(~ DV) +
  theme_bw() +
  labs(x = "Estimated Impact of Electing a Woman",
       y = "Municipality ID Excluded")


p_muni
```

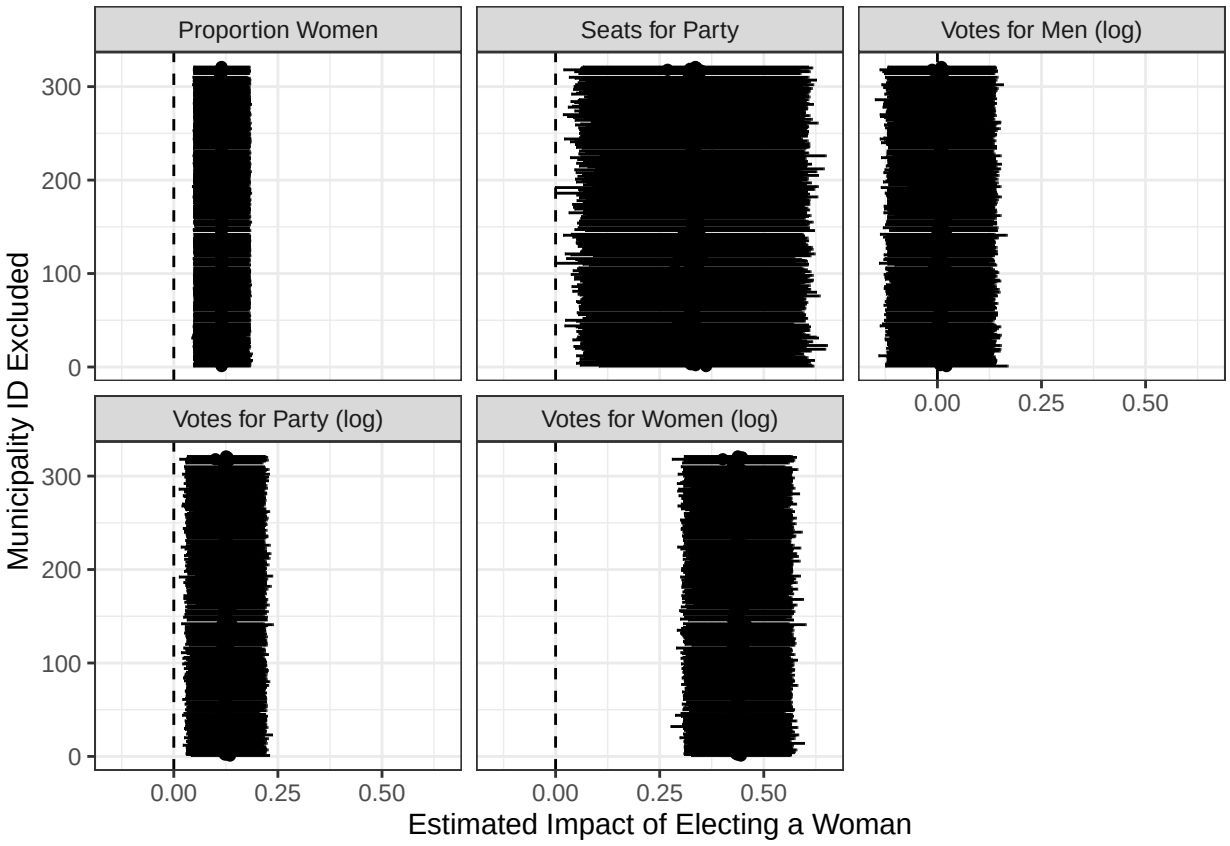
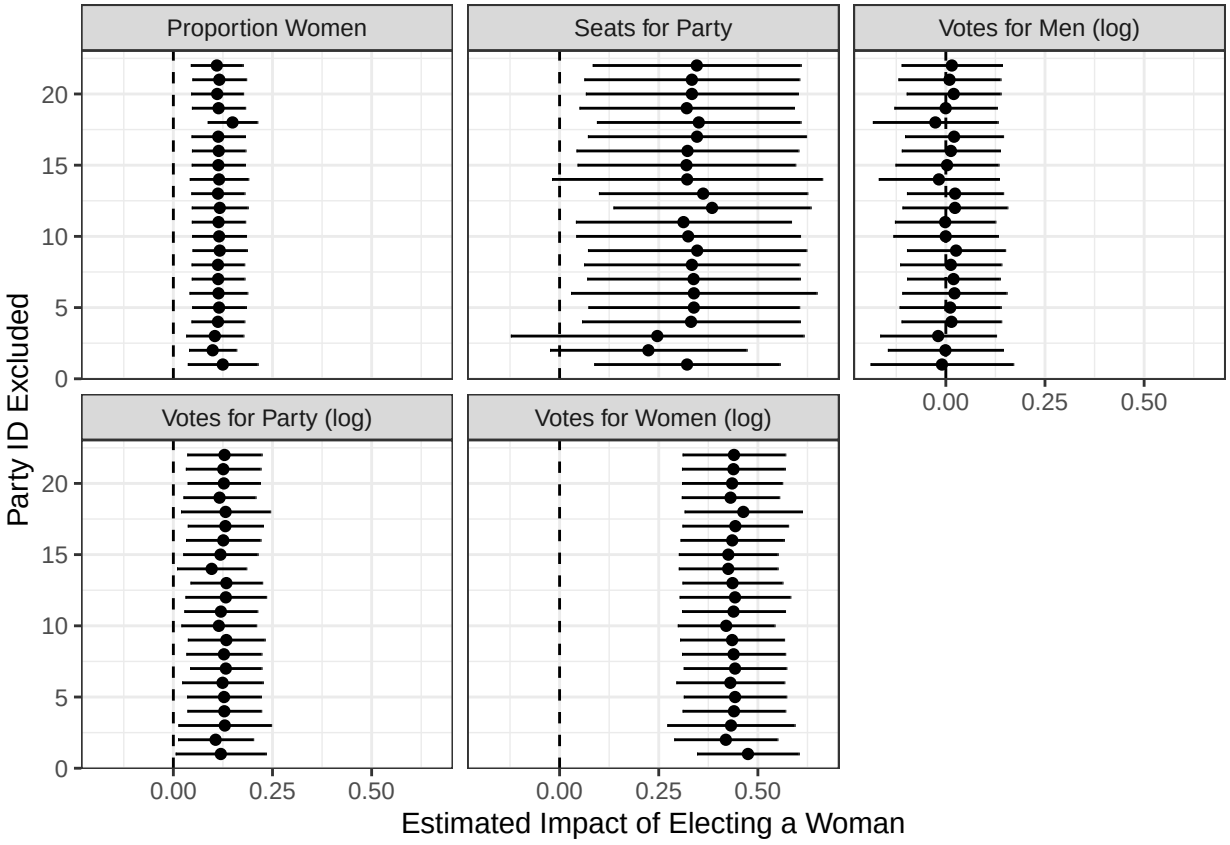


Figure R2

```
# leave out parties
close$party_id <- close %>%
  group_by(party) %>%
  group_indices()

leave_party_res <- list()

for(i in unique(close$party_id)){

  this_close <- close %>% filter(party_id != i)

  fem_elect <- felm(prop_female_t1 ~ female_win |0|0|KOMMUNE + party,
                    data = this_close)

  fem_vote <- felm(log(fem_vote_t1+1) ~ female_win |0|0|KOMMUNE + party,
                   data = this_close)

  male_vote <- felm(log(male_vote_t1) ~ female_win |0|0|KOMMUNE + party,
                    data = this_close)

  party_vote_mod <- felm(log(party_vote_t1) ~ female_win |0|0|KOMMUNE + party,
                         data = this_close)

  mandates_mod <- felm(n_elect_t1 ~ female_win |0|0|KOMMUNE + party,
                       data = this_close)

  leave_party_res[[i]] <- data.frame(
    est = c(coef(fem_elect)[2], coef(fem_vote)[2],
            coef(male_vote)[2], coef(party_vote_mod)[2],
            coef(mandates_mod)[2]),
    se = c(sqrt(diag(vcov(fem_elect)))[2], sqrt(diag(vcov(fem_vote)))[2],
           sqrt(diag(vcov(male_vote)))[2], sqrt(diag(vcov(party_vote_mod)))[2],
           sqrt(diag(vcov(mandates_mod)))[2]),
    iter = i,
    DV = c("Proportion Women", "Votes for Women (log)",
           "Votes for Men (log)", "Votes for Party (log)",
           "Seats for Party")
  )


}


leave_party_res <- do.call("rbind.data.frame",
                           leave_party_res)

p_party <- ggplot(leave_party_res, aes(x = est, y = iter) )+
  geom_point() +
  geom_errorbarh(aes(xmin = est - 1.96*se,
                     xmax = est + 1.96*se), height = 0) +
  geom_vline(xintercept = 0, lty = 2) +
  facet_wrap(~ DV) +
```

```
  theme_bw() +
  labs(x = "Estimated Impact of Electing a Woman",
       y = "Party ID Excluded")

p_party
```



Appendix S

Table S1

```
top2_cand <-
  readRDS("e:/workdata/706687/projects/women_spearhead/data/top2_cand_data.rds")


fem_elect <- felm(prop_female_t1 ~ female_win*scale(mid_pers_vote) |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 18))

fem_vote <- felm(log(fem_vote_t1+1) ~ female_win*scale(mid_pers_vote) |0|0|KOMMUNE + party,
                 data = filter(top2_cand, fractile_thres <= 18))

male_vote <- felm(log(male_vote_t1) ~ female_win*scale(mid_pers_vote) |0|0|KOMMUNE + party,
                  data = filter(top2_cand, fractile_thres <= 18))

party_vote_mod <- felm(log(party_vote_t1) ~ female_win*scale(mid_pers_vote) |0|0|KOMMUNE + party,
                       data = filter(top2_cand, fractile_thres <= 18))

mandates_mod <- felm(n_elect_t1 ~ female_win*scale(mid_pers_vote) |0|0|KOMMUNE + party,
                     data = filter(top2_cand, fractile_thres <= 18))

stargazer(fem_elect, fem_vote, male_vote, party_vote_mod, mandates_mod,
          covariate.labels = c("Woman win t=0", "Votes for Candidate",
                               "Votes X Woman win"),
          dep.var.labels = c(" Share Women Elected", "Log Votes Women", "Log Votes Men",
                             "Log Votes Party", "Seats Party"),
          add.lines = list(c("Bandwidth",
                             rep(18, 5))),
          font.size = "scriptsize",
          df = FALSE, omit.stat = c("rsq", "adj.rsq", "ser"),
          #out = "E:/workdata/706687/projects/women_spearhead/plots/votes_moderator.tex",
          title = "The Moderating Effect of Votes for the Marginal Candidate",
          label = "tab:visible_mod")
```

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac at gmail.com % Date and time: to, aug 15, 2024 - 17:10:34

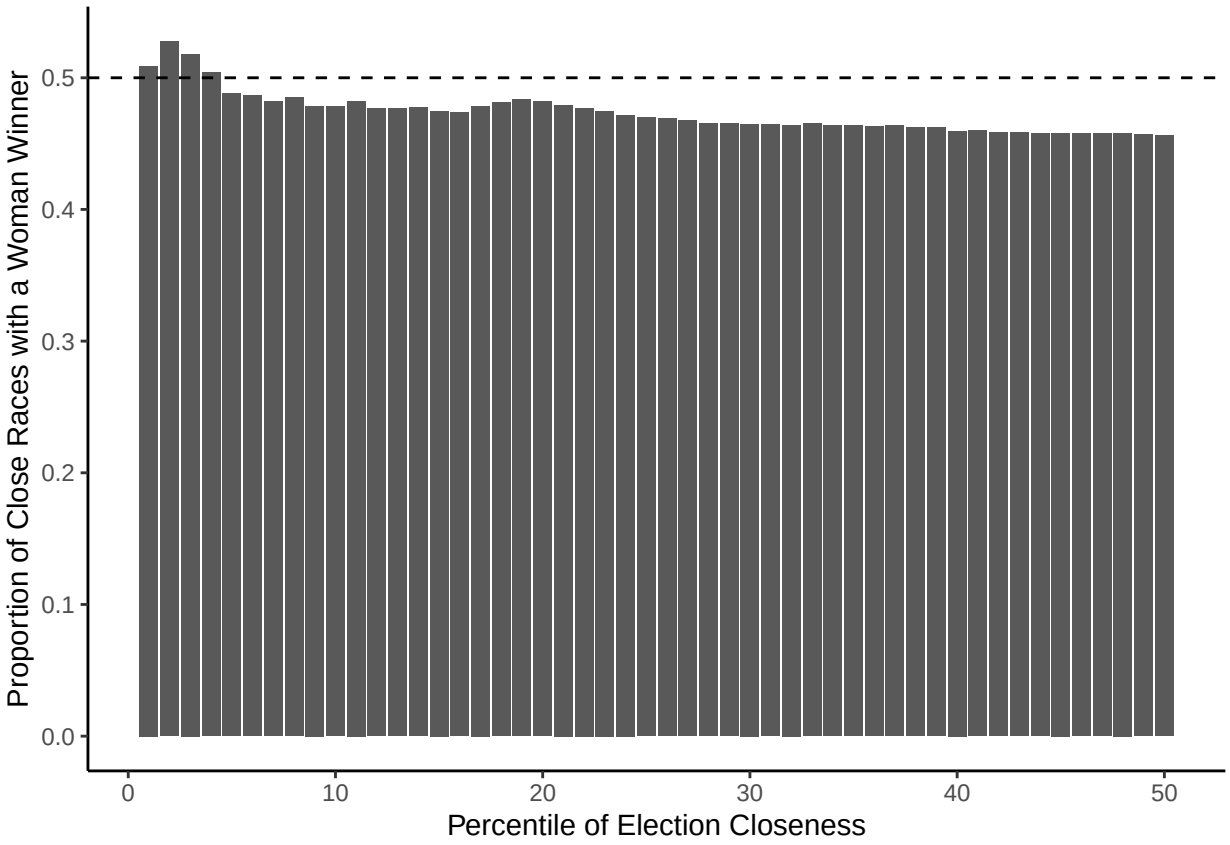
Table 17: The Moderating Effect of Votes for the Marginal Candidate

	Dependent variable:				
	Share Women Elected	Log Votes Women	Log Votes Men	Log Votes Party	Seats Party
	(1)	(2)	(3)	(4)	(5)
Woman win t=0	0.107^{***}	0.413^{***}	0.020	0.123^{***}	0.316^{***}
	(0.032)	(0.079)	(0.067)	(0.046)	(0.112)
Votes for Candidate	-0.014	0.397^{**}	0.298^{**}	0.283^{**}	-0.455^{**}
	(0.013)	(0.182)	(0.120)	(0.113)	(0.179)
Votes X Woman win	0.025	0.202	0.153^{***}	0.177^{***}	0.511^{***}
	(0.019)	(0.131)	(0.038)	(0.054)	(0.144)
Constant	0.238^{***}	5.799^{***}	7.327^{***}	7.494^{***}	4.556^{***}
	(0.018)	(0.291)	(0.297)	(0.240)	(0.971)
Bandwidth	18	18	18	18	18
Observations	819	819	819	819	819
Note:					*p<0.1; **p<0.05; ***p<0.01

Appendix T

Figure T1

```
prop_treat <- rep(NA, 50)
top2_cand <- as.data.frame(top2_cand)

for(i in 1:50){

  prop_treat[i] <- mean(top2_cand$female_win[top2_cand$fractile_thres<= i])


}

prop_treat <- data.frame(prop_treat, bw = 1:50)


p2 <- ggplot(prop_treat, aes(x = bw, y = prop_treat)) +
  geom_bar(stat = "identity") +
  theme_classic() +
  scale_y_continuous(breaks = seq(0, 0.5, 0.1)) +
  labs(x = "Percentile of Election Closeness",
       y = "Proportion of Close Races with a Woman Winner") +
  geom_hline(yintercept = 0.5, lty = 2)

p2
```


Proportion of Close Races with a Woman Winner
0.0
0.1
0.2
0.3
0.4
0.5
0
10
20
30
40
50
Percentile of Election Closeness