# The Ideal Electoral System

Daniel K Baissa

3/28/2022

## Electoral Engineering

Today we have access to machine learning models that can help use greatly improve our ability to develop electoral systems.

## Carey and Hix Replication

I will start by replicating a model from the Carry and Hix dataset in order to ensure that the data and models line up before improving the specification.

### The first model from Carey and Hix 2011

Here I will replicate the coefficients for their model and omit the Panel Corrected Standard Errors

```
CH <- df %>%
  select(disprop, dist_mag_medians, dm_asym, pres, legal_thresh, MMPL, compensatory,
         ethnic_fract_fearon, hybrid , election_yr, pol_freedom, econ_freedom,
         population, gdp_head, growth , age_dem, federal, latitude, col_uk,
         col_sp_po, col_oth, americas, former_com, pacific, s_asia,  gini, africa_me)

m.lm <- lm(disprop ~ ., data = CH)
summary(m.lm)
```

```
##
## Call:
## lm(formula = disprop ~ ., data = CH)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.6826  -2.1786  -0.3893   1.4741  21.9200
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)         -2.444e+02  3.592e+01  -6.803 2.55e-11 ***
## dist_mag_medians    -8.296e-03  7.142e-03  -1.161 0.245929
## dm_asym              1.005e+01  7.573e-01  13.269  < 2e-16 ***
## pres                -3.524e+00  7.741e-01  -4.553 6.45e-06 ***
## legal_thresh         3.709e-01  1.431e-01   2.592 0.009778 **
```

```
## MMPL                 -2.676e+00  1.064e+00  -2.516 0.012132 *
## compensatory          7.740e-01  8.842e-01   0.875 0.381754
## ethnic_fract_fearon   2.714e+00  9.936e-01   2.731 0.006503 **
## hybrid                6.055e-01  6.138e-01   0.986 0.324298
## election_yr           1.234e-01  1.796e-02   6.871 1.64e-11 ***
## pol_freedom           3.674e-01  2.283e-01   1.609 0.108122
## econ_freedom         -2.367e-01  4.164e-01  -0.568 0.570050
## population           -8.889e-03  2.285e-03  -3.890 0.000112 ***
## gdp_head             -2.185e-01  6.890e-02  -3.171 0.001601 **
## growth               -9.004e-02  5.024e-02  -1.792 0.073631 .
## age_dem              -5.024e-02  1.398e-02  -3.594 0.000353 ***
## federal              -1.144e+00  6.500e-01  -1.760 0.078945 .
## latitude              5.973e+00  2.286e+00   2.613 0.009197 **
## col_uk                7.579e-01  6.666e-01   1.137 0.256010
## col_sp_po             2.572e+00  9.154e-01   2.810 0.005122 **
## col_oth               1.987e+00  7.060e-01   2.814 0.005054 **
## americas              2.334e+00  8.707e-01   2.681 0.007548 **
## former_com           -1.960e+00  1.004e+00  -1.953 0.051321 .
## pacific               3.301e+00  9.394e-01   3.514 0.000476 ***
## s_asia                5.887e+00  1.117e+00   5.271 1.91e-07 ***
## gini                 -5.597e-02  2.688e-02  -2.082 0.037797 *
## africa_me             1.450e+00  1.046e+00   1.387 0.166053
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.227 on 582 degrees of freedom
## Multiple R-squared:  0.5614, Adjusted R-squared:  0.5419
## F-statistic: 28.66 on 26 and 582 DF,  p-value: < 2.2e-16
```

### Enter BART

Now let's run Carey and Hix's model using Bayesian Additive Regression Trees. The key difference is that we are going to remove their asymptotic term from the model since BART will automatically allow for curvilinear fits. Then we will rename the y variable "y".

```
X <- df %>%
  select(disprop,
         dist_mag_medians, pres, legal_thresh, MMPL, compensatory,
         ethnic_fract_fearon, hybrid , election_yr, pol_freedom, econ_freedom,
         population, gdp_head, growth , age_dem, federal, latitude, col_uk,
         col_sp_po, col_oth, americas, former_com, pacific, s_asia,  gini, africa_me) %>%
  na.omit()

y <- X$disprop

X <- X %>%
  select(!disprop)
```

```
m.BART <- bartMachine(X = X, y = y)
```

```
## bartMachine initializing with 50 trees...
## bartMachine vars checked...
```

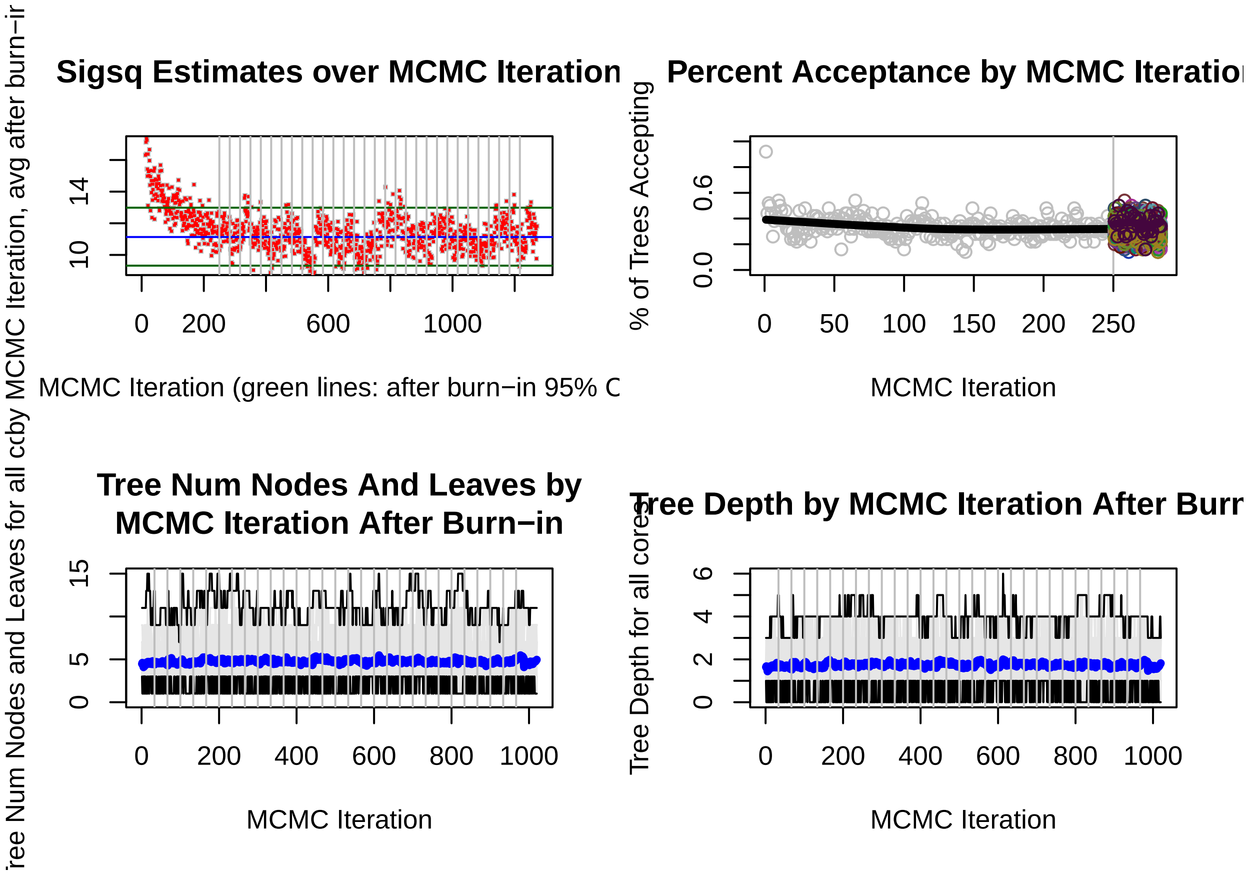
```
## bartMachine java init...
## bartMachine factors created...
## bartMachine before preprocess...
## bartMachine after preprocess... 26 total features...
## bartMachine sigsq estimated...
## bartMachine training data finalized...
## Now building bartMachine for regression...
## evaluating in sample data...done
```

```
m.BART
```

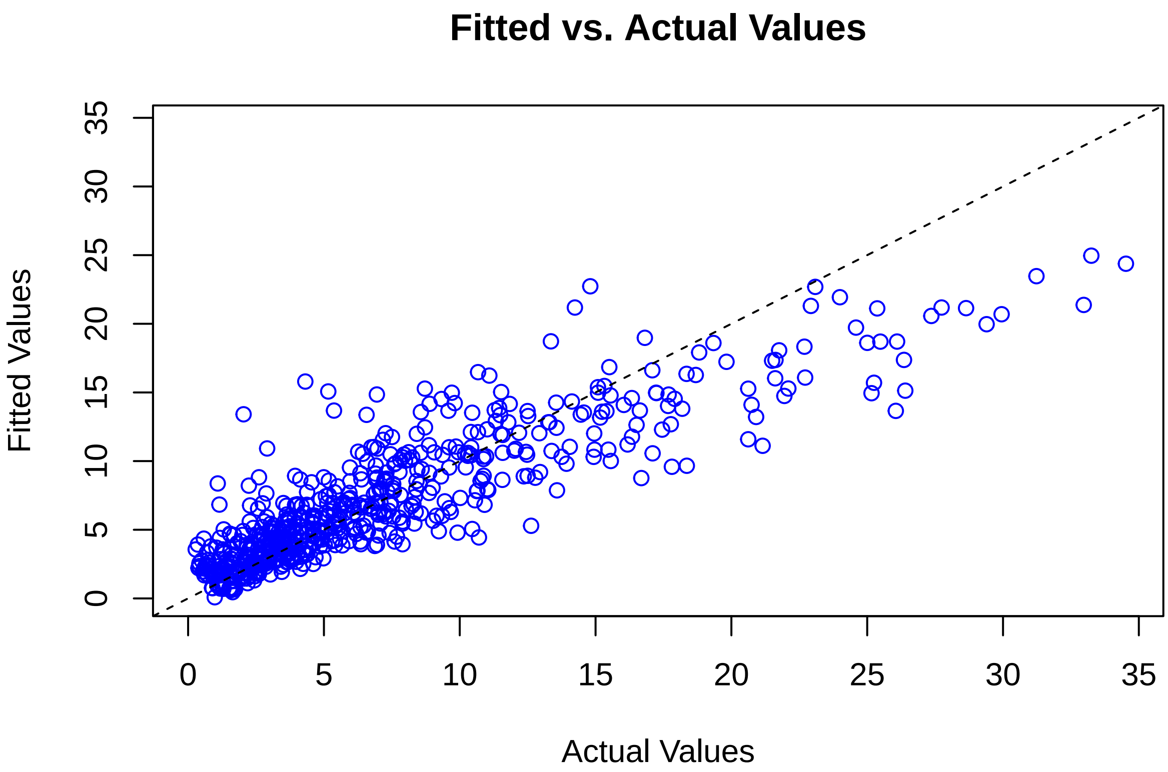
```
## bartMachine v1.2.6 for regression
##
## training data n = 609 and p = 25
## built in 7 secs on 30 cores, 50 trees, 250 burn-in and 1000 post. samples
##
## sigsq est for y beforehand: 22.281
## avg sigsq estimate after burn-in: 11.13047
##
## in-sample statistics:
##  L1 = 1269.5
##  L2 = 5521.96
##  rmse = 3.01
##  Pseudo-Rsq = 0.7672
## p-val for shapiro-wilk test of normality of residuals: 0
## p-val for zero-mean noise: 0.89028
```

We can see here that the Pseudo R Squared for the BART model is significantly higher than that of their OLS model.

```
plot_convergence_diagnostics(m.BART)
```

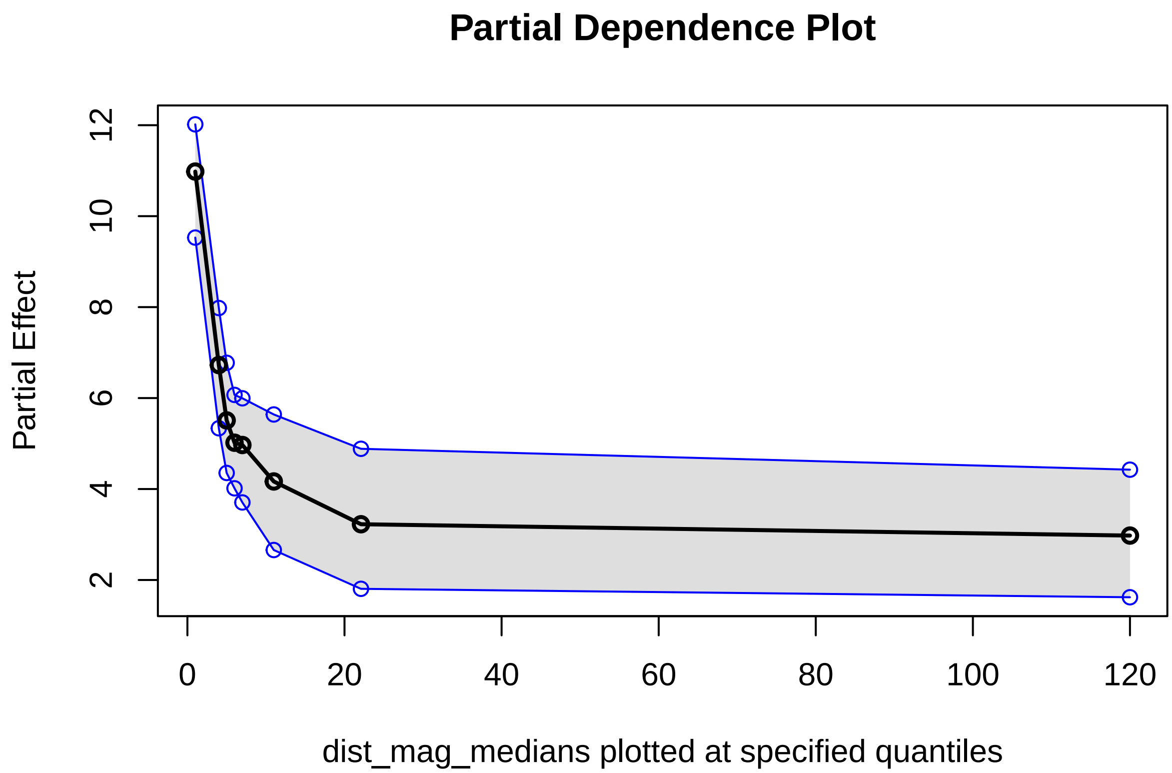



```
plot_y_vs_yhat(m.BART)
```




```r
pd_plot(m.BART,  m.BART[["training_data_features"]][1])
```

```
## ........
```




Great! The main effect mirrors the researcher's findings very well, but with much wider confidence intervals.

## Optimal Governance

Carey and Hix focused used their models to estimate which district magnitudes would be best for optomizing governance. This means for example, they took the median of the `disprop` variable and then said values below the median were idea and those above where less than ideal.

Let's try something a bit more varied. Let's try spliting the data into 3rds and then take the bottom third as ideal.

```
X <- df %>%
  select(disprop,
         dist_mag_medians, pres, legal_thresh, MMPL, compensatory,
         ethnic_fract_fearon, hybrid , election_yr, pol_freedom, econ_freedom,
         population, gdp_head, growth , age_dem, federal, latitude, col_uk,
         col_sp_po, col_oth, americas, former_com, pacific, s_asia,  gini, africa_me) %>%
  mutate(ideal_disprop = ifelse(disprop < quantile(disprop, probs = seq(0, 1, by = 1/3))[[2]], 1, 0)) %
  na.omit()

y <- as.factor(X$ideal_disprop)

X <- X %>%
  select(!c(disprop, ideal_disprop))
```

```r
m.BART <- bartMachine(X = X, y = y)
```

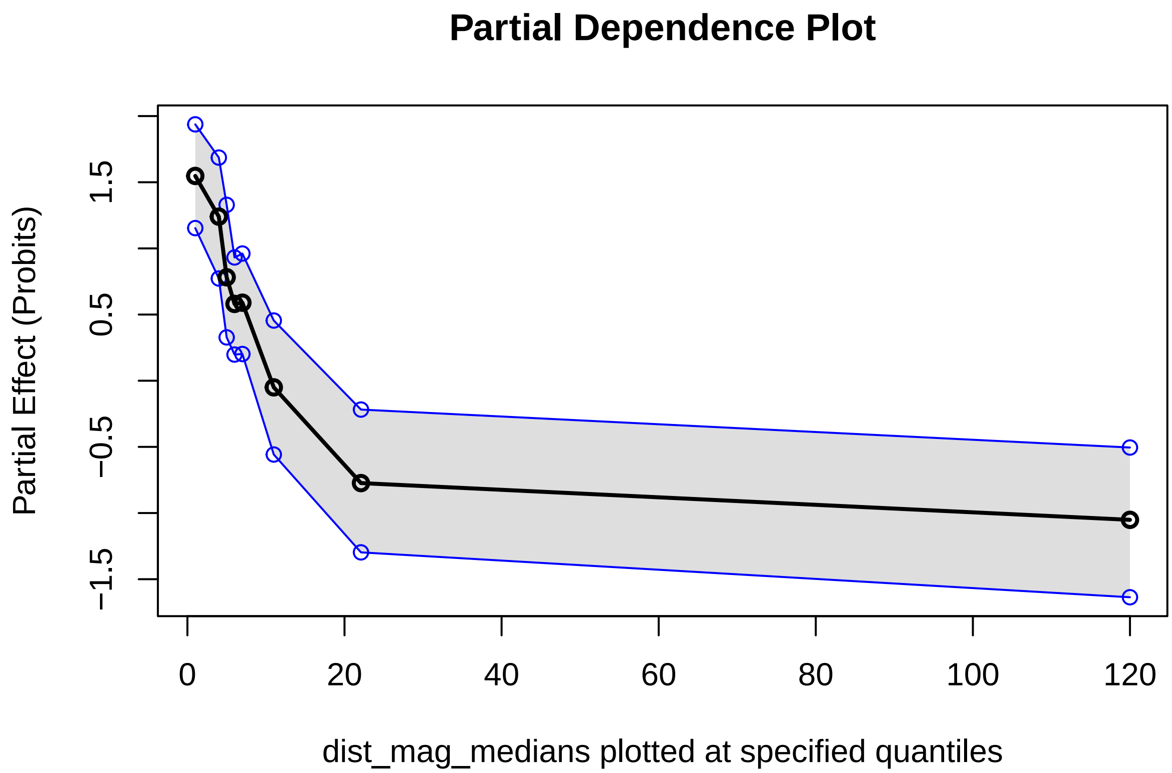
```
## bartMachine initializing with 50 trees...
## bartMachine vars checked...
## bartMachine java init...
## bartMachine factors created...
## bartMachine before preprocess...
## bartMachine after preprocess... 26 total features...
## bartMachine sigsq estimated...
## bartMachine training data finalized...
## Now building bartMachine for classification where "0" is considered the target level...
## evaluating in sample data...done
```

```r
m.BART
```

```
## bartMachine v1.2.6 for classification
##
## training data n = 609 and p = 25
## built in 3.3 secs on 30 cores, 50 trees, 250 burn-in and 1000 post. samples
##
## confusion matrix:
##
##            predicted 0 predicted 1 model errors
## actual 0       384.000      22.000        0.054
## actual 1        59.000     144.000        0.291
## use errors       0.133       0.133        0.133
```

```r
pd_plot(m.BART,  m.BART[["training_data_features"]][1])
```

```
## ........
```




Here we see that a high district magnitude will be needed holding all else equal.

```
y_hat_class = predict(m.BART, X, type = "class")


test <- cbind(df, "Predicted Ideal" = y_hat_class) %>%
  select(country, `Predicted Ideal`, election_yr, growth, gini, federal, pres)
```

### Accountability

```
X <- df %>%
  select(enps,
         dist_mag_medians, pres, legal_thresh, MMPL, compensatory,
         ethnic_fract_fearon, hybrid , election_yr, pol_freedom, econ_freedom,
         population, gdp_head, growth , age_dem, federal, latitude, col_uk,
         col_sp_po, col_oth, americas, former_com, pacific, s_asia,  gini, africa_me) %>%
  mutate(ideal_enps = ifelse(enps < quantile(enps, probs = seq(0, 1, by = 1/3))[[2]], 1, 0)) %>%
  na.omit()

y <- as.factor(X$ideal_enps)

X <- X %>%
  select(!c(enps, ideal_enps))
```

```r
m.BART <- bartMachine(X = X, y = y)
```

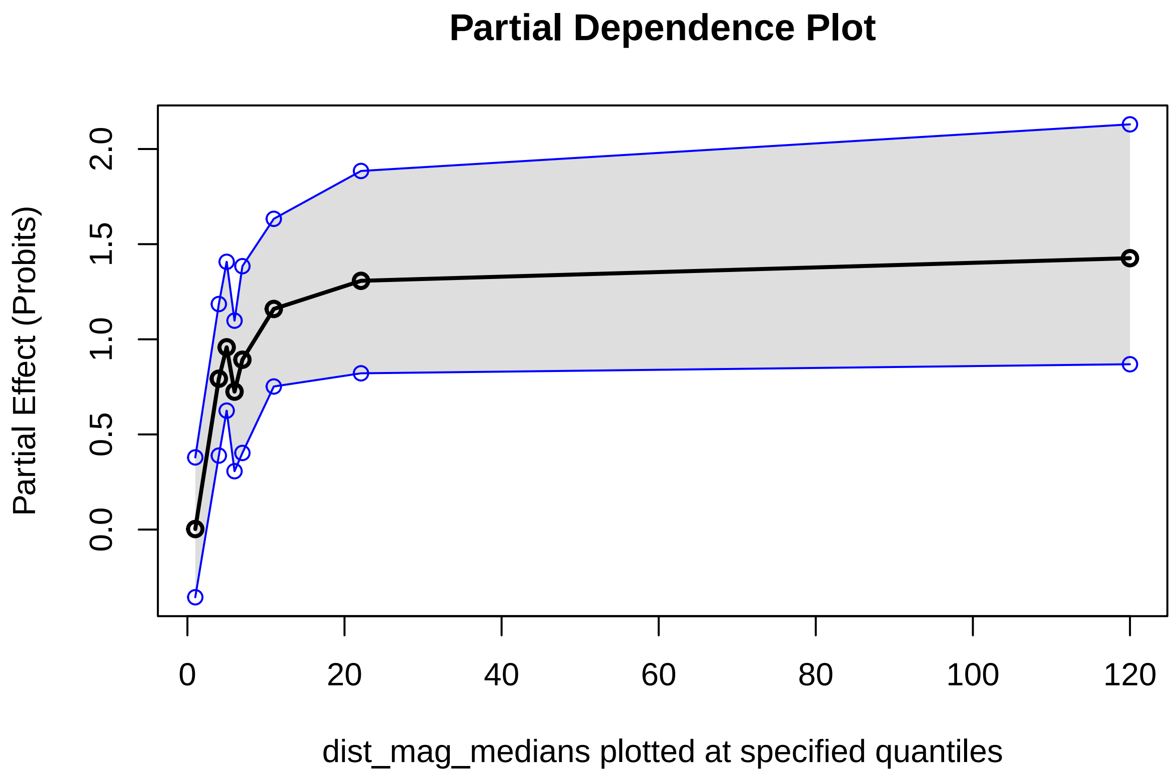
```
## bartMachine initializing with 50 trees...
## bartMachine vars checked...
## bartMachine java init...
## bartMachine factors created...
## bartMachine before preprocess...
## bartMachine after preprocess... 26 total features...
## bartMachine sigsq estimated...
## bartMachine training data finalized...
## Now building bartMachine for classification where "0" is considered the target level...
## evaluating in sample data...done
```

```r
m.BART
```

```
## bartMachine v1.2.6 for classification
##
## training data n = 609 and p = 25
## built in 2.8 secs on 30 cores, 50 trees, 250 burn-in and 1000 post. samples
##
## confusion matrix:
##
##            predicted 0 predicted 1 model errors
## actual 0       389.000       17.00        0.042
## actual 1        65.000      138.00        0.320
## use errors       0.143        0.11        0.135
```

```r
pd_plot(m.BART,  m.BART[["training_data_features"]][1])
```

```
## ........
```




## Are there "Ideal" countries?

Let's say we want better than median Accountability and better than 2/3rds of representation

```
X <- df %>%
  select(enps, disprop, country,
         dist_mag_medians, pres, legal_thresh, MMPL, compensatory,
         ethnic_fract_fearon, hybrid , election_yr, pol_freedom, econ_freedom,
         population, gdp_head, growth , age_dem, federal, latitude, col_uk,
         col_sp_po, col_oth, americas, former_com, pacific, s_asia,  gini, africa_me) %>%
  mutate(ideal_enps = ifelse(enps < quantile(enps, probs = seq(0, 1, by = 1/2))[[2]], 1, 0)) %>%
  mutate(ideal_disprop = ifelse(enps < quantile(disprop, probs = seq(0, 1, by = 1/3))[[2]], 1, 0)) %>%
  # select(country, ideal_enps, ideal_disprop) %>%
  mutate(ideal = ifelse(ideal_enps == 1 & ideal_disprop == 1, 1, 0)) %>%
  na.omit()

y <- as.factor(X$ideal)

X <- X %>%
  select(!c(enps, ideal_enps, ideal_disprop, ideal, country, disprop))
```

```
m.BART <- bartMachine(X = X, y = y)
```

```
## bartMachine initializing with 50 trees...
```

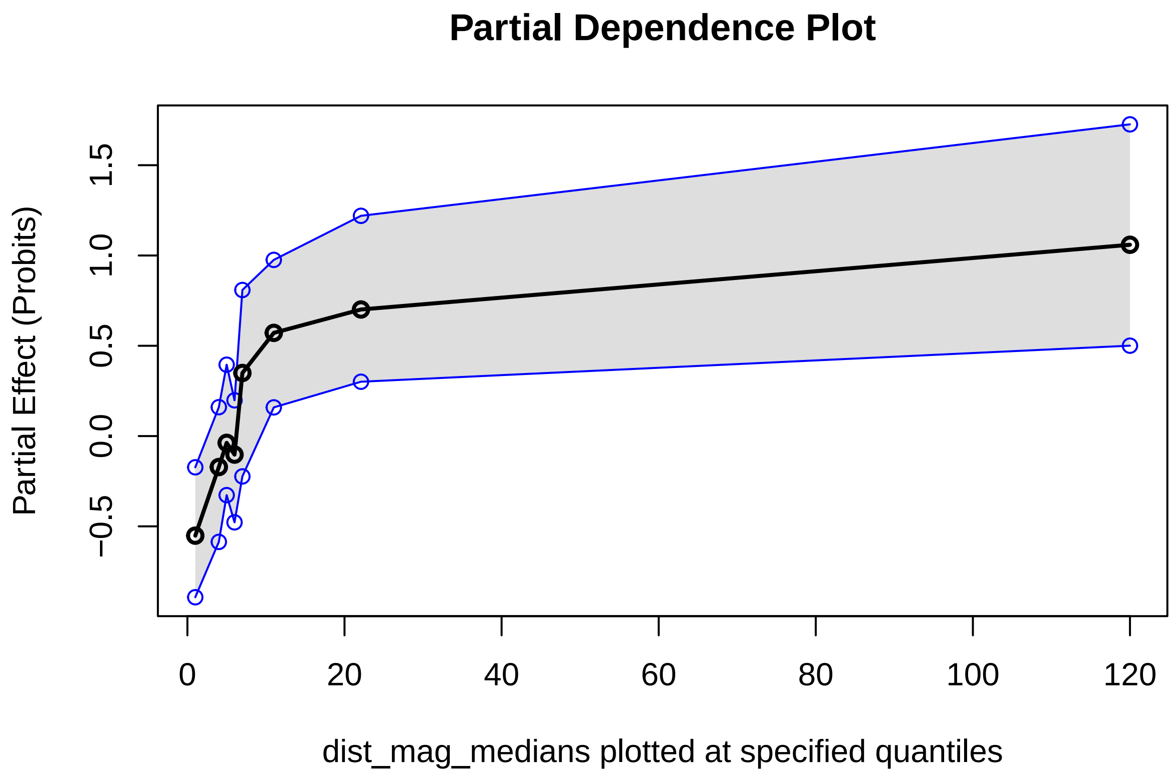
```
## bartMachine vars checked...
## bartMachine java init...
## bartMachine factors created...
## bartMachine before preprocess...
## bartMachine after preprocess... 26 total features...
## bartMachine sigsq estimated...
## bartMachine training data finalized...
## Now building bartMachine for classification where "0" is considered the target level...
## evaluating in sample data...done
```

```
m.BART
```

```
## bartMachine v1.2.6 for classification
##
## training data n = 609 and p = 25
## built in 3.4 secs on 30 cores, 50 trees, 250 burn-in and 1000 post. samples
##
## confusion matrix:
##
##            predicted 0 predicted 1 model errors
## actual 0        259.00      46.000        0.151
## actual 1         42.00     262.000        0.138
## use errors        0.14       0.149        0.144
```

```
pd_plot(m.BART,  m.BART[["training_data_features"]][1])
```

```
## ........
```




Here we can see that District Magnitude is Significant and there is no reason to limit the threshold to an arbitrarily low number of seats.

## Can we improve a country?

Let's take Chile, can we improve Chile?

```
Chile <- df %>%
  select(enps, disprop, country,
         dist_mag_medians, pres, legal_thresh, MMPL, compensatory,
         ethnic_fract_fearon, hybrid , election_yr, pol_freedom, econ_freedom,
         population, gdp_head, growth , age_dem, federal, latitude, col_uk,
         col_sp_po, col_oth, americas, former_com, pacific, s_asia,  gini, africa_me) %>%
  mutate(ideal_enps = ifelse(enps < quantile(enps, probs = seq(0, 1, by = 1/2))[[2]], 1, 0)) %>%
  mutate(ideal_disprop = ifelse(enps < quantile(disprop, probs = seq(0, 1, by = 1/3))[[2]], 1, 0)) %>%
  # select(country, ideal_enps, ideal_disprop) %>%
  mutate(ideal = ifelse(ideal_enps == 1 & ideal_disprop == 1, 1, 0)) %>%
  filter(country == "Chile")
```

Let's confirm that Chile is not seen as an ideal country.

```
Chile_changed = predict(m.BART, X, type = "class")

Chile_changed <- cbind(Chile_changed, X)
```

```
Chile_changed %>%
  filter(ethnic_fract_fearon < 0.498) %>%
  filter(ethnic_fract_fearon > 0.496)
```

```
##     Chile_changed dist_mag_medians pres legal_thresh MMPL compensatory
## 126             0                2    1            0    0            0
## 127             0                2    1            0    0            0
## 128             0                2    1            0    0            0
##     ethnic_fract_fearon hybrid election_yr pol_freedom econ_freedom population
## 126               0.497      0        1993           8            2   13.91608
## 127               0.497      0        1997           8            2   14.82768
## 128               0.497      0        2001           9            2   15.59634
##      gdp_head   growth age_dem federal  latitude col_uk col_sp_po col_oth
## 126  7.905548 9.081370       5       0 0.3153333      0         1       0
## 127  9.997491 7.470669       9       0 0.3153333      0         1       0
## 128 10.538120 2.368420      13       0 0.3153333      0         1       0
##     americas former_com pacific s_asia gini africa_me
## 126        1          0       0      0 48.9         0
## 127        1          0       0      0 48.9         0
## 128        1          0       0      0 48.9         0
```

It's not.

Now let's see what would happen if we changed their electoral rules to be more representative.

```
X2 <- X %>%
  mutate(dist_mag_medians = ifelse(ethnic_fract_fearon < 0.498 & ethnic_fract_fearon > 0.496,
                                   20 , dist_mag_medians))
```

```
Chile_changed = predict(m.BART, X2, type = "class")

Chile_changed <- cbind(Chile_changed, X2)

Chile_changed %>%
  filter(ethnic_fract_fearon < 0.498) %>%
  filter(ethnic_fract_fearon > 0.496)
```

```
##     Chile_changed dist_mag_medians pres legal_thresh MMPL compensatory
## 126             0               20    1            0    0            0
## 127             0               20    1            0    0            0
## 128             0               20    1            0    0            0
##     ethnic_fract_fearon hybrid election_yr pol_freedom econ_freedom population
## 126               0.497      0        1993           8            2   13.91608
## 127               0.497      0        1997           8            2   14.82768
## 128               0.497      0        2001           9            2   15.59634
##      gdp_head   growth age_dem federal  latitude col_uk col_sp_po col_oth
## 126  7.905548 9.081370       5       0 0.3153333      0         1       0
## 127  9.997491 7.470669       9       0 0.3153333      0         1       0
## 128 10.538120 2.368420      13       0 0.3153333      0         1       0
##     americas former_com pacific s_asia gini africa_me
## 126        1          0       0      0 48.9         0
## 127        1          0       0      0 48.9         0
## 128        1          0       0      0 48.9         0
```

Its not enough…

Let's also add federlaism?

```
X2 <- X %>%
  # mutate(dist_mag_medians = ifelse(ethnic_fract_fearon < 0.498 & ethnic_fract_fearon > 0.496,
  #                                  6 , dist_mag_medians))
  # # %>%
  mutate(federal = ifelse(ethnic_fract_fearon < 0.498 & ethnic_fract_fearon > 0.496,
                          1 , federal))

Chile_changed = predict(m.BART, X2, type = "class")

Chile_changed <- cbind(Chile_changed, X2)

Chile_changed %>%
  filter(ethnic_fract_fearon < 0.498) %>%
  filter(ethnic_fract_fearon > 0.496)
```

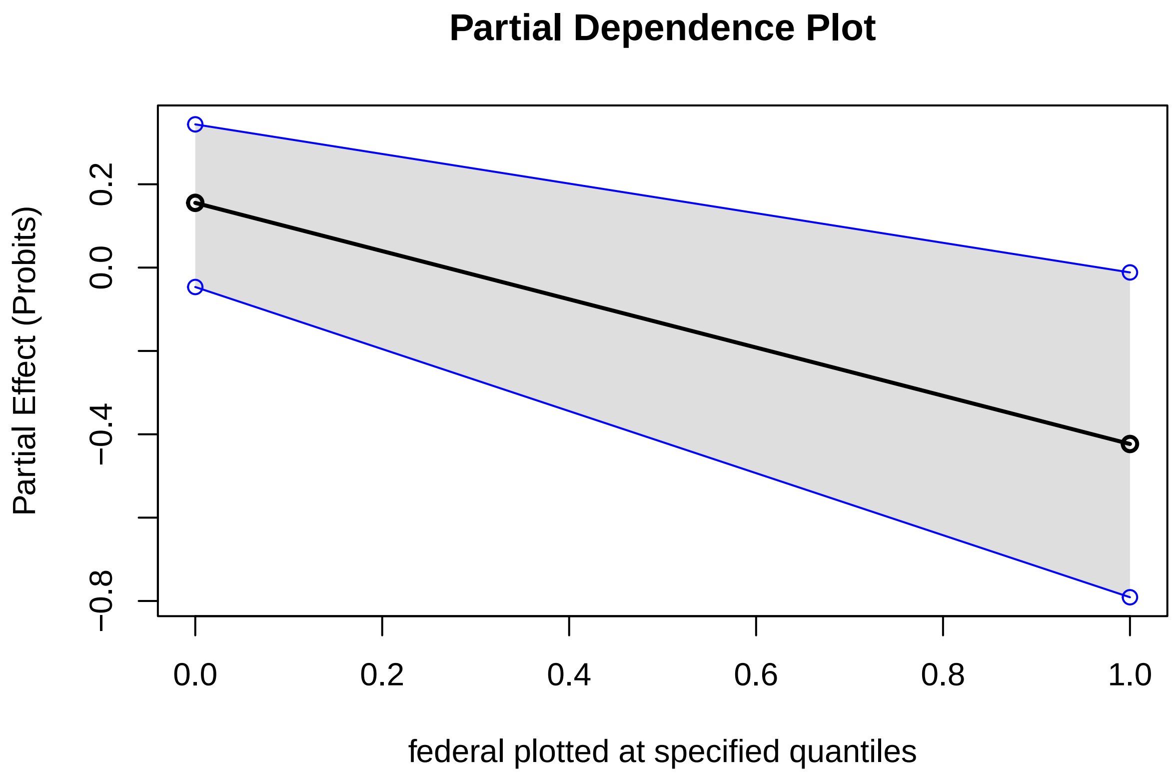
```
##     Chile_changed dist_mag_medians pres legal_thresh MMPL compensatory
## 126             1                2    1            0    0            0
## 127             1                2    1            0    0            0
## 128             1                2    1            0    0            0
##     ethnic_fract_fearon hybrid election_yr pol_freedom econ_freedom population
## 126               0.497      0        1993           8            2   13.91608
## 127               0.497      0        1997           8            2   14.82768
## 128               0.497      0        2001           9            2   15.59634
##      gdp_head   growth age_dem federal  latitude col_uk col_sp_po col_oth
## 126  7.905548 9.081370       5       1 0.3153333      0         1       0
## 127  9.997491 7.470669       9       1 0.3153333      0         1       0
## 128 10.538120 2.368420      13       1 0.3153333      0         1       0
##     americas former_com pacific s_asia gini africa_me
## 126        1          0       0      0 48.9         0
## 127        1          0       0      0 48.9         0
## 128        1          0       0      0 48.9         0
```

Federalism seems to fix it?!?! Federalism is more important than magnitude.

Since that is the case, let's look at if Federalism is statistically significant.

```
pd_plot(m.BART,  m.BART[["training_data_features"]][15])
```

```
## ..
```




Interesting! On average Federalism makes things worse. In Chile, however, Federalism makes things better. This means that the BART model is finding tailored electoral and governmental rules for each country.