Data Analysis of Roll Call Data

Table of contents

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

0.0.1 MODEL 1: Simple Linear Probability Model (OLS)

```
ols_1 <- lm(Vote_change_dummy ~ ., data = df_ols)
summary(ols_1)
```

```
Call:
lm(formula = Vote_change_dummy ~ ., data = df_ols)

Residuals:
     Min       1Q   Median       3Q      Max
```

```
-0.36323 -0.08900 -0.03073  0.04225  0.85614

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         6.820e-02  3.189e-02   2.139 0.034359 *
partyR             -1.518e-01  4.286e-02  -3.542 0.000556 ***
amount.oil.113     -8.330e-07  8.584e-07  -0.970 0.333698
amount.coal.113    -4.450e-06  5.966e-06  -0.746 0.457167
amount.mining.113   5.810e-06  4.063e-06   1.430 0.155178
amount.gas.113     -9.351e-07  4.958e-06  -0.189 0.850709
amount.env.113     -1.753e-07  3.153e-06  -0.056 0.955762
amount.alt_en.113  -6.919e-07  5.805e-06  -0.119 0.905314
amount.oil.114      2.799e-06  1.049e-06   2.669 0.008613 **
amount.coal.114    -2.812e-06  1.163e-05  -0.242 0.809344
amount.mining.114  -8.828e-07  6.142e-06  -0.144 0.885932
amount.gas.114      1.996e-06  5.492e-06   0.363 0.716855
amount.env.114      9.669e-07  3.081e-06   0.314 0.754191
amount.alt_en.114  -4.219e-06  6.599e-06  -0.639 0.523823
amount.oil.115     -5.681e-07  1.536e-06  -0.370 0.712118
amount.coal.115     1.117e-05  1.040e-05   1.074 0.284856
amount.mining.115  -5.952e-06  6.088e-06  -0.978 0.330121
amount.gas.115      3.506e-06  5.637e-06   0.622 0.535031
amount.env.115     -3.020e-06  5.404e-06  -0.559 0.577231
amount.alt_en.115  -7.277e-07  5.408e-06  -0.135 0.893185
amount.oil.116      4.503e-07  1.041e-06   0.432 0.666166
amount.coal.116    -7.205e-06  9.238e-06  -0.780 0.436873
amount.mining.116  -1.140e-06  2.627e-06  -0.434 0.664976
amount.gas.116     -6.622e-06  5.546e-06  -1.194 0.234662
amount.env.116     -1.180e-06  5.540e-06  -0.213 0.831740
amount.alt_en.116   7.405e-06  4.567e-06   1.622 0.107386
amount.oil.117     -8.185e-07  8.387e-07  -0.976 0.330981
amount.coal.117    -1.329e-05  1.234e-05  -1.077 0.283553
amount.mining.117   1.028e-05  7.695e-06   1.336 0.184041
amount.gas.117      6.090e-06  4.280e-06   1.423 0.157162
amount.env.117     -1.725e-06  3.179e-06  -0.543 0.588222
amount.alt_en.117  -2.717e-06  2.382e-06  -1.141 0.256168
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1938 on 127 degrees of freedom
  (401 observations deleted due to missingness)
Multiple R-squared:  0.3721,    Adjusted R-squared:  0.2188
F-statistic: 2.428 on 31 and 127 DF,  p-value: 0.0002924
```

here we can see the results are not very significant, only the party variable gives us significant results.

```
# view(df_sum)
ols_sum <- lm(Vote_change_dummy ~ ., data = df_sum)
summary(ols_sum)
```

```
Call:
lm(formula = Vote_change_dummy ~ ., data = df_sum)

Residuals:
     Min       1Q   Median       3Q      Max
-0.34468 -0.03988 -0.02993 -0.02358  0.96847

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.897e-02  1.435e-02   2.019 0.044020 *
partyR       5.263e-03  1.893e-02   0.278 0.781097
oil_sum     -3.839e-08  8.031e-08  -0.478 0.632798
coal_sum    -1.428e-06  1.049e-06  -1.361 0.173919
gas_sum      2.001e-06  5.404e-07   3.702 0.000235 ***
mining_sum   2.365e-07  6.302e-07   0.375 0.707595
env_sum     -4.296e-07  3.240e-07  -1.326 0.185466
alt_en_sum  -3.039e-07  5.164e-07  -0.589 0.556430
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1938 on 552 degrees of freedom
Multiple R-squared:  0.05964,   Adjusted R-squared:  0.04772
F-statistic: 5.002 on 7 and 552 DF,  p-value: 1.651e-05
```

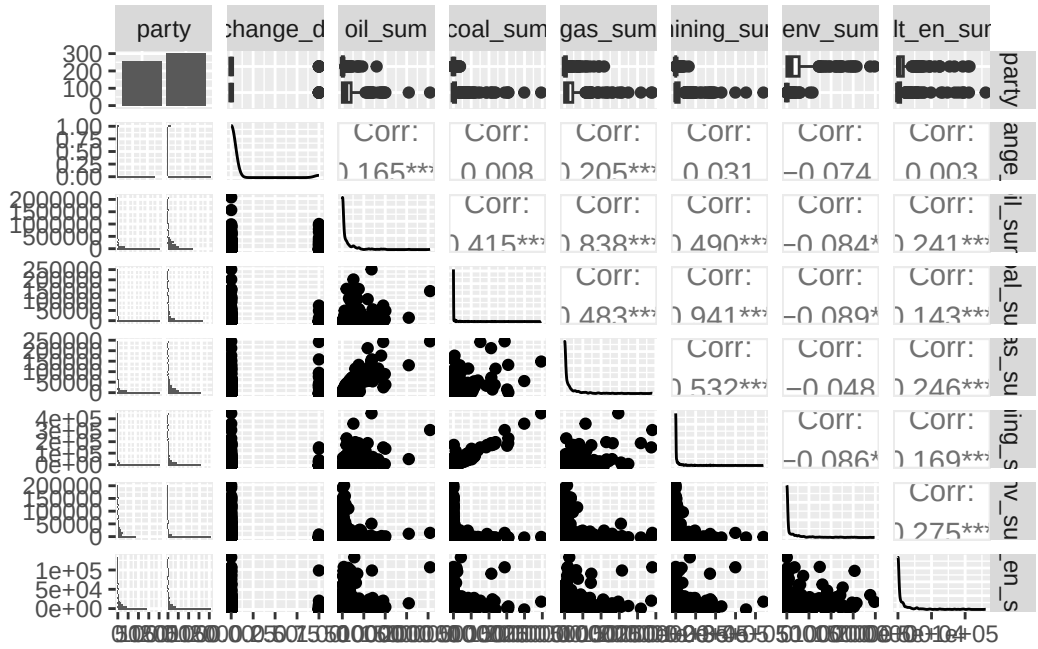
interestingly, at least amount_gas is more significant here, but party variable does not produce any significant results.

0.0.2 plot correllelograms of variables

Results are not very significant, we need to add more variables, but also we need to check each session with the vote changes.

```
# ggpairs(df_ols)
ggpairs(df_sum)
```

```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Interpretation...

0.0.3 113th session contributions & vote change

I am looking at each session independently, aka the vote of the representative in the 113th session and the corresponding contributions received. I am regressing these values with the Vote_change_dummy. This does not necessarily mean that a representative had to have changed their vote in the next session, but just how representatives who eventually changed their votes voted, and how their campaign contributions looked like in the 113th session.

```r
df_113 <- filter_session_data(df, "113")
ols_113 <- lm(Vote_change_dummy ~ ., data = df_113)
summary(ols_113)
```

```
Call:
lm(formula = Vote_change_dummy ~ ., data = df_113)

Residuals:
     Min       1Q   Median       3Q      Max
-0.66110 -0.04054 -0.01896 -0.00082  0.96801

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        7.503e-01  8.503e-02   8.823 3.35e-16 ***
partyR            -7.528e-01  8.547e-02  -8.809 3.69e-16 ***
Vote3+            -7.297e-01  8.691e-02  -8.396 5.47e-15 ***
amount.oil.113     1.719e-07  2.960e-07   0.581 0.562013
amount.coal.113   -2.441e-06  3.114e-06  -0.784 0.433975
amount.mining.113  1.104e-06  2.041e-06   0.541 0.589194
amount.gas.113     7.807e-06  2.318e-06   3.368 0.000892 ***
amount.env.113    -2.256e-07  1.452e-06  -0.155 0.876617
amount.alt_en.113 -5.082e-06  3.266e-06  -1.556 0.121151
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1846 on 222 degrees of freedom
  (329 observations deleted due to missingness)
Multiple R-squared:  0.3352,    Adjusted R-squared:  0.3112
F-statistic: 13.99 on 8 and 222 DF,  p-value: < 2.2e-16
```

b0 is 0.75, aka the change of a representative changing their votes is to 75% not affected by the following coefficients, and holding all else constant. b0 is significant at a 0 level.

b1 is the coefficient which determines that being in the republican party is associated with a -0.752 change in vote. Meaning compared to democrats, republicans in the 113th session are less likely to change their votes. b1 is significant at a 0 level.

b2 describes that voting pro-environment on the methane bill introduced in the 113th congress decreases the chance of changing votes by 0.73, this coefficient is significant on a 0 level.

from the amount variables. only b6 is significant on a high level, which means that for a one unit increase in natural gas contributions to a representatives campaign, will increase their change of a change in vote by 0.0000078, aka 0.00078 percent.

0.0.4 114th session contributions & vote change

```
df_114 <- filter_session_data(df, "114")
ols_114 <- lm(Vote_change_dummy ~ ., data = df_114)
summary(ols_114)
```

```
Call:
lm(formula = Vote_change_dummy ~ ., data = df_114)

Residuals:
     Min       1Q   Median       3Q      Max
-0.50487 -0.05104 -0.02991 -0.00187  1.00908

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        4.917e-01  8.180e-02   6.011 6.31e-09 ***
partyR            -4.937e-01  8.093e-02  -6.100 3.89e-09 ***
Vote4+            -4.564e-01  8.168e-02  -5.588 5.87e-08 ***
amount.oil.114     3.856e-07  3.124e-07   1.234  0.21815
amount.coal.114   -3.401e-06  4.247e-06  -0.801  0.42406
amount.mining.114  4.663e-07  2.481e-06   0.188  0.85107
amount.gas.114     6.422e-06  2.015e-06   3.187  0.00162 **
amount.env.114    -9.729e-07  1.141e-06  -0.853  0.39473
amount.alt_en.114 -2.893e-06  2.993e-06  -0.967  0.33453
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.194 on 256 degrees of freedom
  (295 observations deleted due to missingness)
Multiple R-squared:  0.2208,    Adjusted R-squared:  0.1965
F-statistic: 9.068 on 8 and 256 DF,  p-value: 5.782e-11
```

0.0.5 115th session contributions & vote change

```
# incl. two votes
df_1151 <- filter_session_data(df, "1151")
df_1152 <- filter_session_data(df, "1152")
ols_1151 <- lm(Vote_change_dummy ~ ., data = df_1151)
summary(ols_1151)
```

```
Call:
lm(formula = Vote_change_dummy ~ ., data = df_1151)

Residuals:
     Min       1Q   Median       3Q      Max
-0.34501 -0.06536 -0.04049 -0.01533  0.98827

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         5.294e-02  9.233e-02   0.573   0.5669
partyR             -4.248e-02  9.116e-02  -0.466   0.6416
Vote51+             1.084e-02  9.180e-02   0.118   0.9061
amount.oil.115      2.124e-07  4.041e-07   0.526   0.5995
amount.coal.115    -4.463e-06  4.513e-06  -0.989   0.3237
amount.mining.115   4.377e-07  2.991e-06   0.146   0.8838
amount.gas.115      5.727e-06  2.208e-06   2.593   0.0101 *
amount.env.115     -3.224e-06  1.661e-06  -1.942   0.0533 .
amount.alt_en.115   7.731e-07  2.299e-06   0.336   0.7369
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2185 on 255 degrees of freedom
  (296 observations deleted due to missingness)
Multiple R-squared:  0.08148,   Adjusted R-squared:  0.05266
F-statistic: 2.828 on 8 and 255 DF,  p-value: 0.005056
```

```
ols_1151 <- lm(Vote_change_dummy ~ ., data = df_1151)
summary(ols_1151)
```

```
Call:
```

```
lm(formula = Vote_change_dummy ~ ., data = df_1151)

Residuals:
     Min       1Q   Median       3Q      Max
-0.34501 -0.06536 -0.04049 -0.01533  0.98827

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         5.294e-02  9.233e-02   0.573   0.5669
partyR             -4.248e-02  9.116e-02  -0.466   0.6416
Vote51+             1.084e-02  9.180e-02   0.118   0.9061
amount.oil.115      2.124e-07  4.041e-07   0.526   0.5995
amount.coal.115    -4.463e-06  4.513e-06  -0.989   0.3237
amount.mining.115   4.377e-07  2.991e-06   0.146   0.8838
amount.gas.115      5.727e-06  2.208e-06   2.593   0.0101 *
amount.env.115     -3.224e-06  1.661e-06  -1.942   0.0533 .
amount.alt_en.115   7.731e-07  2.299e-06   0.336   0.7369
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2185 on 255 degrees of freedom
  (296 observations deleted due to missingness)
Multiple R-squared:  0.08148,   Adjusted R-squared:  0.05266
F-statistic: 2.828 on 8 and 255 DF,  p-value: 0.005056
```

not much change in the 115th session. is this because in between the 115th session there were no vote changes? There were 7 vote changes within the 115th session...

0.0.6 116th session contributions & vote change

```
df_116 <- filter_session_data(df, "116")
df_116 <- lm(Vote_change_dummy ~ ., data = df_116)
summary(df_116)
```

```
Call:
lm(formula = Vote_change_dummy ~ ., data = df_116)

Residuals:
     Min       1Q   Median       3Q      Max
```

```
-0.25094 -0.05440 -0.03430 -0.00344  1.05898

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       -1.113e-01  8.704e-02  -1.278   0.2023
partyR             1.090e-01  8.593e-02   1.268   0.2060
Vote6+             1.669e-01  8.563e-02   1.949   0.0524 .
amount.oil.116     1.827e-07  3.373e-07   0.542   0.5884
amount.coal.116   -1.685e-06  3.493e-06  -0.483   0.6298
amount.mining.116 -4.438e-07  1.557e-06  -0.285   0.7759
amount.gas.116     3.448e-06  2.044e-06   1.687   0.0927 .
amount.env.116    -2.385e-06  1.215e-06  -1.963   0.0507 .
amount.alt_en.116  9.334e-07  1.640e-06   0.569   0.5698
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1858 on 259 degrees of freedom
  (292 observations deleted due to missingness)
Multiple R-squared:  0.07076,   Adjusted R-squared:  0.04206
F-statistic: 2.465 on 8 and 259 DF,  p-value: 0.0137
```

0.0.7 117th session contributions & vote change

```r
df_117 <- filter_session_data(df, "117")
df_117 <- lm(Vote_change_dummy ~ ., data = df_117)
summary(df_117)
```

```
Call:
lm(formula = Vote_change_dummy ~ ., data = df_117)

Residuals:
     Min       1Q   Median       3Q      Max
-0.76499 -0.05377 -0.02651  0.01391  1.00875

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       -6.769e-01  9.131e-02  -7.413 2.24e-12 ***
partyR             6.509e-01  8.953e-02   7.270 5.36e-12 ***
Vote7+             7.320e-01  8.971e-02   8.160 2.07e-14 ***
```

```
amount.oil.117     4.815e-07  2.965e-07   1.624   0.1058
amount.coal.117   -8.706e-06  4.099e-06  -2.124   0.0347 *
amount.mining.117  5.518e-06  3.092e-06   1.784   0.0757 .
amount.gas.117     5.647e-07  1.784e-06   0.317   0.7518
amount.env.117    -1.292e-06  8.082e-07  -1.599   0.1111
amount.alt_en.117 -1.708e-06  1.337e-06  -1.277   0.2028
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1719 on 234 degrees of freedom
  (317 observations deleted due to missingness)
Multiple R-squared:  0.2791,    Adjusted R-squared:  0.2544
F-statistic: 11.32 on 8 and 234 DF,  p-value: 1.536e-13
```

0.0.8 MODEL 2: Subsample OLS

here I only look at those who did change their voting behavior. For this I created a seperate df out of the original df, where i only included representatives who changed their minds. I marked the changes, where from pro-environment vote "+" to anti-environment "-" = 0, and from anti-environment to pro-environment = 1. ### vote changers dataframe

```
mind_changers <- df %>% filter(Vote_change_dummy == 1)
# view(mind_changers)

knitr::kable(mind_changers, format = "html")
```

last_name	first_name	name	party	state	member_id	Vote3	Vote4
Costa	Jim	jim costa	D	CA	C001059	-	+
Coffman	Mike	mike coffman	R	CO	C001077	-	-
Ros-Lehtinen	Ileana	ileana ros lehtinen	R	FL	NA	-	+
Bishop	Sanford D.	sanford d bishop	D	GA	B000490	+	-
Roskam	Peter J.	peter j roskam	R	IL	R000580	-	-
Upton	Fred	fred upton	R	MI	U000031	-	-
Smith	Christopher H.	christopher h smith	R	NJ	S000522	-	-
Lance	Leonard	leonard lance	R	NJ	L000567	-	-
Hanna	Richard	richard hanna,richard l hanna	R	NY	H001051	-	+
Reed	Tom	tom reed	R	NY	R000585	-	-
Meehan	Patrick	patrick meehan	R	PA	M001181	-	-
Sanford	Mark	mark sanford	R	SC	S000051	-	-
Cuellar	Henry	henry cuellar	D	TX	C001063	-	-
Green	Gene	gene green	D	TX	G000410	-	+

last_name	first_name	name	party	state	member_id	Vote3	Vote4
Veasey	Marc A.	marc a veasey	D	TX	V000131	+	-
Vela	Filemon	filemon vela	D	TX	V000132	-	+
Reichert	Dave	dave reichert	R	WA	NA	-	-
Knight	Steve	steve knight	R	CA	NA	NA	-
Stefanik	Elise	elise stefanik,elise m stefanik	R	NY	S001196	NA	+
Costello	Ryan	ryan costello,ryan a costello	R	PA	C001106	NA	-
Gaetz	Matt	matt gaetz	R	FL	G000578	NA	NA
Gonzalez	Vicente	vicente gonzalez	D	TX	G000581	NA	NA
Waltz	Michael	michael waltz	R	FL	W000823	NA	NA

1 OLS

```
knitr::kable(df_subsample, format = "html")
```

last_name	first_name	name	change	votes	party	state	member_i
Costa	Jim	jim costa	1	change1	D	CA	C001059
Costa	Jim	jim costa	0	change2	D	CA	C001059
Costa	Jim	jim costa	1	change3	D	CA	C001059
Coffman	Mike	mike coffman	1	change1	R	CO	C001077
Ros-Lehtinen	Ileana	ileana ros lehtinen	1	change1	R	FL	NA
Bishop	Sanford D.	sanford d bishop	0	change1	D	GA	B000490
Bishop	Sanford D.	sanford d bishop	1	change2	D	GA	B000490
Bishop	Sanford D.	sanford d bishop	0	change3	D	GA	B000490
Bishop	Sanford D.	sanford d bishop	1	change4	D	GA	B000490
Roskam	Peter J.	peter j roskam	1	change1	R	IL	R000580
Upton	Fred	fred upton	1	change1	R	MI	U000031
Smith	Christopher H.	christopher h smith	1	change1	R	NJ	S000522
Smith	Christopher H.	christopher h smith	0	change2	R	NJ	S000522
Lance	Leonard	leonard lance	1	change1	R	NJ	L000567
Hanna	Richard	richard hanna,richard l hanna	1	change1	R	NY	H001051
Reed	Tom	tom reed	1	change1	R	NY	R000585
Meehan	Patrick	patrick meehan	1	change1	R	PA	M001181
Sanford	Mark	mark sanford	1	change1	R	SC	S000051
Cuellar	Henry	henry cuellar	1	change1	D	TX	C001063
Green	Gene	gene green	1	change1	D	TX	G000410
Veasey	Marc A.	marc a veasey	0	change1	D	TX	V000131
Veasey	Marc A.	marc a veasey	1	change2	D	TX	V000131

last_name	first_name	name	change	votes	party	state	member_i
Vela	Filemon	filemon vela	1	change1	D	TX	V000132
Reichert	Dave	dave reichert	1	change1	R	WA	NA
Knight	Steve	steve knight	1	change1	R	CA	NA
Knight	Steve	steve knight	0	change2	R	CA	NA
Stefanik	Elise	elise stefanik,elise m stefanik	0	change1	R	NY	S001196
Costello	Ryan	ryan costello,ryan a costello	1	change1	R	PA	C001106
Gaetz	Matt	matt gaetz	1	change1	R	FL	G000578
Gonzalez	Vicente	vicente gonzalez	1	change1	D	TX	G000581
Waltz	Michael	michael waltz	0	change1	R	FL	W000823

```r
df_subsample <- subset(df_subsample, select = -c(
    last_name, first_name, member_id, name, Vote_count, District,
    Vote3, Vote4, Vote51, Vote52, Vote6, Vote7, Vote_change, state, Vote_change_dummy
))

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

```
Call:
lm(formula = change ~ ., data = df_subsample)

Residuals:
         1          2          3          4          6          7          8
 5.000e-01 -8.333e-01  3.333e-01  7.440e-16 -1.667e-01  5.000e-01 -3.333e-01
         9         11         16         19         21         22         23
-3.313e-17 -3.662e-16 -4.772e-16 -1.143e-15 -3.333e-01  3.333e-01 -3.313e-17

Coefficients: (24 not defined because of singularities)
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        1.279e-01  7.321e-01   0.175    0.872
voteschange2       3.333e-01  6.086e-01   0.548    0.622
voteschange3       1.667e-01  7.136e-01   0.234    0.830
voteschange4       8.333e-01  9.379e-01   0.889    0.440
partyR             1.335e+00  2.574e+00   0.519    0.640
amount.oil.113     1.382e-05  1.272e-05   1.087    0.357
amount.coal.113   -4.443e-04  6.692e-04  -0.664    0.554
amount.mining.113  7.962e-05  1.886e-04   0.422    0.701
amount.gas.113    -9.826e-05  1.222e-04  -0.804    0.480
amount.env.113    -7.094e-04  7.206e-04  -0.985    0.397
```

```
amount.alt_en.113  1.370e-03  1.596e-03   0.858    0.454
amount.oil.114            NA         NA      NA       NA
amount.coal.114           NA         NA      NA       NA
amount.mining.114         NA         NA      NA       NA
amount.gas.114            NA         NA      NA       NA
amount.env.114            NA         NA      NA       NA
amount.alt_en.114         NA         NA      NA       NA
amount.oil.115            NA         NA      NA       NA
amount.coal.115           NA         NA      NA       NA
amount.mining.115         NA         NA      NA       NA
amount.gas.115            NA         NA      NA       NA
amount.env.115            NA         NA      NA       NA
amount.alt_en.115         NA         NA      NA       NA
amount.oil.116            NA         NA      NA       NA
amount.coal.116           NA         NA      NA       NA
amount.mining.116         NA         NA      NA       NA
amount.gas.116            NA         NA      NA       NA
amount.env.116            NA         NA      NA       NA
amount.alt_en.116         NA         NA      NA       NA
amount.oil.117            NA         NA      NA       NA
amount.coal.117           NA         NA      NA       NA
amount.mining.117         NA         NA      NA       NA
amount.gas.117            NA         NA      NA       NA
amount.env.117            NA         NA      NA       NA
amount.alt_en.117         NA         NA      NA       NA

Residual standard error: 0.7454 on 3 degrees of freedom
  (17 observations deleted due to missingness)
Multiple R-squared:  0.4167,    Adjusted R-squared:  -1.528
F-statistic: 0.2143 on 10 and 3 DF,  p-value: 0.9726
```

I tried regressing the change variable with each individual contributions, i.e. amount.coal.113, amount.oil.113, etc. but the results were neither significant, nor did they often appear. many had NAs.

```r
df_subsample_sum <- subset(df_subsample_sum, select = -c(
    last_name, first_name, member_id, name, Vote_count, District,
    Vote3, Vote4, Vote51, Vote52, Vote6, Vote7, Vote_change, state, Vote_change_dummy
))

ols_subsample_sum <- lm(change ~ ., data = df_subsample_sum)
summary(ols_subsample_sum)
```

```
Call:
lm(formula = change ~ ., data = df_subsample_sum)

Residuals:
    Min      1Q  Median      3Q     Max
-0.8396 -0.2432  0.1327  0.2473  0.7695

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.094e-01  3.045e-01   2.001   0.0591 .
voteschange2 -3.492e-01  2.584e-01  -1.352   0.1916
voteschange3 -2.006e-01  4.314e-01  -0.465   0.6470
voteschange4  4.203e-01  5.719e-01   0.735   0.4710
partyR        1.433e-01  3.070e-01   0.467   0.6458
oil_sum       3.967e-07  1.059e-06   0.374   0.7121
coal_sum     -5.561e-06  4.570e-05  -0.122   0.9044
gas_sum       2.314e-06  5.609e-06   0.413   0.6843
mining_sum   -4.936e-07  2.015e-05  -0.024   0.9807
env_sum      -2.517e-07  3.498e-05  -0.007   0.9943
alt_en_sum   -3.930e-06  8.332e-06  -0.472   0.6422
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4789 on 20 degrees of freedom
Multiple R-squared:  0.2273,    Adjusted R-squared:  -0.1591
F-statistic: 0.5882 on 10 and 20 DF,  p-value: 0.8049
```

So i tried to regress the change with the summarised variables, i.e. oil_sum, coal_sum, etc. and now the variables aren't NA, just not significant.

1.0.1 OLS vote change and party affiliation

```
# ols_1 <- lm(Vote_change_dummy ~ party, data = df)
# summary(ols_1)

# ols_2 <- lm(Vote_change_dummy ~ party, data = df_subsample_sum)
# summary(ols_2)
```

regressing the vote change dummy on party affiliation, we find that the coefficient for the Republican party is 0.2, the intercept shows how likely a democrat is to change their voting

behaviour, at 0.31, we can see that democrats are more likely to change their voting behaviour on methane related issues, with only the democrat variable at 0.05 being significant.

1.0.2 MODEL 3: Subsample OLS with (time and state) Fixed Effects

```
# remove cols not necessary for analysis
df_fe <- df_fe %>% select(-c(
    last_name, first_name, name, District, Vote_count, Vote_change,
    first_contribution, Vote_change, Vote_change_dummy, member_id, Vote3, Vote4, Vote51, Vote
))
view(df_fe)

fixed <- plm(vote_change_type ~ ., data = df_fe, index = c("state", "year"), model = "within"
```

```
Warning in pdata.frame(data, index): duplicate couples (id-time) in resulting pdata.frame
 to find out which, use, e.g., table(index(your_pdataframe), useNA = "ifany")
```

```
summary(fixed)
```

```
Oneway (individual) effect Within Model

Call:
plm(formula = vote_change_type ~ ., data = df_fe, model = "within",
    index = c("state", "year"))

Unbalanced Panel: n = 6, T = 1-4, N = 14

Residuals:
          1           4           5           6          10          11
 5.3333e-01 -5.3333e-01  1.7222e-16  2.3700e-16 -4.0000e-01 -1.3333e-01
         12          13          15          20          26          27
 7.7843e-17  5.3333e-01  0.0000e+00  0.0000e+00 -1.3333e-01 -1.9971e-16
         28          30
 1.3333e-01 -8.8690e-17

Coefficients: (30 dropped because of singularities)
                             Estimate  Std. Error t-value Pr(>|t|)
vote_change_year2016-2017  7.3333e-01  7.0553e-01  1.0394   0.4078
vote_change_year2017-2018 -4.0000e-01  9.2376e-01 -0.4330   0.7072
vote_change_year2018-2019  6.6667e-02  7.0553e-01  0.0945   0.9333
```

```
vote_change_year2019-2021  5.3333e-01  9.6148e-01  0.5547   0.6349
amount.oil.113            -1.7760e-05  1.9702e-05 -0.9014   0.4625
amount.mining.113          4.8437e-04  5.0413e-04  0.9608   0.4380

Total Sum of Squares:    2.4167
Residual Sum of Squares: 1.0667
R-Squared:      0.55862
Adj. R-Squared: -1.869
F-statistic: 0.421875 on 6 and 2 DF, p-value: 0.82568
```

```
knitr::kable(df_fe, format = "html")
```

party	state	vote_change_type	vote_change_year	year	amount.oil.113	amount.coal.113	amount.m
D	CA	1	2013-2016	2016	82750	5000	
D	CA	0	2018-2019	2019	82750	5000	
D	CA	1	2019-2021	2021	82750	5000	
R	CO	1	2017-2018	2018	210275	12750	
R	FL	1	2013-2016	2016	NA	NA	
D	GA	0	2013-2016	2016	5500	0	
D	GA	1	2016-2017	2017	5500	0	
D	GA	0	2017-2018	2018	5500	0	
D	GA	1	2018-2019	2019	5500	0	
R	IL	1	2017-2018	2018	42250	2000	
R	MI	1	2017-2018	2018	224300	44633	
R	NJ	1	2016-2017	2017	NA	NA	
R	NJ	0	2019-2021	2021	NA	NA	
R	NJ	1	2017-2018	2018	21050	0	
R	NY	1	2013-2016	2016	NA	NA	
R	NY	1	2019-2021	2021	120021	0	
R	PA	1	2016-2017	2017	50000	1500	
R	SC	1	2016-2017	2017	NA	NA	
D	TX	1	2018-2019	2019	99125	0	
D	TX	1	2013-2016	2016	96700	0	
D	TX	0	2013-2016	2016	62350	0	
D	TX	1	2016-2017	2017	62350	0	
D	TX	1	2013-2016	2016	13550	0	
R	WA	1	2016-2017	2017	NA	NA	
R	CA	1	2016-2017	2017	NA	NA	
R	CA	0	2017-2018	2018	NA	NA	
R	NY	0	2019-2021	2021	NA	NA	
R	PA	1	2016-2017	2017	NA	NA	

party	state	vote_change_type	vote_change_year	year	amount.oil.113	amount.coal.113	amount.m
R	FL	1	2019-2021	2021	NA	NA	
D	TX	1	2017-2018	2018	NA	NA	
R	FL	0	2019-2021	2021	NA	NA	

time & state fixed effects – This model eliminates omitted variable bias caused by excluding unobserved variables that evolve over time but are constant across entities.

fixed effects remove the effect of those time-invariant characteristics so we can assess the net effect of the predictors on the outcome variable. I.e. we fix time and thus only look at one time period each.

1.0.3 MODEL 4: Probit and Logit for Robustness