

Homework 3

Kaleb Luna | kl36927

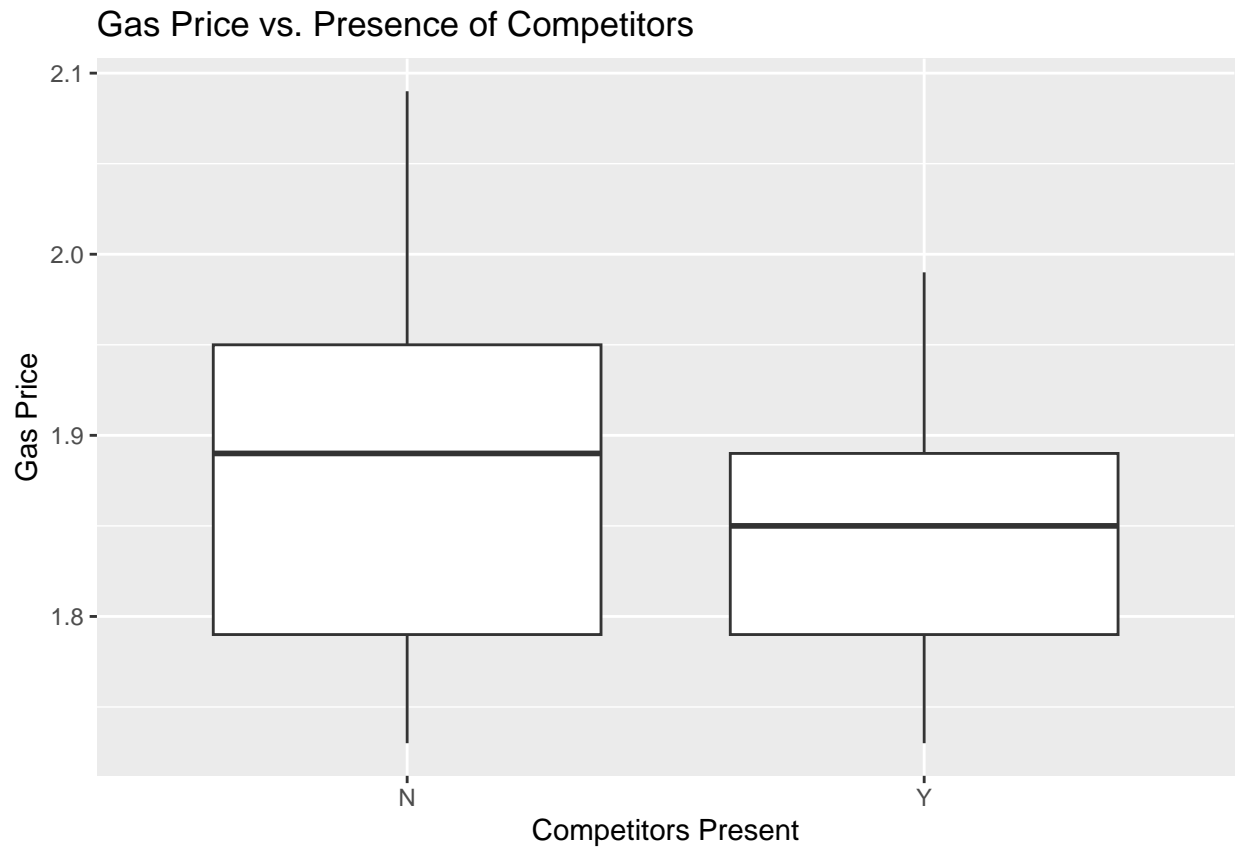
2025-02-12

GitHub Link: <https://github.com/kalebjoel/Homework3.git>

Problem 1: Gas Stations

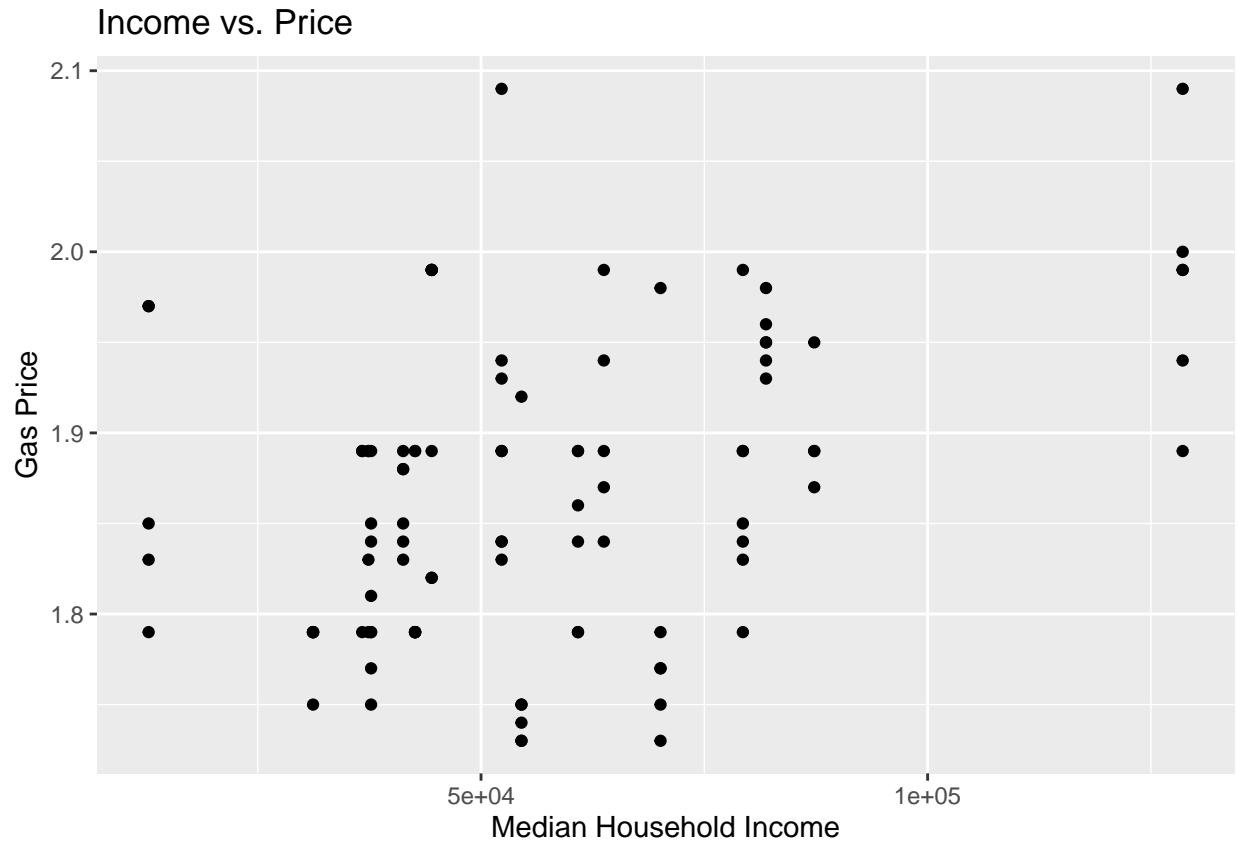
Assessments of Proposed Theories

Theory A: “Gas stations charge more if they lack direct competition in sight.”



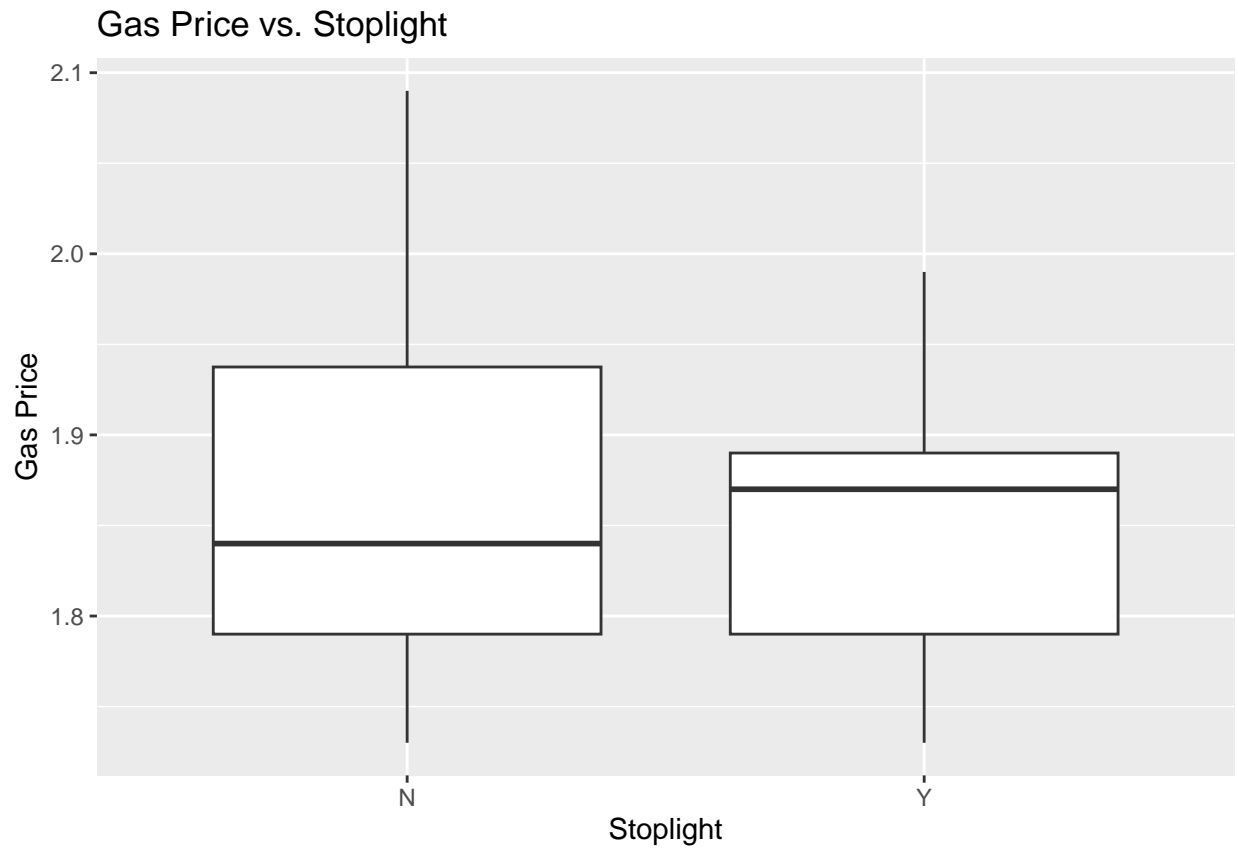
Based on the data above, the theory that **gas stations charge more if they lack direct competition in sight** is supported as **the median gas price for gas stations without the presence of competitors is higher than the median gas price for gas stations in the presence of competitors**. This conveys that gas stations with no present competition nearby typically raise their prices as drivers are faced with little to no alternatives.

Theory B: “The richer the area, the higher the gas prices.”



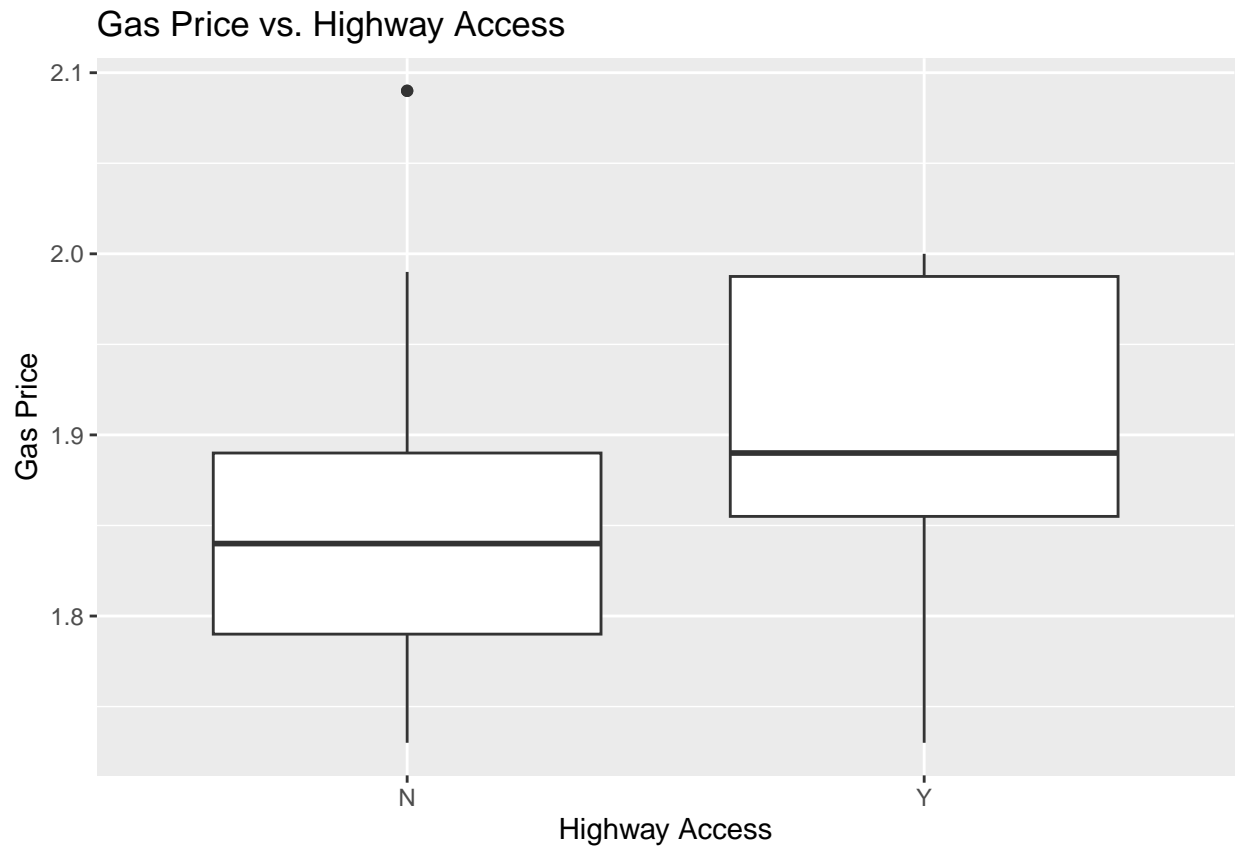
The scatter plot displaying the relationship between **Median Household Income** and **Gas Price** shows a weak, positive relationship, indicating that although it shows a correlation between the two variables, it isn't plausible to conclude that increasing **Gas Prices** are consistent with increasing **Median Household Income**, with a correlation coefficient of approximately 0.40.

Theory C: “Gas stations at stoplights charge more.”



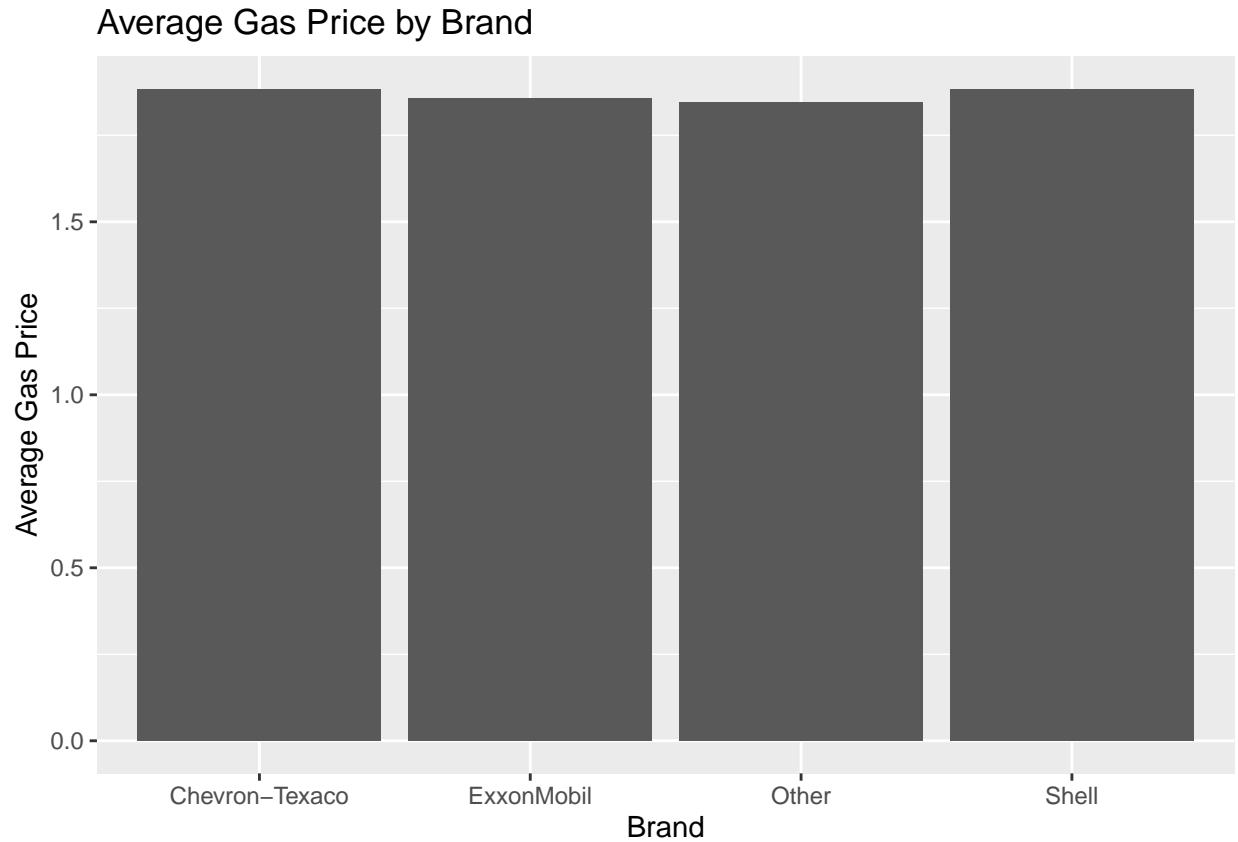
According to the figure above, the box plot that shows **gas stations at stoplights** has a **higher median price** than **gas stations that are not located at stoplights**, indicating that the theory that gas stations at stoplights charge more is supported. Gas stations located at stoplights may use their location to their advantage as drivers tend to see their stop as a decision out of convenience.

Theory D: “Gas stations with direct highway access charge more.”



According to the figure above, the box plot that shows **gas stations with highway access** contains a **higher median gas price** than **gas stations without direct highway access**, indicating that **the theory that gas stations with direct highway access charge more is supported**. Similar to previous mentioned theories, gas stations often capitalize on the convenience factor that most drivers prefer, as drivers are typically faced with limited options at this point.

Theory E: “Shell charges more than all other non-Shell brands.”



According to the bar plot comparing the average gas prices by brand, it can be seen that **there is no significant difference between the four major gas fuel brands**, including Chevron-Texaco, Exxon Mobil, Shell, and other, indicating that **the theory that Shell charges more than all other non-shell brands is not supported**. This suggests that gas prices are more likely to be influenced by other factors including location, competitors, and convenience as found in previous figures.

Problem 2: Mercedes S-Class Vehicles

Part A

```
## [1] "The 95% Confidence Interval for the proportion of the average mileage of 2011 S-Class 63 AMG's: 2  
## [2] "The 95% Confidence Interval for the proportion of the average mileage of 2011 S-Class 63 AMG's: 3
```

Part B

```
## [1] "95% CI for proportion of Black 2014 550s: 0.42"  
## [2] "95% CI for proportion of Black 2014 550s: 0.45"
```

Problem 3: TV Shows

Part A: *“Living with Ed”* vs. *“My Name is Earl”*

Which show makes people happier? “Living with Ed” or “My Name is Earl”? To investigate this question, I filtered the dataset to distinguish the responses from only the two shows of interests and obtain the mean Q1_Happy rating from both TV shows. Then, I calculated the 95% confidence intervals of the difference of means in order to test the significance of the data, which provides a measure of uncertainty when estimating the ratings between the two shows. The difference of means, standard errors, and the confidence interval are approximately -0.150, -0.00525, and -0.14 to -0.16, respectively. The confidence interval does not contain a value of 0; therefore, it can be concluded that one show produces higher happiness ratings on a consistent basis.

Part B: *“The Biggest Loser”* vs. *“The Apprentice: Los Angeles”*

To explore the question of which reality show made people feel more annoyed when watching, I filtered the dataset to only contain responses from the two shows of interest: “The Biggest Loser” and “The Apprentice: Los Angeles.” Then, I calculated the 95% confidence interval for the difference in means. The difference of means, difference of standard errors, and the confidence interval are approximately -0.271, -0.000436, and -0.27, respectively. The results as a whole do provide convincing evidence that one show is rated “more annoying” than the other as the confidence interval does not contain 0.

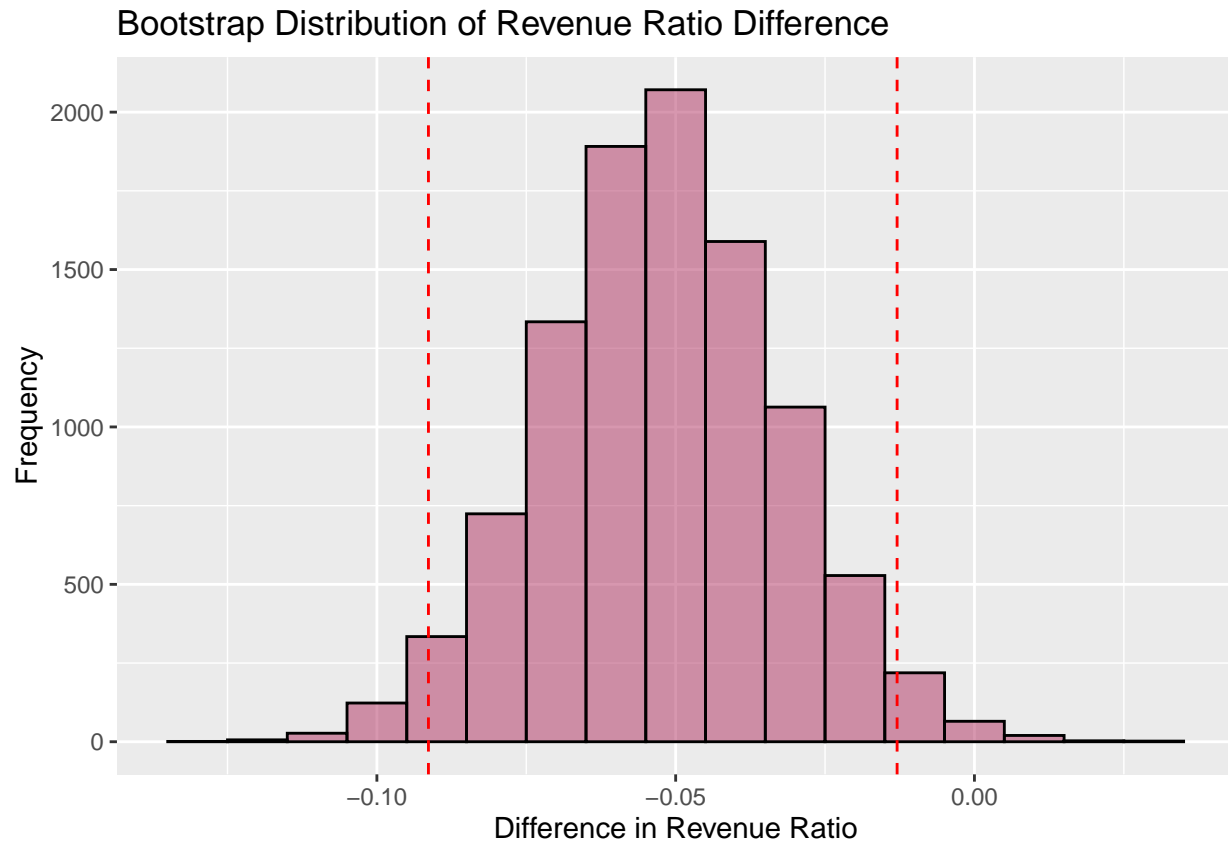
Part C: *“Dancing with the Stars”*

To investigate the question of what proportion of American viewers find the show “Dancing with the Stars” confusing, I filtered the data set to include only the responses for the show of interest and calculated the proportion of American viewers who gave it a confusing rating and a 95% confidence interval. With a calculated proportion of approximately 0.0773 and a confidence interval of 0.04 to 0.12, this provides an estimate of how viewers often find the show confusing when watching. Because the confidence interval is relatively high, it can be concluded that a majority of proportion of viewers agree that they find “Dancing with the Stars” confusing.

Problem 4: EBay

Observed difference in revenue ratio: -0.05228145

95% Confidence Interval: -0.09137524 to -0.0129432



When exploring the question of whether or not Google paid search advertising affects eBay's revenue ratio at the DMA level, I calculated the revenue ratio by dividing the revenue after by the revenue before for each DMA and compared the mean revenue ratios between the treatment (paused ads) and control groups (unpaused ads). In this bootstrap method, 10,000 Monte Carlo simulations were utilized in order to calculate the 95% confidence intervals for the difference of means. The difference in revenue ratios was -0.0523, with a confidence interval of -0.0914 to -0.0129, which suggests that pausing Google paid search advertising is statistically significant in the decrease of the revenue ratio. Overall, this result suggests that Google paid search advertising does in fact increase eBay's revenue.