5-2018

The NFL and Trump: Did Protesting Cause the Decline?

Jameson Osmond
Utah State University

The NFL and Trump: did protesting cause the decline?

by

Jameson Osmond

**Capstone submitted in partial fulfillment of
the requirements for graduation with**

UNIVERSITY HONORS

With majors in

**Finance
in the Department of Finance and Economics;**

**Economics
in the Department of Finance and Economics;**

**International Business
in the Department of Management**

Approved:

Capstone Mentor
Dr. Briggs Depew

Departmental Honors Advisor
Paul Fjeldsted, CFA

University Honors Program Director
Dr. Kristine Miller

**UTAH STATE UNIVERSITY
Logan, UT**

Spring 2018

Acknowledgements

I want to thank Dr. Briggs Depew. He was an invaluable resource, whose expertise and guidance made this project possible. Thank you so much, Professor!

Word Count: 5,939

Contents

I. Abstract

This project is a study in econometric modeling of the effects of the protests of the national anthem in the NFL during the 2016-2017 seasons. The project is created to determine the accuracy or lack thereof of President Donald Trump's statement that the cause of the decline in viewership and ratings (and thus business) of NFL games was caused by protests that deterred US viewers.

Using viewership and rating data, along with various protest indexes created by collecting game-level protest data, an econometric model was constructed to allow for control over various endogenous and exogenous variables that surround NFL in-season data. These variables include: home vs. away team; time of day; day of the week; which week the game occurred in the regular season; win/loss record of the opposing teams at the time of each game; which network or networks televised the game; percentage of the national audience with access to the televised game; and, importantly, the method of protesting of the NFL players. Fixed effects were also incorporated to control as much as possible for holidays – i.e. the U.S. holiday of Thanksgiving, and the change of seasons (from Week 17 in year one to Week 1 in the next, and so on).

The data was run through various regressions to test for the individual effect of each form of protest and their aggregates against a one-week lagged effect on viewership, in both the binary and continuous forms.

According to the findings of the model, dummy forms of the protest variables did not have a significant enough effect on viewership to

reject the null hypothesis, but four of the direct forms of protesting were significant at the 1% level, allowing the rejection of the null and suggesting that the number of NFL protesters at each game did indeed have a negative effect on NFL viewership in the United States.

II. Introduction

As the newly elected President, Donald Trump has become known for his direct style of confrontation on economic and social issues, especially through his tweets. The Twitter platform's restricted word count on posts compound this directness, as his comments are forced to be concise in order to express his point quickly. On September 27, 2017, this directness was exemplified when Trump made the following comment in his twitter feed: "NFL attendance and ratings are WAY DOWN. Boring games yes, but many stay away because they love our country. League should back U.S." [1]

This tweet came in response of a (then) year-long string of protests of the national anthem by NFL football players, beginning in the 2016 preseason and continuing into the 2017 season. The movement began when one player, Colin Kaepernick, knelt during the anthem in protest to "what he felt were wrongdoings against African Americans and minorities in the United States."[2] During the 2016 season, many other players began to imitate Kaepernick's protests in various ways during the anthem, including by kneeling, raising fists, linking arms, holding shoulders, staying in tunnels or lockers room, and others.[3] In 2017, the

[1] Davis, J. (2017). *Trump Calls for Boycott if N.F.L. Doesn't Crack Down on Anthem Protests. Nytimes.com*. Retrieved from https://www.nytimes.com/2017/09/24/us/politics/trump-calls-for-boycott-if-nfl-doesnt-crack-down-on-anthem-protests.html

[2] Colin Kaepernick explains why he sat during national anthem. (2018). NFL.com. Retrieved from http://www.nfl.com/news/story/0ap3000000691077/article/colin-kaepernick-explains-why-he-sat-during-national-anthem

protests continued, and drew as much, if not more, media attention, resulting in this eventual input of the President.

This project is created in response to the President's comments. Nominally, viewership data of NFL games over the last few years has dropped significantly over the timeframe of the protests, but there are many variables that must be controlled for to determine whether or not the protests are a statistically significant contributor of the decline. In order to account for these variables, this analysis aggregated data of protests and viewership and analyzed it in an econometric regression model, controlling for: the home vs. away team, time of day, day of the week, which week the game occurred in the regular season, win/loss record of the opposing teams at the time of each game, which network or networks televised the game, and percentage of the national audience with access to the televised game. Fixed effects were also incorporated into the data to account for various time-specific events.

III. Literature Review

Determining the economic effects of social actions occurring in sporting events is a forward-thinking concept whose validity can rightfully be questioned. However, although limited in number, this specific type of analysis can be backed by professional and well-known precedent.

A. Football and Family Violence

Publicized in 2009, a study conducted by David Card and Gordon Dahl tested for the effects of upsets in professional football games on reports of domestic violence.[4] The pair used a Poisson model to test for

---

[3] *GamesCricketRugbyEnduranceCFL, &., Sports.* (2018). *NFL players who protested during national anthem in Week 2.* ESPN.com. Retrieved from http://www.espn.com/blog/nflnation/post/_/id/214580/nfl-players-who-protested-during-national-anthem-in-week-2

the probability of the occurrence of a report of domestic abuse after the result of a regular season football game, controlling for variables such as home team vs. away team, day of the week, and expected winner/loser. This last variable fulfills a similar role as the win/loss rate in my analysis, except that their expectations are based off of Las Vegas betting markets and not past records. They also incorporate fixed effects such as non-game day(s).

Their analysis finds a positive correlation between upset losses in football games and cases of male-to-female domestic violence, with no corresponding female-to-male increase. Upset wins resulted in a dampened effect on cases of domestic violence. The results were specific enough to allow the pair to find significant correlations between reports of domestic violence and games with an unusually high number of sacks and/or interceptions. These results are a positive sign for this research, and help verify the validity of the findings.

B. Super Bowl and Birthweight

In a study of children born from 1969 to 2004, a positive correlation was found of incidences of low birthweight of newborn children and Super Bowl wins in the counties where children were born.[5] The study used publicly available data of the birth weight of children from mothers that lived "no further than one county away from an NFL stadium." The team not only found that winning was associated with more incidences of lower birth weight (which they interpret to be correlated with stress events), but that upset win and losses generated more substantial differences in the birth weights.

---

[4] *FAMILY VIOLENCE AND FOOTBALL: THE EFFECT OF UNEXPECTED EMOTIONAL CUES ON VIOLENT BEHAVIOR* (2018). *Nber.org.* Retrieved from http://www.nber.org/papers/w15497.pdf

[5] *Emotional Cues and Low Birth Weight: Evidence from the Super Bowl* (2018). Conference.iza.org. Retrieved from http://conference.iza.org/conference_files/riskonomics2012/rees_d4915.pdf

The study follows a similar structure to this analysis, controlling for various other factors that could affect birth weight, yet goes further with the data set, segmenting the mothers as well as the children into brackets in order to test for other factors such as levels of substance abuse among mothers near the Super Bowl against birthweight and other outcomes.

IV. Data sources

A. Data on viewership and ratings

This data set was readily available from *statista.com,* which was an aggregate and "cleaned" data set on viewership and ratings records of NFL games compiled from: Sports TV Ratings, Showbuzz Daily, Programming Insider, ESPN.com, NBCUniversal, and CBS.[6] This analysis originally relied on the veracity of the data, but each data point was later verified.[7] The restrictions of this data set are discussed further in the "Data analysis challenge" section.

B. Data on protests

Data collection on protests was not readily available by season. Various websites, such as espn.com and sportsillustrated.com contained weekly summaries of protests, but these were in paragraph form and required manually entry into excel.[8] The restrictions of this data set are discussed

[6] *National Football League (NFL) television ratings and viewership figures in the United States in the 2017/2018 season, g.* (2018). *NFL viewership figures 2017-18 | Statistic. Statista.* Retrieved from https://www.statista.com/statistics/749140/nfl-viewership-figures-ratings/

[7] *NFL Ratings 2017 Compared to Last Year - Page 3 of 4 - Sports Media Watch.* (2018). *Sports Media Watch.* Retrieved from http://www.sportsmediawatch.com/nfl-tv-ratings-viewership-nbc-cbs-fox-espn-nfln-regular-season-playoffs/3/

[8] This source contains links to reports of every other week's protests in the 2017 season: *NFL players who protested during national anthem in Week 17.* (2018). *ESPN.com.* Retrieved from http://www.espn.com/blog/nflnation/post/_/id/264431/nfl-players-who-protested-during-national-anthem-in-week-17

further in the "Data analysis challenge" section.

C. Data on win/loss records

Data on win/loss record was not readily available on a game-by-game basis. Nfl.com contained a usable source of win/loss records by week, post-game, which was used to extrapolate the necessary data. The restrictions of this data set are discussed further in the "Data analysis challenge" section.[9]

V. A simple model of protests on viewership

A. Variable setup

The basic layout of the model of the effects of protesting on viewership attempts to isolate the protest data from other significant variables that relate directly with the surrounding season's games.

The chosen data set contains information on the 2015-2017 seasons. This allows the model to have a control season as a baseline, which then incorporates the protest data from the following 2016-2017 seasons. The high-level raw data on each game in each season is entered into five different variables: the first two are the home and away variables. Home teams draw more viewers than away teams at each stadium, and a team with a particularly active fan base or a particularly inactive fan base at home could skew the data. This

---

The 2016 season data was sourced from weekly data gathered by ESPN, such as: *NFL players who protested during the national anthem in Week 17.* (2017). *ESPN.com.* Retrieved April 19, from http://www.espn.com/blog/nflnation/post/_/id/227095/nfl-players-who-protested-during-the-national-anthem-in-week-17

[9] *NFL.com Scores*. (2015). *Nfl.com*. Retrieved from http://www.nfl.com/scores/2015/REG1
NFL.com Scores. (2016). *Nfl.com*. Retrieved from http://www.nfl.com/scores/2016/REG1
NFL.com Scores. (2017). *Nfl.com*. Retrieved from http://www.nfl.com/scores/2017/REG1

"geographic popularity" variable controls for variability expressed in the viewership data of the game.

The second through fifth variables are the times of day, week, and season the games occurred, and are included because (in general) games held before 4:30 pm draw fewer viewers than games held after that time, and need to be addressed in the model. Often, the games held before 4:30 are held at 1:00 pm, and multiple networks will televise between 5-10 games at once, further reducing the numbers of viewers of each game, which skews the data if left unaddressed. In a similar manner, games held on Sunday will draw more viewers than games held on Thursday, and so on.

Viewership numbers have a tendency to decline from Week 1 – Week 17, for reasons that can be surmised as early fervor that tapers off as winners and losers are established (especially between playoff and non-playoff teams), the depreciation of novelty and excitement, etc. If this variable were not accounted for, any analysis of the data would be entirely useless, as the unobserved term in the model would hugely bias the protest data index.

Another variable incorporated into the data set is the win/loss records of each team, and or this model, the data is reported as the win/loss record of each team *before* each game. Data reported in this form is most effective in accounting for potential variance in viewership numbers if a team is passing through an unusually high- or unusually low-performing season.

Finally, multiple protest variables are included in the set. These variables look at the protest data both in aggregate (all forms of protesting) and individually, to test for potential different effects caused by the different forms of protest. The aggregate data is split into six different variables: kneeling, raised fist, shoulder hold, sitting in tunnel or locker room, linking arms, and other (random protesting acts that were not continued beyond one or two games), along with the clustered (aggregate) variable.

From these variables, many dummy variables are added to the model, which are described in the following section.

B. Model setup

The econometric model used is a simple log-level analysis in a panel model. This allows for changes in the variable protest to have a constant percent association with the outcome. Doing so handles the issue of widely varied nominal increases and decreases in viewership numbers among teams with different levels of average viewership by standardizing the data.

Various questions had to be answered to set up an accurate time series model, namely: which outcomes to test for (including the addition of the necessary dummy variables to improve the testing), how to organize the data (meaning how to merge the seasonal data sets/add fixed effects), and how to incorporate a lag effect.[10]

Outcomes: The testable outcomes follow a basic y_t=Xß relationship, meaning that the relationship of the predictor variables over time (X) and the observed result of the relationship (Y) are used to discover the relationship coefficient (β).In this model, these tests could be: whether the aggregate protests by game have an effect on overall viewership,

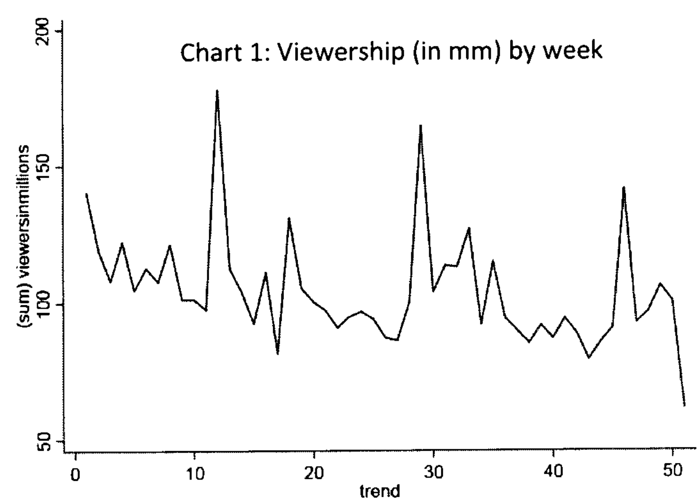
[10] In the remainder of this section, the data is to referred to in "weekly aggregates." The distinction is explained in the *Data Analysis Challenge* section.

whether each team's protests have an effect on the viewership of their individual games, whether each type of protest has a larger or smaller effect on viewership than another, or whether different numbers of protestors in each game cause different levels of viewership change.

These tests can also be binary, meaning if *any* protesting, regardless of the number of protestors, has as effect on viewership numbers. This test was included in the model, and doing so required that binary variables be added to the data sets. In this case, it was a simple addition: another column was added to the data set, and any game with any form of protest (where protestors of any form >0) was designated one, and any game with no form of protest was designated zero.

<u>*Organizing the data:*</u> Data organization faced two challenges: time gaps in seasonal data sets and one-time events.

A panel model walks through all the data in the data set as if each point were equidistant from the other. However, there is a large time gap between each season of NFL games from each week 17 to each week 1, and as is mentioned, viewership data in week 1 tends to be higher than in week 17. This can be seen in the Chart 1, where week 18 (week 1 in year 2) sees a large spike in the data, which happens again in week 35 (week 1 in year 3). In order to account for this time disparity and skew in the data, a "fixed effect" variable is added to each week 1. A fixed effect is an assumption that there is a unique event that is correlated



with an independent variable, and accounting for it as a variable in the model removes its impact on the analysis.

Similarly, another fixed effect in the model is Thanksgiving. Viewership numbers during the week of Thanksgiving jump disproportionately high compared the rest of the model, which can be seen in each of the largest three spikes in the model. A variable is added to account for that disproportionate increase and remove its impact from the model.

Lag effect: One of the key points of analysis in the model is the use of a lag effect, which allows the model to test for the effect protests have on games *in the week following,* rather than the same week the games occur, which would not be as realistic. Most viewers, if they were to choose to disassociate themselves with the NFL games, would do so after seeing the protests in the current game, because it is likely that they would have already decided to commit the time to view to the current one.

In order to incorporate a lag effect, the data set is modified to effectively shift every week's protest data to the next week, leaving all the other data alone. [11] This simulates a lag effect – to the data, it appears that the data in week 2 is being affected by the protest in week 2, but that protest data actually comes from week 1, keeping the model simple and clean. In order to make the shift work, week 17, which becomes week 18 with the shift, is dropped from the data set (the same with week 34 and 51).

Each type of protest variable (kneel, fist, shoulder, etc) is replaced with a new one (kneel_lag, fist_lag, shoulder_lag) in the regression.

VI. Data analysis challenge

Throughout the discussion of the setup of the model, the analysis method refers to "aggregated weekly data" rather than "game-by-game" data. This aggregation came about by necessity during the course of research, and was challenging enough to change the direction of the outcome testing/ analysis.

The issue occurred between the sourced viewership data set and compiled "protest index" data set. Initially, the sourced viewership data appeared to have both viewership and rating records for each individual game; however, when the protest variables were added to each game's record, only about 1 in every 3 games with protestors was connected with corresponding game in the data set. A quick verification determined that the viewership data set, while still reporting the entire viewership of all NFL games over the 2015-2017 seasons, only individually listed <u>nationally</u> televised games, and grouped together the viewership data of most of the <u>regional</u> games.

A regional game is only broadcast to a certain portion of the country, a method used by NFL committees and broadcast agencies that groups together counties to assure that certain games' viewership numbers are at their maximum, tailoring the broadcasts to certain audiences.[11] This grouping process is called regionalization, and during that time NFL rules prohibit other NFL games from being shown on local television

[11] For example, in in the 2017 season, a portion of Mississippi was broadcast a game between the Chiefs and the Giants at Kansas City (two regions unrelated to Mississippi) because Eli Manning, the Giants Quarterback, "starred there...and CBS believe[d] that a lot of people in the state still want[ed] to watch him play." This occurred even though a (notably) higher-rated game (Baltimore at Green Bay) would be happening at the same time. Draper, K. (2017). *Why People in Mississippi Have to Watch the Giants. Nytimes.com.* Retrieved from https://www.nytimes.com/2017/11/19/sports/football/why-people-in-mississippi-have-to-watch-the-giants.html?smid=tw-nytsports&smtyp=cur

stations while the "local" team is playing a sold out, "locally" televised home game.

This issue would be easily fixed by adding a variable to the model that accounted for the percentage of viewers that have access to watch each regional game; however, the available data only listed the viewership data for the regional game *that garnered the highest viewership and rating data* – all other regional games during that time (each Sunday) were grouped together as "Various (week x)" in the set.

This complicated the analysis. If the other 2/3 of the protest data were not included, any conclusions that could be drawn from the analysis would underrepresent the number of games with protesters. But there also did not exist any plausible way to incorporate the protest data into the grouped games, barring any complicated "popularity" rating that could arguably divvy up viewership numbers to each game. However, this would be highly uncertain – broadcasters have control to arbitrarily assign games to certain regions to maximize total overall viewership of whichever game they choose, and an attempt to estimate the per game viewership would be inconclusive.

However, even considering these incongruities in the data sets, the data sets were not incompatible. In order to overcome the challenge in combining the sets, all data from each relevant variable was grouped together by week. This allowed all of the protest data to be inserted into the data set without increasing volatility from the "Various" data points. A separate index of protest data listed by week was created to replace individual protest data in the analysis, and viewership data was aggregated.

Doing so caused the following ramifications:

No team-by-team analysis: As previously mentioned, one of the potential outcomes that could be inspected by this model is the effect of player protests in each NFL team on that team's future viewership data, controlling for other variables like the opposing team, home vs away, etc. Doing so would allow for interesting analysis about the preferences of consumers in the geographic areas (or some other such segmentation) surrounding each team, i.e. how sensitive the viewers in each segment are to protests. With this type of analysis, each team's individual risk of loss could be directed analyzed and even quantified, and would be the ideal outcome variable to test. Aggregating the data by week eliminates any opportunity to study this outcome variable. Instead, outcome variables can only focus on whether or not the forms of protest had an overall effect on viewership of games in the league.

No win-loss rate incorporation: Aggregating the data by week removes the ability to include variables on the win/loss rate of each team before each game. The win/loss record of the league as a whole is always static, and neutralizes the impact of this variable.[12] However, some bias may push the numbers up or down if a widely viewed team performs well above or below average.

No "percentage of viewers with access" analysis: Because all data points are aggregated by week, an analysis of a "percentage of viewers with access" variable is not performed.[13]

---

[12] A potential solution to this issue would be to incorporate the Las Vegas betting outcomes on NFL games, and assign some variable that could relate the estimated outcomes of the games with increased or decreased levels of viewership. Doing so, however, would require its own analysis and did not fit into the timeframe of this project.

[13] The highest rated regional games are individually listed, and the variance in the viewership in these games could be accounted for individually, but is not performed.

Dummy variables are unaffected, both in their binary lagged and continuous lagged forms, as they are simply aggregated along with the rest of the data.

Accounting for the day in which the games occurred, and their time, is still possible. Even though the data is aggregated by week, most (if not all) games grouped into the "various" category occurred on Sunday at 1pm, allowing the model to control for each data point's day and time of occurrence.

Home team and away team variables are accounted for in the same way – it is a categorical variable that is accounted for separately. Control for the "network" variable is also performed.

Final model equation: The final regression equations run into the model are built off the following format:

$$Ln(v) = X\beta + \alpha p + e$$

X is the vector of controls; in the controls we allow protest to vary according to the indicator variable, p, as a continuous/level number of protestors.

The equation test for the percent change on viewership, controlling for home vs away team, the day the game occurred, the broadcast agency used, the Sunday night fixed effect, the time of the game (afternoon or evening), and the Thanksgiving fixed effect. Each regression equation tests against one of the protest variables - either a cluster or an individual form of protest - whether binary or continuous. There are multiple equations because in this model the protest variables can only be addressed one at a time so as not to violate Gauss-Markov Assumption 4: the Zero Conditional Mean.[14]

VII. Findings

The final results of the regression analyses are grouped into the following four categories: Binary – all games; Continuous – all games; Binary – only national games; Continuous – only national games. Given the above-mentioned fact that no analysis is performed on the variance of individually listed regional games, it is appropriate to present the data both including and excluding those games. See Table 1 for the breakdown of the findings:

Binary – All Games

The categorical results of the Binary – All Games regressions are very similar. The R^2 and Root MSE in each test are within 1% of the average. R^2 represents "how much variance in the data can be explained by the model," and Root MSE is the mean squared error – the difference between the estimated outcomes from the model and the actual outcomes.

What these numbers test for is the probability that each form of protest, regardless of the number of protesters, is a significant regressor affecting viewership in its standardized form (remembering that each regression result comes from an isolated test – none of these variables were run simultaneously). As a note, in most econometric models, a regressor is considered significant if the "P" value is less than either 10%, 5%, or 1%, with one percent being the most significant,

[14] Angrist, J. D. and Pishke, J. (2014). *Mastering 'Metrics: The Path from Cause to Effect.* Princeton University Press. See also, Lee, Q. (2017). *OLS, BLUE and the Gauss Markov Theorem | UW Economics Society. Uweconsoc.com.* Retrieved from http://uweconsoc.com/ols-blue-and-the-gauss-markov-theorem/

stating that “there is a 99% chance that regressor x is a significant variable.” In these tests, only the “other” protests are considered significant at the 10%, 5%, and 1% levels, suggesting that the simple act of a protest of an NFL player during the national anthem does not cause a decrease in NFL viewership.

Continuous – All Games

The categorical results of the Continuous – All Games regressions are also very similar. The R^2 and Root MSE in each test are within 1% of the average.

The results of the Continuous chart are much more significant: every single regressor is significant at the 10% level, all but the “other” protest category are significant at the 5% level, and all but the “fist”, “shoulder”, and “other” categories are significant at the 1% level. This suggests that the actual number of NFL players protesting had a significant negative correlation on NFL viewership. For example, the regression on the “kneeling” regressor states in sentence form: “It can be said with 99.7% confidence that for every additional player that kneels during the national anthem during a given week, the percentage of viewership of NFL games will drop .001% the next week.” Or, for the “raising a fist” regressor: “It can be said with 96.6% confidence that for every additional player that kneels during the national anthem during a given week, the percentage of viewership of NFL games will drop .0045% the next week.”

An interesting point to note is that when taken as a whole, the number of NFL players protesting the anthem is significant at the .1% level, although the effect is reduced to a .0001647% drop for each additional protestor, suggesting that there may be some bias in each form of protest that is addressed when all forms of protesting are grouped together.

To put these numbers in perspective, as of 2016, the NFL had 32 teams with a maximum roster of 53 players, setting the maximum number of players in the league to 1696; if every single player protested the anthem in some form, viewership in the next week would be expected to drop .2793%, i.e. by ~391 thousand viewers from Week 1 to Week 2 of 2017(including viewers of more than one game per week).

Binary – Only National Games

As can be seen in Table 1, removing the regional games from the analysis drops the total number of observations to 172 from 259, meaning that, ceteris paribus, the protest variables would need a higher t-stat to have a significant p-value. The regression of the overall R^2 of the analysis increases to a range of .937-.942 in each test, up from .88 in the "All Games" regressions, and visually shows the removal of the aforementioned variance in regional game viewership in the data analysis section. MSE also drops to a range of .156-.162 from .18.

Interestingly, by removing the regional games from the data set, the binary combined regressor becomes a significant at the 10% level, dropping from 14% to 6.2%. It could infer that: (i) viewers are more sensitive to any protests in the national scene than they are with any protests from their own region or (ii) that the extra variance in viewership data from the regional games skewed to the "All Games" analysis, which was removed from the data set in this regression. It may also be due to other unknown factors.

Continuous – Only National Games

As with the Binary – Only National games summary chart, almost 90 observations are removed from the analysis, and the R^2 (.93) and Root MSE (.16) values improve, all values ranging within .002 of the average.

By removing the regional games, the statistical significance of the regressors in these tests decreases, notably so in some cases. What these results may suggest is that: (i) regional games are more heavily influenced by the number of protestors than national ones (an interesting finding given the result of the binary test), (ii) there may have been a skew in the data from regional games, or (iii) teams that consistently had many protesting players were mostly televised regionally or were pushed out of nationally televised games into regionally televised ones, reducing the numbers of incidence of protests in this data set. It may also be due to other unknown factors.

VIII. Conclusion and Further Research

In summary, this econometric analysis shows that NFL protests do have a significant effect on the standardized metric of viewership. In three of the four methods of testing (all except the "Binary –All Games" regressions) most test are at least significant at the 10% level, and the methods of protesting are significant at the 1% level among most "Continuous – All Games" tests.

According to this model, it can be said with at least 90% confidence that President Trump was accurate in his Twitter post statement that the NFL protests have had a negative impact on viewership.

Improving the regression: There are five areas of study that could increase the accuracy of the analysis:

- First, adding an additional base year as a control. Public data on viewership from known sources extends back an additional year. Adding the 2014 season to the data set could add more accuracy to the trends already in force before the inception of the protests.
- Second, collecting individual viewership data of regional games. As discussed, regionalization and grouping of viewership data prevented the analysis of impactful variables such as team-by-team protest risk. Gaining access to that data will create an opportunity for powerful analysis with clearly quantifiable outcomes.
- Third, addressing potential Zero Conditional Mean violations in the data set. One other factor not considered in this analysis is the potential violations of the Zero Conditional Mean assumption in the "broadcaster" regressors. Although the effect may be small, further research into effective methods of accounting for correlations between broadcasting agencies would hone the accuracy of the analysis.
- Fourth, inclusion of the rise or fall in popularity of other sports. There may be an all-encompassing trend in viewership data among all sports that could be affecting the data set.
- Fifth, verifying the accuracy of streaming data inclusion/non-inclusion. Streaming viewership numbers are small in comparison to TV viewership (1 or 2% of the size), but, as with the other factors, may increase the accuracy of the model.[15]

---

[15] *NFL Ratings 2017 Compared to Last Year - Sports Media Watch.* (2018). *Sports Media Watch.* Retrieved from http://www.sportsmediawatch.com/nfl-tv-ratings-viewership-nbc-cbs-fox-espn-nfln-regular-season-playoffs/

Table 1

Tests for significance: estimated effect of forms of protest on NFL Viewership
(Dependent Variables: Form of protest)

	N= 259 Avg. R^2: .8797	N= 259 Avg. R^2: .8807	N= 172 Avg. R^2: .9394	N= 172 Avg. R^2: .9385
	Binary – All	Continuous - All	Binary - Regional	Continuous – Regional
Any Protest	-0.07758	-0.000165***	-0.09932*	-0.000120**
	(0.05170)	*(0.000043)*	*(0.05195)*	*(0.000055)*
Kneeling	-0.07758	-0.001009***	-0.09932*	-0.000676
	(0.05170)	*(0.0003187)*	*(0.05195)*	*(0.000449)*
Sitting	-0.02453	-0.002907***	-0.00747	-0.002639**
	(0.04729)	*(0.001074)*	*(0.05806)*	*(0.001264)*
Tunnel	-0.05122	-0.000973***	-0.02955	-0.000986**
	(0.04764)	*(0.00026)*	*(0.06342)*	*(0.000393)*
Fist	-0.07758	-0.004497**	-0.09932*	-0.004359**
	(0.05170)	*(0.002065)*	*(0.05195)*	*(0.002140)*
Shoulder	-0.07248	-0.002870**	-0.03190	-0.001001
	(0.05099)	*(0.001145)*	*(0.06668)*	*(0.001624)*
Link Arms	-0.05680	-0.000227***	-0.05944	-0.000162*
	(0.03727)	*(0.000060)*	*(0.04054)*	*(0.000087)*
Other	-0.13800***	-0.01693*	-0.14365**	-0.017434
	(0.0473)	*(0.008456)*	*(0.06666)*	*(0.010684)*

Each point estimate and the corresponding standard error (in parenthesis) is estimated from separate OLS regression. *: significant at 0.10, **: significant at 0.05, ***: significant at 0.01 level

IX. References

Angrist, J. D. and Pishke, J. (2014). *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press.

Colin Kaepernick explains why he sat during national anthem. (2018). NFL.com. Retrieved from http://www.nfl.com/news/story/0ap3000000691077/article/colin-kaepernick-explains-why-he-sat-during-national-anthem

Davis, J. (2017). *Trump Calls for Boycott if N.F.L. Doesn't Crack Down on Anthem Protests. Nytimes.com*. Retrieved from https://www.nytimes.com/2017/09/24/us/politics/trump-calls-for-boycott-if-nfl-doesnt-crack-down-on-anthem-protests.html

Emotional Cues and Low Birth Weight: Evidence from the Super Bowl (2018). Conference.iza.org. Retrieved from http://conference.iza.org/conference_files/riskonomics2012/rees_d4915.pdf

FAMILY VIOLENCE AND FOOTBALL: THE EFFECT OF UNEXPECTED EMOTIONAL CUES ON VIOLENT BEHAVIOR (2018). *Nber.org.* Retrieved from http://www.nber.org/papers/w15497.pdf

GamesCricketRugbyEnduranceCFL, &., Sports. (2018). *NFL players who protested during national anthem in Week 2.* ESPN.com. Retrieved from http://www.espn.com/blog/nflnation/post/_/id/214580/nfl-players-who-protested-during-national-anthem-in-week-2

Lee, Q. (2017). *OLS, BLUE and the Gauss Markov Theorem | UW Economics Society. Uweconsoc.com*. Retrieved from http://uweconsoc.com/ols-blue-and-the-gauss-markov-theorem/

National Football League (NFL) television ratings and viewership figures in the United States in the 2017/2018 season, g. (2018). *NFL viewership figures 2017-18 | Statistic. Statista.* Retrieved from https://www.statista.com/statistics/749140/nfl-viewership-figures-ratings/

NFL Ratings 2017 Compared to Last Year - Page 3 of 4 - Sports Media Watch. (2018). *Sports Media Watch.* Retrieved from http://www.sportsmediawatch.com/nfl-tv-ratings-viewership-nbc-cbs-fox-espn-nfln-regular-season-playoffs/3/

NFL players who protested during national anthem in Week 17. (2018). *ESPN.com.* Retrieved from http://www.espn.com/blog/nflnation/post/_/id/264431/nfl-players-who-protested-during-national-anthem-in-week-17

NFL players who protested during the national anthem in Week 17. (2017). *ESPN.com.* Retrieved from http://www.espn.com/blog/nflnation/post/_/id/227095/nfl-players-who-protested-during-the-national-anthem-in-week-17

NFL.com Scores. (2015). *Nfl.com.* Retrieved from http://www.nfl.com/scores/2015/REG1

NFL.com Scores. (2016). *Nfl.com.* Retrieved from http://www.nfl.com/scores/2016/REG1

NFL.com Scores. (2017). *Nfl.com.* Retrieved from http://www.nfl.com/scores/2017/REG1

NFL Ratings 2017 Compared to Last Year - Sports Media Watch. (2018). *Sports Media Watch.* Retrieved from http://www.sportsmediawatch.com/nfl-tv-ratings-viewership-nbc-cbs-fox-espn-nfln-regular-season-playoffs/

Capstone Reflection

My capstone project was a sizeable learning experience for me. I actually started working on the project back in Fall 2016 when I learned about it in the Huntsman Scholar program. I had no idea what I wanted to do when I first started, and I spent the better part of a year trying to decide what I wanted to do. Eventually, I decided on a project about differential tuition. The idea came about when I took an Econometrics class with Briggs Depew. I really enjoyed the class, and the study of econometrics was more than just an enjoyable school topic – I was interested in seeing if there was anything I could do to apply it to my life. It happened that at that time there was a lot of discussion in the business school about differential tuition, and although I did not enjoy being involved in those conversations – since I felt that most people wished to simply express their opinions not base any of them in fact – I realized that it would be a perfect opportunity to use econometric modeling to create some tests.

From what I was learning in econometrics, I would be able to set up a model that could control for many of the exogenous variables that were the source of a lot of the bias in opinions about differential tuition. I talked with my professor in that class, Briggs Depew, and we were able to work it out into a type of honors capstone. This is where the challenges began. My original goal was to set up a model that could test to see if differential tuition across the US was meeting the outcomes it says it is implemented for. However, many schools have different goals, and so my first challenge was deciding how to go about standardizing the test to account for the different goals among schools. In the end, I couldn't even get that far. As I started looking for data about certain statistics (funding channels for the differential tuition, salary data, enrollment data, etc) neither I nor Dr. Depew could find any available public database that was sufficient for the tests; in order to get that data, I would need to get it myself. We decided to scale down the project and focus on Utah, because I could use a GRAMA request to get the data from public universities in the state.

That didn't work either – it took much longer than I planned to prepare an accurate GRAMA request that would be able to get me the information that I needed, and in January Dr. Depew and I decided to change the project to the NFL protests. It was a timely and relevant project – Trump had recently spoken out about it, and it had been going on for two years without anyone doing a study on the data. The data (we thought) would be easily accessible online, and I could quickly get down to the analysis, so I started the data gathering. It ended up taking over month however, when, as I mention in the actual project, I realized that only the national games had publicly available data.

This was one of the bigger learning points of the project for me, because we were at a wall –there was no time to change the project, but the data was unusable as it was. So, we had to find a way around it, which required me to rely a lot of Dr. Depew's experience in the field. We were able to change the setup of the model to aggregate the protest data, bypassing the need to have individualized data from regional games. It was something that I had never learned in class, and it was a great experience to learn a different style of modeling. I was able to finish the project soon afterward – I had already collected all the data, and with the STATA program, I was easily able to incorporate the protest index.

Lessons Learned

Over the course of the last year and a half, I think that there are three major lessons that I learned about capstone research and writing:

First, it was good that I switched the project multiple times. Doing a capstone project requires well over a hundred of hours of work, and had I decided to work on any of the early ideas for my project, I would not have had the desire to put in an adequate level of effort and analysis into the ideas. Switching my ideas multiple time allowed me to decide what interested me, refined my approach to the project (the scope of analysis, necessary data requirements, outcome variables, etc), and helped me to commit to the level of work necessary to have a measure of success/results.

Second, having a good mentor was critical. Dr. Depew is an extremely skilled economist, and was able to efficiently guide me to the right answers for my project. When I encountered complex problems, I was always able to go back to him to fix errors and receive feedback. Without his help, my project would have been entirely different, and I was so grateful to have his support and leadership.

Third, it was good that I started (relatively) early. My first project failed because I often delayed working on the GRAMA requests. Coming into the NFL protest project, I set a much better standard by biting off small chunks of the project at a time, which (i) made the project manageable both in terms of time and stress, and (ii) alerted me to unforeseen challenges months in advance, allowing Dr. Depew and I to tackle them without having to frantically rush to put something together. I would have preferred if I had started three or four more months earlier, but that was all the time we had, and I tried to make good on it.

As a whole, the project was a growth experience that stretched my ability to manage my time and work independently. I am happy I did the project and am pleased with the results that both yielded results and left room for me to analyze the topic further in the future, which I think is a good sign.

Authors Biography

Jameson Osmond was born in Orem, Utah, and raised in Provo, Utah; Issaquah, Washington; and South Jordan, Utah. He graduated from Herriman High School in 2012, entering Utah State University that fall as a Huntsman Scholar, Pearson Scholar, and presidential scholarship recipient. He began his studies in Business Administration, but after serving a two year mission in Tegucigalpa, Honduras for The Church of Jesus Christ of Latter-day Saints he switched to a triple major in Finance, Economics, and International Business, with a minor in Human Resource Management.

After taking the Econometrics class offered at the university, Jameson began working with Dr. Briggs Depew on a project concerning differential tuition. Soon after, they began working on this project of NFL protests.

Jameson will graduate in the spring of 2018, and hopes to continue his work in venture capital for a few years before continuing with his graduate studies. He enjoys business and hopes to work in a startup in the future.