Introduction: In the 2000 Presidential Election, did victory elude Al Gore by the length of a butterfly…ballot? In this election, Palm Beach County, Florida adopted a new kind of ballot, known as the "Butterfly Ballot" because of its unique voting design. This new design may have misled voters to incorrectly vote for Buchanan, the Reform candidate, rather than Al Gore, the Democratic candidate. This would eventually cause the Presidential vote to tip to Republican in Florida, leading to a George Bush victory. In this analysis, I will study whether voters mistakenly voted for Buchanan because of ballot modification.

Discussion: In order to study this election, I have county-level election outcomes for all US counties across the country, including demographic data and previous election voting tendencies. I also have data regarding the precincts within Palm Beach County (PBC), and the frequency of votes for the candidates in each. I've chosen three approaches, beginning with an analysis of the entire country, and then proceeding with an analysis of the state of Florida, and then ending with an analysis of Palm Beach County, in order to determine if the Butterfly Ballot misled voters.

Initially, I will create a model analyzing votes for Buchanan across all counties across the country. Because of the massive amounts of data from all the US counties, it will be difficult to see whether Palm Beach County is an outlier in the state of Florida, or even in the country. But because of this large amount of data, I can create a model that is very accurate in predicting the number of votes for Palm Beach County.

Second, in order to see whether Palm Beach County was an outlier in Florida, I will also create a model that only models the votes for Buchanan strictly across all counties in the state of Florida.

Third, in order to rule out the idea of there being a high number of individuals who are Pro-Reform Party in Palm Beach and to discover if the Butterfly Ballot misled voters, I will analyze the votes cast in the Palm Beach County Polls on Election Day and the Palm Beach County Absentee Ballots.

All Counties Analysis:

Below is a plot of the log transformed votes for Buchanan for each county versus the log transformed total number of votes per county. The highlighted red point is Palm Beach County. From this plot it appears the Palm Beach County is not an outlier, and is in line with the upward trend.

Figure 1.1

Votes for Buchanan by County
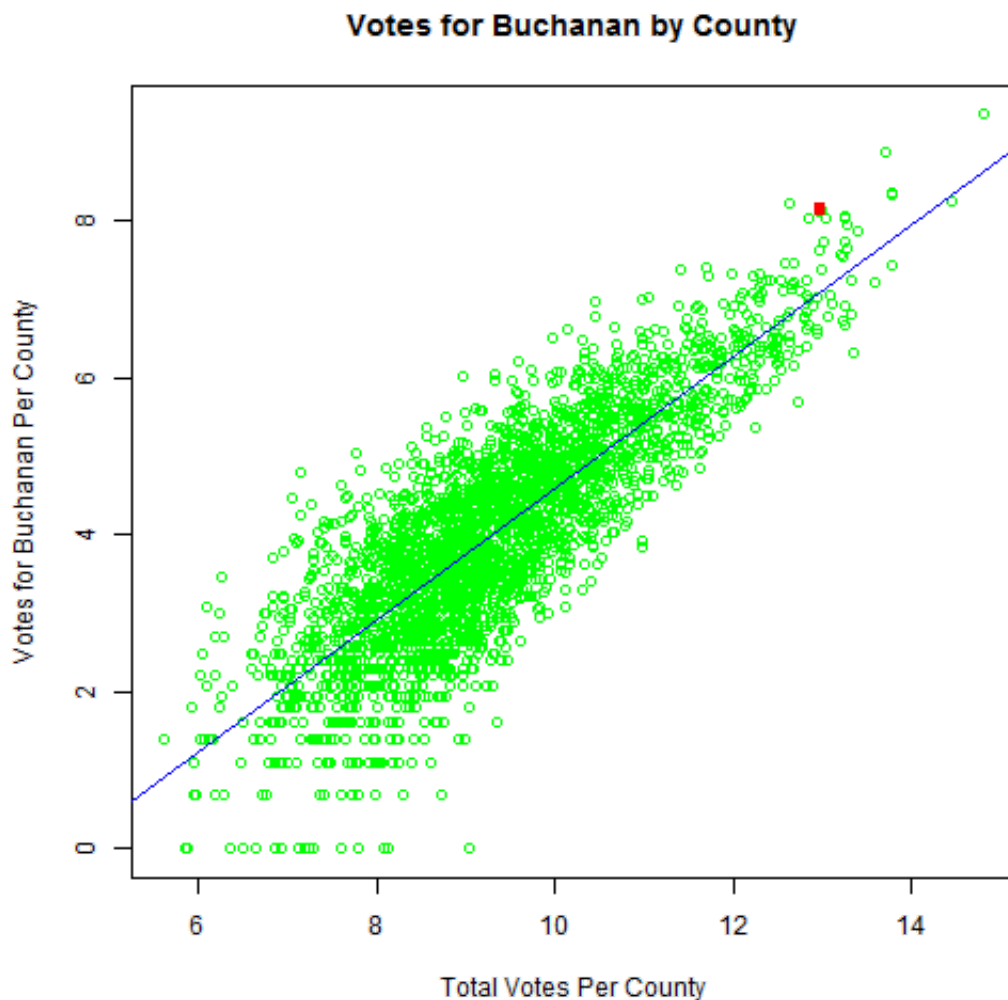
Before creating a model, I log transformed the total number of votes, the votes for Buchanan, the population size, income, density, and the percentage of Asians, such that the data could follow a more normal distribution. I removed 26 observations from the data set (0.85%), because some of the log transformed data resulted in values of infinity. Because of the small sample size, I do not see this impacting the regression results.

Model: I utilized backward selection to create the best model. This model regresses the number of Buchanan votes on the total number of votes, population, income, density, demographics (White, Asian, Native, Hispanic), the percentage of voters who voted for Perot in 1996 and 1992, and the percentage of voters who voted Republican in 1992. All of the coefficients are significant.

Below is a summary of the model.

Figure 1.2

```
##         Adjusted R^2    AIC    BIC
## Model        0.8078  -3085  -3013
```

As one can see from **Figure 1.2** above, this model has a very high adjusted R^2, and very low AIC and BIC values, and thus appears to be a good fit for all the data. I will next observe Cook's distance for any potential outliers and leverage points to improve the fit. A plot of the values for Cook's distance can be seen below (**Figure 1.3**).

Figure 1.3

Cook's Distance All Counties Model
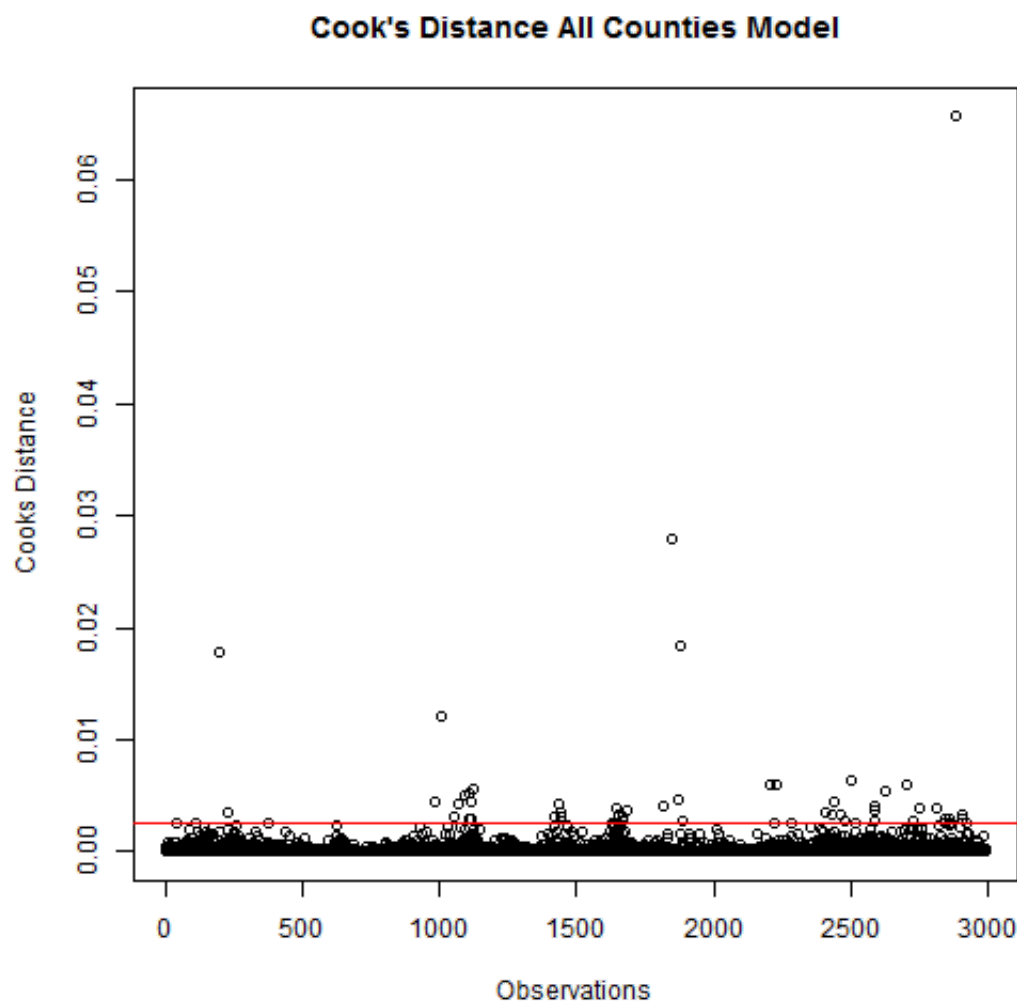
I have set a cutoff for hat values at 0.0025, as indicated by the red horizontal line. I will remove the points above this line, and then refit the model. Palm Beach County is not considered a high leverage point in this model as compared other counties across the United States. As noted in the discussion, because of the mass amount of data however, it might be difficult to see if this is an outlier in the state of Florida. However, the model we fitted will be more accurate because of the large quantity of data.

```
##              allmod$coeff
## (Intercept)    2.228e+00
## total          1.240e+00
## pop           -2.994e-01
## inc           -2.506e-01
## density        2.465e-05
## white         -2.769e-01
## asian          1.915e-02
## native         1.189e-01
## hispanic      -9.846e-02
## pperot96       8.751e-01
## pperot92       3.412e-01
## prep92         1.679e-01
```

Now based on the new coefficients for the model, the predicted number of the actual number of votes for Buchanan from Palm Beach County is approximately 1558, which is well short of the registered 3,014 votes for Buchanan. Because of the difference of 1456 votes between the actual and the fitted values, there is belief that maybe that the Butterfly Ballots did affect voting results.

But in order to see if this difference in actual and fitted values is significant and unusual, I will need to compare these results to other counties strictly within the state of Florida.

State of Florida:

Below is a plot of the number of votes for Buchanan in Florida by county.

Figure 1.4

Votes for Buchanan in Florida by County
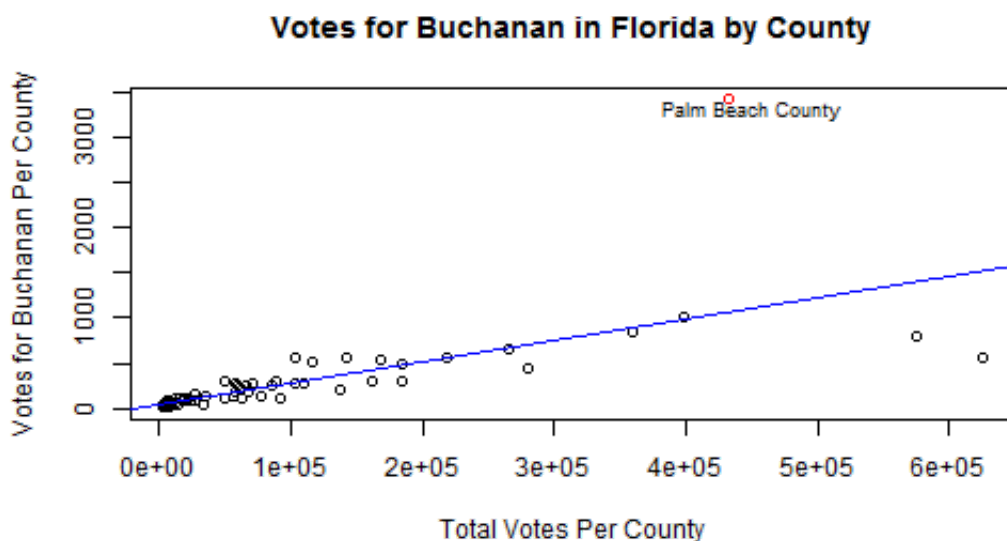
Analysis: Before creating the model for votes for Buchanan in Florida, I log transformed many of the demographic variables, so that the data could assume a more normal distribution. These included the number of votes for Buchanan, the total number of votes, population, income, density, white, black, asian, native, and hispanic.

Model:

I utilized backward selection to create the best model, and utilized a deviance table to remove any insignificant coefficients. Thus I ended up choosing a model that regressed the votes for Buchanan on Hispanic and Native percentage, the total number of votes cast, income, and the percentage that voted for Perot 1992 and 1996. I removed the coefficients for percentage of Asians and individuals with a college degree from the model because they would not improve the fit.

Thus, with this model, I will analyze the studentized residuals and look for observations with high values for Cook's distance. Based on the initial analysis performed, I will be expecting Palm Beach County to be an outlier.

Figure 1.5
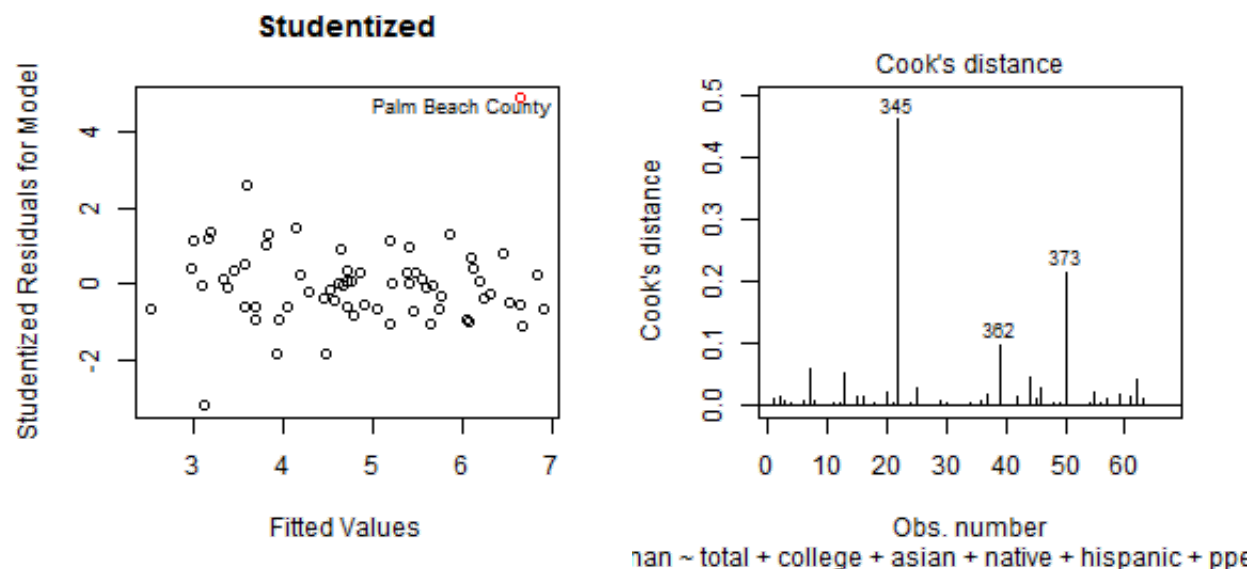
After observing the studentized residuals of our original model, it appears we have a very strong outlier in observation 373, which happens to be Palm Beach County. Upon calculating Cook's distance, Palm Beach and observation 345 (Glade's County) have very high influence on the model. Thus, is further evidence that the Butterfly Ballot in Palm Beach County potentially misled voters. I will remove the outliers from the model, and then attempt to calculate the expected number of votes from Palm Beach County.

Figure 1.6

```
##                        Adjusted R^2    AIC
## Model W/Outliers               0.9064 -127.8
## Model W/Outliers Removed       0.9352 -160.1
```

As can be seen from **Figure 1.6** above, after removing the outliers, the adjusted R^2 for this model increases from 0.9049 to 0.9352; all the coefficients for each term slightly change. The AIC value is also slightly lower, moving from -127.5683 to -160.1156.

By looking at the updated coefficients of this model, we predict the estimated number of votes for Buchanan from Palm Beach County, which comes out to approximately 660 votes, a far cry from the 3,014 votes received. Thus, it is becoming clear that Palm Beach County votes for Buchanan was an anomaly in the state of Florida, and perhaps a result of the Butterfly Ballot. However, in order to pinpoint that the Butterfly Ballot was the issue, and not just an abnormally large population of Reform voters in Palm Beach County, we will next analyze votes from precincts in Palm Beach County, and compare the Butterfly Ballot results to the absentee ballot results.

Figure 1.7

```
##               mod2$coeff
## (Intercept)    -3.5309
## total           0.8848
## hispanic       -0.2573
## native          0.3950
## pperot96        6.7140
## pperot92       -3.8789
```

This is our original graph and trend line from the original data (**Figure 1.4**). The green dot represents our predicted value for Palm Beach County. It is much closer to the original trend line.

Votes for Buchanan in Florida by County
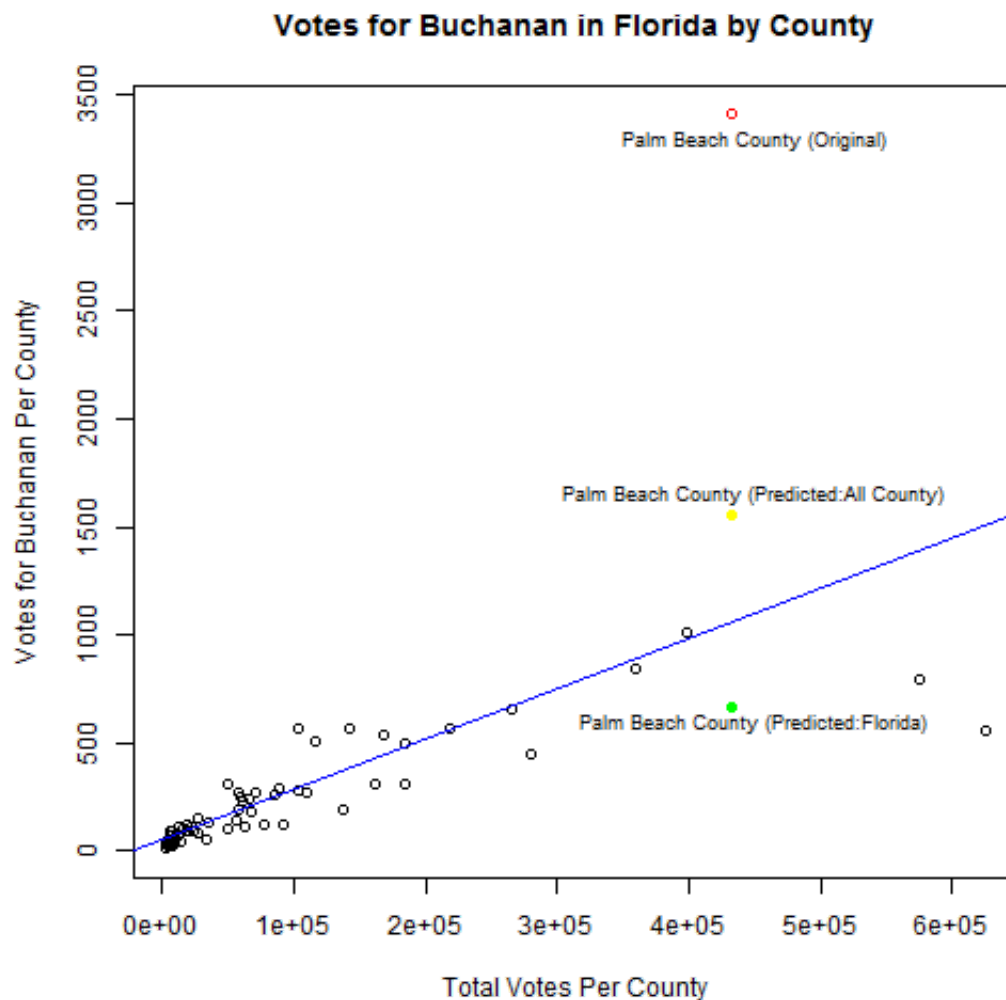
Palm Beach Precincts:

One of our initial concerns is whether or not the Butterfly Ballot on Election Day within Palm Beach County affected election results. This Butterfly Ballot was not offered to absentee voters, and thus we can assume they had the normal ballots. An initial look at the mosaic plot below (**Figure 1.8**), reveals an interesting information on voting tendencies within Palm Beach County.

Figure 1.8

Voting Tendencies In Palm Beach County
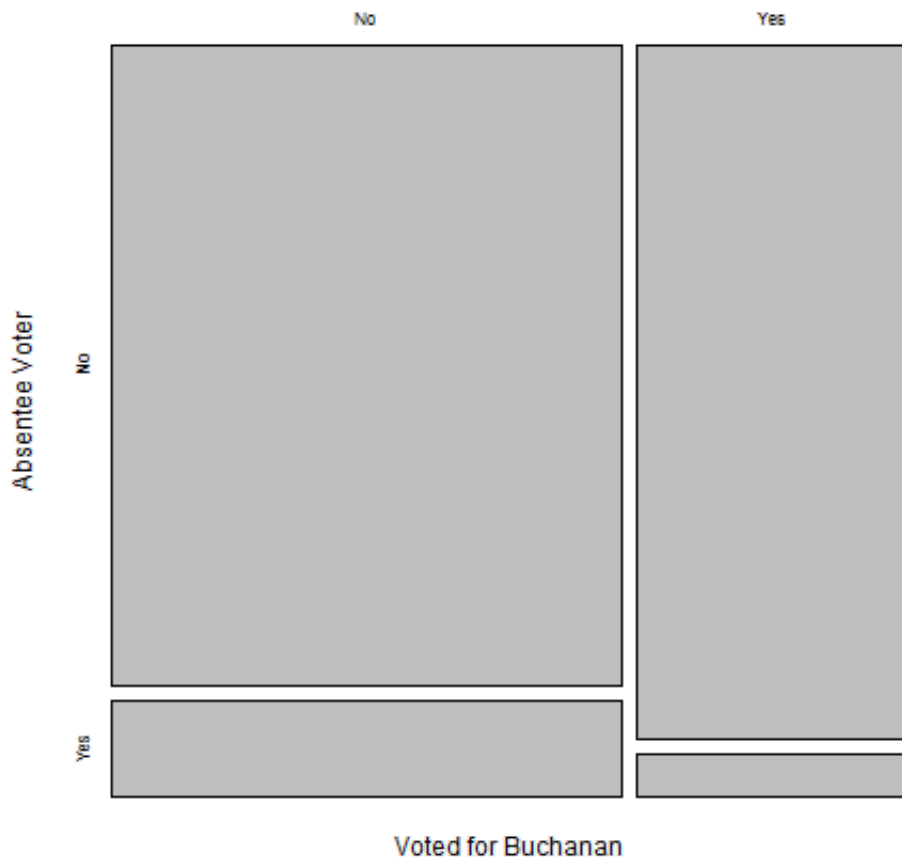
Even though absentee voters were a small portion of all total voters, an overwhelming majority of absentee voters did not vote for Buchanan. However, voters on Election Day who were subjected to the use of the Butterfly Ballot appear to have favored Buchanan. Was the Butterfly Ballot responsible for this?

Analysis II:

Before I begin discussing my model analysis, I would like to state an assumption that I took into consideration. I assume that individuals are voting within their favored party lines. For example, if an individual voted for the reform candidate, Deckard, I would also assume that he or she would vote for Buchanan. Based on this assumption, it would not make sense if an individual voted for Nelson (Democrat), and then voted Buchanan for President. The only way cross-party voting could happen is if voters at the polls in Palm Beach County misinterpreted the Butterfly Ballot.

In order to verify whether the Butterfly Ballot affected voting, I created 4 models split up into pairs of absentee voters and butterfly ballot voters. These models can be found in **Table 1.1** in the Appendix.

In these models, I tested the significance of regressing the number of votes on Nelson, the democratic candidate. Based on my assumption made above, I am assuming no cross-party voting: voters vote in line for their favored parties in all elections. Thus, if it appears that the interaction between Nelson and Buchanan (Democratic and Reform Candidates, respectively)is significant, it suggests that people potentially made mistakes when using the butterfly ballot, and accidentally voted for two parties. In each pair of models, I created an analysis of deviance table to determine whether this interaction was significant in determining the number of votes.

The second model in each figure below dropped the interaction between Nelson and Buchanan.

Figure 1.9 Absentee Voters

```
##                               Res.Df        RSS Df Sum of Sq
F Pr(>F)
## W/Inter of Nelson and Buch    247 25069138 NA          NA
NA     NA
## Interaction Removed           248 25102062 -1    -32924
0.3244 0.5695
```

Figure 1.10 Non-Absentee Voters

```
##                               Res.Df        RSS Df Sum of Sq
F    Pr(>F)
## W/Inter of Nelson and Buch    2142  95544262 NA          NA
NA        NA
## Interaction Removed           2143 104297176 -1  -8752914
196.2 1.004e-42
```

As we can see from the anova table above (**Figure 1.9**), for absentee ballots, the interaction term between Nelson and Buchanan is not significant. Thus, it appears that voters stayed within their own party lines (individuals voted for the same party through all elections).

However, in the last two models (**Figure 1.10**), the interaction between Nelson and Buchanan is significant!

Thus, it appears that cross-voting did occur, which would violate the assumption mentioned above. Thus, this gives an indication that the butterfly ballots potentially caused people to incorrectly vote for the wrong candidates.

Conclusion: After careful analysis, it appears Al Gore lost the election because voters in Palm Beach County were misled when filling out their Butterfly Ballots. The number of votes for Buchanan from this county was very large compared to other counties in Florida. In addition, both estimates (from Florida data and all county data) produced numbers much less than the

3,014 votes Buchanan received. In order to verify that the Butterfly ballots were the issue, and to rule out the idea that there was an abnormally large number of Reform party candidates in Palm Beach, I analyzed the interaction of voting for Nelson (Democrat) and Buchanan, and this interaction was significant. This violates the assumption that voters stay within party lines, and thus, it becomes clear that the Butterfly Ballot misled voters to incorrectly vote for Buchanan rather than Gore.

APPENDIX

Table 1.1

```
     Response Predictors         Interaction Terms
Absentee?
mod1 "Freq"    "Buchanan,Nelson" "Buchanan, Nelson" "Yes"
mod2 "Freq"    "Buchanan,Nelson" "None"             "Yes"
mod3 "Freq"    "Buchanan,Nelson" "Buchanan, Nelson" "No"
mod4 "Freq"    "Buchanan,Nelson" "None"             "No"
```