

# Data Analysis

*Matthew Brown*

## Multiple Regression Model 1

**Margin = Fundraising + Performance**

The initial multiple regression model compared the dependent variable (the margin of victory or loss in an election) to the independent variable (the share of spending) and a control variable (the margin received in the previous election). This model demonstrates that there is a statistically significant relationship between the dependent variable and both the independent variable and the control variable.

```
m1<-lm(Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI)
summary(m1)
```

```
##
## Call:
## lm(formula = Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.44144 -0.08010  0.00608  0.07869  0.39438
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.32686    0.01311  -24.936  <2e-16 ***
## Senate7$share_of_spent  0.59770    0.02330   25.653  <2e-16 ***
## Senate7$Previous_RPVI  0.21273    0.02449    8.686  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.126 on 532 degrees of freedom
## (103 observations deleted due to missingness)
## Multiple R-squared:  0.7481, Adjusted R-squared:  0.7472
## F-statistic: 790 on 2 and 532 DF, p-value: < 2.2e-16
```

```
coef(m1)
```

```
##              (Intercept) Senate7$share_of_spent  Senate7$Previous_RPVI
##              -0.3268644          0.5977026          0.2127322
```

## Multiple Regression Model 2

Margin = Fundraising + Performance + Change in National Political Climate

In order to reduce the residuals and produce a more predictive model, a second multiple regression model which was tested comparing the dependent variable to the independent variable and two control variables (the margin received in the previous election and the shift in the congressional popular vote between the previous election and the one being analyzed). This model demonstrated that there is statistical significance between the shift in national congressional popular vote and the margin of victory or loss. The median residual decreased by 17% and the maximum and minimum residuals both decreased in magnitude.

```
m2<-lm(Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI + Senate7$YPVI2)
summary(m2)
```

```
##
## Call:
## lm(formula = Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI +
##     Senate7$YPVI2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42969 -0.07439  0.00471  0.07931  0.36096
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.31826    0.01283  -24.798 < 2e-16 ***
## Senate7$share_of_spent  0.58155    0.02283   25.473 < 2e-16 ***
## Senate7$Previous_RPVI  0.23821    0.02423    9.830 < 2e-16 ***
## Senate7$YPVI2       0.34571    0.06121    5.648 2.64e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1225 on 531 degrees of freedom
## (103 observations deleted due to missingness)
## Multiple R-squared:  0.7624, Adjusted R-squared:  0.761
## F-statistic: 567.9 on 3 and 531 DF, p-value: < 2.2e-16
```

```
coef(m2)
```

```
##              (Intercept) Senate7$share_of_spent Senate7$Previous_RPVI
##              -0.3182581          0.5815470          0.2382058
##              Senate7$YPVI2
##              0.3457135
```

## Multiple Regression Model 3

Margin = Fundraising + Performance + Change in National Political Climate + Incumbency

In order to further reduce the residuals and produce a more predictive model, a third multiple regression model which was tested comparing the dependent variable to the independent variable and three control variables (the margin received in the previous election, the shift in the congressional popular vote between the previous election and the one being analyzed, and whether or not the candidate is an incumbent). This model demonstrated that there is no statistical significance between the incumbency of a candidate and the margin of victory or loss.

```
m3<-lm(Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI + Senate7$YPVI2 + Senate7$Inc)
summary(m3)
```

```
##
## Call:
## lm(formula = Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI +
##     Senate7$YPVI2 + Senate7$Inc)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42293 -0.06873  0.00346  0.07842  0.36135
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.31403    0.01294  -24.263  < 2e-16 ***
## Senate7$share_of_spent  0.54548    0.02834   19.250  < 2e-16 ***
## Senate7$Previous_RPVI  0.22718    0.02470    9.198  < 2e-16 ***
## Senate7$YPVI2       0.35822    0.06128    5.845 8.84e-09 ***
## Senate7$IncTRUE      0.03486    0.01633    2.135  0.0332 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1221 on 530 degrees of freedom
## (103 observations deleted due to missingness)
## Multiple R-squared:  0.7644, Adjusted R-squared:  0.7626
## F-statistic: 429.9 on 4 and 530 DF,  p-value: < 2.2e-16
```

```
coef(m3)
```

```
##              (Intercept) Senate7$share_of_spent Senate7$Previous_RPVI
##             -0.31403374           0.54548494           0.22717795
##             Senate7$YPVI2           Senate7$IncTRUE
##             0.35822045           0.03486359
```

## Multiple Regression Model 4

**Margin = Fundraising + Performance + Change in National Political Climate + Incumbency of Opponent**

In order to further reduce the residuals and produce a more predictive model, a third multiple regression model which was tested comparing the dependent variable to the independent variable and three control variables (the margin received in the previous election, the shift in the congressional popular vote between the previous election and the one being analyzed, and whether or not the opponent candidate is an incumbent). This model demonstrated that there is no statistical significance between the incumbency of a candidate's opponent and the margin of victory or loss.

```
m4<-lm(Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI + Senate7$YPVI2 + Senate7$AInc)
summary(m4)
```

```
##
## Call:
## lm(formula = Senate7$RPVI ~ Senate7$share_of_spent + Senate7$Previous_RPVI +
##     Senate7$YPVI2 + Senate7$AInc)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.44064 -0.07250  0.00370  0.07993  0.36454
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.30169    0.01940  -15.551 < 2e-16 ***
## Senate7$share_of_spent  0.56292    0.02808   20.047 < 2e-16 ***
## Senate7$Previous_RPVI  0.23125    0.02498    9.256 < 2e-16 ***
## Senate7$YPVI2      0.35188    0.06143    5.728 1.7e-08 ***
## Senate7$AIncTRUE   -0.01870    0.01642   -1.139  0.255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1225 on 530 degrees of freedom
## (103 observations deleted due to missingness)
## Multiple R-squared:  0.763, Adjusted R-squared:  0.7612
## F-statistic: 426.5 on 4 and 530 DF, p-value: < 2.2e-16
```

```
coef(m4)
```

```
##              (Intercept) Senate7$share_of_spent Senate7$Previous_RPVI
##              -0.30168973          0.56292113          0.23125249
##              Senate7$YPVI2          Senate7$AIncTRUE
##              0.35187639          -0.01869574
```