

Data Analysis of Roll Call Data

The summary statistics and plots of the data can be found in the plots.qmd file in the same directory.

Simple Linear Probability Model

```
# get all reps who voted only pos. or only neg. and regress with all contributions
df_1 <- df %>% filter(Vote_change_dummy == 0)
df_1$"all_votes" <- 0
# the all_votes column indicates if the rep voted all pos or neg. votes
for (i in 1:nrow(df_1)) {
  if (!is.na(any(df_1[i, c("Vote3", "Vote4", "Vote51", "Vote52", "Vote6", "Vote7")] == "+"))) {
    df_1[i, "all_votes"] <- "1"
  } else if (!is.na(any(df_1[i, c("Vote3", "Vote4", "Vote51", "Vote52", "Vote6", "Vote7")] == "-"))) {
    df_1[i, "all_votes"] <- "0"
  }
}

df_1 <- dummy_cols(df_1, select_columns = "all_votes")
df_1 <- df_1 %>%
  rename("all_votes_plus" = "all_votes_1") %>%
  rename("all_votes_minus" = "all_votes_0")

# view(df_1)
df_1 <- df_1 %>% select(-c(
  member_id, last_name, first_name, name, Vote_count,
  Vote_change, Vote_change_dummy, Vote3, Vote4, Vote51, Vote52, Vote6, Vote7
))
```

```
ols_1 <- lm(all_votes_minus ~ . - all_votes - all_votes_plus - state, data = df_1)
summary(ols_1)
```

Call:

```
lm(formula = all_votes_minus ~ . - all_votes - all_votes_plus -
    state, data = df_1)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.47094	-0.20848	-0.02834	0.07123	0.91849

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-8.537e-02	7.085e-02	-1.205	0.23034
partyR	3.072e-01	6.703e-02	4.583	1.04e-05 ***
District	6.438e-03	2.498e-03	2.577	0.01104 *
GeographicalNE	4.762e-02	9.628e-02	0.495	0.62170
GeographicalSO	2.899e-02	7.088e-02	0.409	0.68315
GeographicalWE	1.721e-04	7.882e-02	0.002	0.99826
pro_env.113	1.388e-07	3.890e-06	0.036	0.97159
anti_env.113	-9.014e-07	9.104e-07	-0.990	0.32391
pro_env.114	-1.840e-06	3.835e-06	-0.480	0.63221
anti_env.114	-1.240e-06	1.020e-06	-1.216	0.22616
pro_env.115	5.711e-06	5.878e-06	0.972	0.33296
anti_env.115	1.784e-06	1.201e-06	1.485	0.13980
pro_env.116	-5.764e-06	4.767e-06	-1.209	0.22874
anti_env.116	1.889e-06	8.086e-07	2.336	0.02097 *
pro_env.117	4.909e-07	2.636e-06	0.186	0.85256
anti_env.117	-1.766e-06	5.487e-07	-3.217	0.00162 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3035 on 134 degrees of freedom

(387 observations deleted due to missingness)

Multiple R-squared: 0.3424, Adjusted R-squared: 0.2688

F-statistic: 4.652 on 15 and 134 DF, p-value: 4.085e-07

LPM per session

```
# Vote 3
df_113 <- filter_session_data(df, "113")
ols_113 <- lm(Vote3_minus ~ . - Vote3 - Vote3_plus, data = df_113)
summary(ols_113)
```

Call:

```
lm(formula = Vote3_minus ~ . - Vote3 - Vote3_plus, data = df_113)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.13349	-0.05287	-0.01330	0.01118	0.94725

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.056e-02	2.750e-02	1.839	0.0673 .
partyR	9.184e-01	2.413e-02	38.067	<2e-16 ***
anti_env.113	2.653e-07	1.417e-07	1.872	0.0625 .
pro_env.113	-8.657e-07	9.395e-07	-0.921	0.3578
GeographicalNE	-4.147e-02	3.324e-02	-1.248	0.2135
GeographicalSO	7.565e-03	2.569e-02	0.295	0.7686
GeographicalWE	-3.807e-02	2.968e-02	-1.282	0.2010
District	1.891e-03	9.142e-04	2.069	0.0398 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1422 on 221 degrees of freedom

(135 observations deleted due to missingness)

Multiple R-squared: 0.9207, Adjusted R-squared: 0.9181

F-statistic: 366.3 on 7 and 221 DF, p-value: < 2.2e-16

```
# Vote 4
df_114 <- filter_session_data(df, "114")
ols_114 <- lm(Vote4_minus ~ . - Vote4 - Vote4_plus, data = df_114)
summary(ols_114)
```

Call:

```
lm(formula = Vote4_minus ~ . - Vote4 - Vote4_plus, data = df_114)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.97668	-0.03042	-0.00147	0.01626	0.93829

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.541e-02	2.636e-02	2.482	0.0137 *
partyR	9.092e-01	2.307e-02	39.404	<2e-16 ***
anti_env.114	2.724e-07	1.242e-07	2.194	0.0291 *
pro_env.114	-3.374e-07	7.745e-07	-0.436	0.6635
GeographicalNE	-6.383e-02	3.157e-02	-2.022	0.0442 *
GeographicalSO	1.006e-02	2.488e-02	0.404	0.6862
GeographicalWE	-3.418e-02	2.823e-02	-1.211	0.2271
District	-4.855e-04	8.779e-04	-0.553	0.5807

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1477 on 255 degrees of freedom

(157 observations deleted due to missingness)

Multiple R-squared: 0.9135, Adjusted R-squared: 0.9112

F-statistic: 384.9 on 7 and 255 DF, p-value: < 2.2e-16

```
# Vote 51
df_1151 <- filter_session_data(df, "1151")
ols_1151 <- lm(Vote51_minus ~ . - Vote51 - Vote51_plus, data = df_1151)
summary(ols_1151)
```

Call:

```
lm(formula = Vote51_minus ~ . - Vote51 - Vote51_plus, data = df_1151)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.97336	-0.02205	0.00698	0.02280	0.94659

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.067e-02	2.733e-02	2.220	0.02733 *
partyR	9.207e-01	2.310e-02	39.853	< 2e-16 ***
anti_env.115	2.603e-07	1.561e-07	1.667	0.09668 .

```

pro_env.115      -2.201e-07  8.797e-07  -0.250  0.80264
GeographicalNE -8.580e-02  3.078e-02  -2.788  0.00571 **
GeographicalSO -1.031e-02  2.573e-02  -0.401  0.68909
GeographicalWE -4.560e-02  2.835e-02  -1.609  0.10896
District        2.224e-04  8.965e-04   0.248  0.80430

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1489 on 252 degrees of freedom

(155 observations deleted due to missingness)

Multiple R-squared: 0.9128, Adjusted R-squared: 0.9104

F-statistic: 377.1 on 7 and 252 DF, p-value: < 2.2e-16

```

# Vote 52
df_1152 <- filter_session_data(df, "1152")
ols_1152 <- lm(Vote52_minus ~ . - Vote52 - Vote52_plus, data = df_1152)
summary(ols_1152)

```

Call:

```
lm(formula = Vote52_minus ~ . - Vote52 - Vote52_plus, data = df_1152)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.97143	-0.01887	0.03011	0.05869	0.96289

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.485e-02	3.616e-02	0.964	0.336
partyR	9.110e-01	3.008e-02	30.282	<2e-16 ***
anti_env.115	-2.157e-08	2.037e-07	-0.106	0.916
pro_env.115	-2.745e-07	1.168e-06	-0.235	0.814
GeographicalNE	-4.209e-02	4.090e-02	-1.029	0.304
GeographicalSO	2.646e-02	3.360e-02	0.787	0.432
GeographicalWE	-6.646e-03	3.790e-02	-0.175	0.861
District	-7.171e-04	1.227e-03	-0.584	0.559

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1947 on 247 degrees of freedom

(156 observations deleted due to missingness)

Multiple R-squared: 0.8524, Adjusted R-squared: 0.8482

F-statistic: 203.8 on 7 and 247 DF, p-value: < 2.2e-16

```
# Vote 6
df_116 <- filter_session_data(df, "116")
ols_116 <- lm(Vote6_minus ~ . - Vote6 - Vote6_plus, data = df_116)
summary(ols_116)
```

Call:

```
lm(formula = Vote6_minus ~ . - Vote6 - Vote6_plus, data = df_116)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.97634	-0.00979	0.01250	0.03225	0.96878

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.456e-02	2.394e-02	0.608	0.544
partyR	9.539e-01	2.175e-02	43.856	<2e-16 ***
anti_env.116	6.537e-08	1.186e-07	0.551	0.582
pro_env.116	2.390e-08	6.675e-07	0.036	0.971
GeographicalNE	-2.268e-02	2.860e-02	-0.793	0.429
GeographicalSO	8.351e-04	2.319e-02	0.036	0.971
GeographicalWE	2.383e-02	2.638e-02	0.903	0.367
District	-9.804e-04	8.500e-04	-1.153	0.250

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1374 on 254 degrees of freedom

(146 observations deleted due to missingness)

Multiple R-squared: 0.9263, Adjusted R-squared: 0.9243

F-statistic: 456.2 on 7 and 254 DF, p-value: < 2.2e-16

```
# Vote 7
df_117 <- filter_session_data(df, "117")
ols_117 <- lm(Vote7_minus ~ . - Vote7 - Vote7_plus, data = df_117)
summary(ols_117)
```

Call:

```
lm(formula = Vote7_minus ~ . - Vote7 - Vote7_plus, data = df_117)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.98236	-0.01476	0.02329	0.03485	0.10408

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.032e-02	2.336e-02	0.870	0.3854
partyR	9.529e-01	2.050e-02	46.477	<2e-16 ***
anti_env.117	-4.467e-08	1.031e-07	-0.433	0.6651
pro_env.117	2.139e-08	4.784e-07	0.045	0.9644
GeographicalNE	-4.910e-02	2.694e-02	-1.822	0.0697 .
GeographicalSO	1.069e-02	2.268e-02	0.471	0.6380
GeographicalWE	7.159e-03	2.493e-02	0.287	0.7742
District	-1.100e-03	8.441e-04	-1.303	0.1940

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1273 on 230 degrees of freedom

(117 observations deleted due to missingness)

Multiple R-squared: 0.936, Adjusted R-squared: 0.9341

F-statistic: 480.8 on 7 and 230 DF, p-value: < 2.2e-16

LPM of only mind changers

```
df_subsample <- df_subsample %>% select(-c(
  last_name, first_name, name, Vote_count, Vote_change,
  Vote_change, Vote_change_dummy, votes, member_id, Vote3, Vote4, Vote51, Vote52, Vote6, V
))
# view(df_subsample)

ols_subsample <- lm(vote_change_to_anti ~ ., data = df_subsample)
summary(ols_subsample)
```

Call:

lm(formula = vote_change_to_anti ~ ., data = df_subsample)

Residuals:

	Min	1Q	Median	3Q	Max
--	-----	----	--------	----	-----

-0.5000 -0.3333 0.0000 0.3750 0.6667

Coefficients: (8 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-7.167e-01	2.576e+00	-0.278	0.790
partyR	-2.071e-01	1.629e+00	-0.127	0.903
District	-1.756e-02	2.279e-02	-0.771	0.470
GeographicalNE	1.269e+00	2.920e+00	0.435	0.679
GeographicalSO	1.256e+00	2.544e+00	0.494	0.639
GeographicalWE	9.631e-01	1.905e+00	0.506	0.631
pro_env.113	1.276e-04	1.488e-04	0.857	0.424
anti_env.113	-4.329e-07	8.008e-06	-0.054	0.959
pro_env.114	NA	NA	NA	NA
anti_env.114	NA	NA	NA	NA
pro_env.115	NA	NA	NA	NA
anti_env.115	NA	NA	NA	NA
pro_env.116	NA	NA	NA	NA
anti_env.116	NA	NA	NA	NA
pro_env.117	NA	NA	NA	NA
anti_env.117	NA	NA	NA	NA

Residual standard error: 0.6009 on 6 degrees of freedom

(17 observations deleted due to missingness)

Multiple R-squared: 0.2417, Adjusted R-squared: -0.6431

F-statistic: 0.2732 on 7 and 6 DF, p-value: 0.9432

Why warning message?