# Golden Arm or Fool's Gold?

The Harker School
ATE Econometrics
Mr. Lizardo
Winter 2023
Athreya Daniel, Rohan Gorti, Zeke Weng

## I. Abstract

The paper attempts to understand whether NFL Front Offices are justified in paying quarterbacks large amounts of money. To examine this question, we explored the relationship between the amount a quarterback is paid as a percentage of total salary that cap in a year with the performances of the player's team—factoring in both regular season and postseason results. We initially theorized that there would be a moderately strong positive relationship because the Quarterback is both the critical locker room presence and the centerpiece of the team's offense. Our results indicate that there is in fact a moderately strong association between the percentage of salary that a team pays its quarterback and the team's winnings (as defined by the win index we outline in Section III).

## II. Introduction and Purpose

On July 6, 2020, Patrick Mahomes signed an extension with a max payout of $503 million and $141 million guaranteed through 2031. But is one player really worth all this money? With an annual cap hit of approximately $50.3 million, Mahomes' contract alone constitutes about 20% of the Chiefs available salary cap in 2022.

The salary cap (arguably the financial bedrock of the National Football League (NFL)) dictates the amount NFL franchises can pay their players annually. This cap is a fixed amount of money (an estimated $208.2 million in 2023) allocated by the league that teams cannot exceed in roster-based spending. Each year, front offices across the league must allocate their available cap across upwards of 70 players (exceeding the 53 man roster size due to roster cuts, mid-season pick ups, and elevated practice squad players) to create the most prolific team on the field. Ideally, teams and players can reach an agreement for the players' value on the field with money as compensation; simply, the better the player, the more they should be paid.

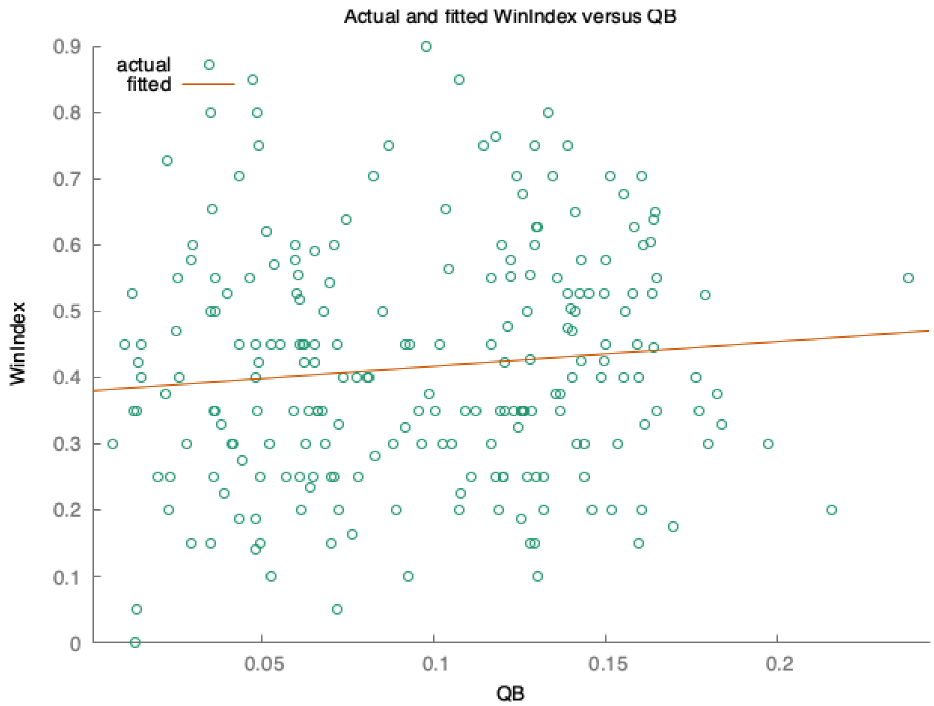
As the facilitator, playmaker, and leader of the team, the quarterback (QB) is the focal point of the NFL offense. Quarterbacks handle the ball on nearly every offensive play from scrimmage, and it is widely held that they are among the most vital components of a successful franchise. One of the most pressing questions faced by any NFL team, therefore, is how much of their salary cap should be allocated to the quarterbacks on their roster in order to maximize their in-season success. Investigating this basic relationship between the percentage of salary cap a team allocates to its quarterbacks and the team's winnings is the primary objective of our paper. However, our investigation follows three key assumptions:

1.) We assume that players are fairly compensated for their performance on the field through the amount they are paid. Therefore, the amount a player is paid acts as a proxy for his performance in our model.
2.) The variable that we want to observe is the QB pay and performance; however other variables like other positions also have an impact on QB pay and performance along with the winning. Hence, we included all the positional groups on the field ranging from QB to DB to K/P/LS.
3.) Players act as another control variable. Having over the standard roster size of 53 is a product of compensation for players away from the team, dead cap, and Injury Reserves. Roster sizes that are on the larger side usually indicate more substitutions/manipulations of the roster which shuffle team chemistry, injuries, and other internal unrest which would negatively impact winning.

## III. Data

The majority of our data is collected from a population of salary metrics collected from league-wide metrics. Though we collect financial metrics from the 2015-2022 NFL seasons, our data remains cross-sectional (rather than time-series) because we seasonal salaries at specific points in time. In other words, our data does not inherently depend on (and is not influenced by) any measure of time.

The key relationship we seek to test in this paper is between the percentage of the salary cap an NFL team allocates to its quarterback and their in-season success. We denote our key independent variable (% of cap spent on QB) by ***QB (1)***, and our key dependent variable (in-season winnings) as ***WinIndex (11)***. A scatterplot of the data representing this relationship as well as a simple regression line is shown below. Based on this plot, we observe a weak, positive relationship between the percentage of the team's salary cap allocated to quarterbacks and the team's winnings (based on our calculated win-index).




We also show descriptions and statistics for all of our variables in the table below:

| Variable(#) | Description | Unit | Source |
|---|---|---|---|
| QB (1) | The percentage of total league cap space allocated to all of the quarterbacks on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| RB/FB (2) | The percentage of total league cap space allocated to all of the halfbacks and fullbacks on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| WR (3) | The percentage of total league cap space allocated to all of the wide receivers on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| TE (4) | The percentage of total league cap space allocated to all of the tight ends on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |

| | | | |
|---|---|---|---|
| OL (5) | The percentage of total league cap space allocated to all offensive linemen (Center, Guard, Tackle) on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| DL (6) | The percentage of total league cap space allocated to all defensive linemen (Defensive End, Defensive Tackle) on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| LB (7) | The percentage of total league cap space allocated to all of the linebackers on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| DB (8) | The percentage of total league cap space allocated to all of the defensive backs on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| K/P/LS (9) | The percentage of total league cap space allocated to all of the kickers, punters, and long snappers on a given team's roster. | N/A | https://www.spotrac.com/nfl/positional/breakdown/ https://www.spotrac.com/nfl/cap/ |
| Players (10) | A quantitative variable indicating the number of players on a team's roster. | Individuals | https://www.spotrac.com/nfl/positional/breakdown/ |
| WinIndex (11) | A measure of a team's in-season success in a given year. We computed our own measure of a team's "winnings" based on a weighted summation of their regular season win percentage (80%) and the number of Wild Card, Divisional Round, Conference Championship, and Super Bowl wins they | N/A | https://calbizjournal.com/nfl-paychecks-how-do-nfl-players-get-paid/ |

| | | | |
|---|---|---|---|
| | achieved (20%). The data was approximately normal and any transformation we tried did not yield more normality. | | |

Table 1: Data Description

To calculate ***WinIndex (11)***, we weigh the Regular Season at 80% because it consists of 16-17 Games and the Postseason at 20% with a total of 4 games playable. The variables ***Wildcard, Div, Conf,*** and ***Superbowl*** are dummy variables, with a value of 1 representing a win and a value of 0 representing a loss. In our specification, a first-round bye is equivalent to a Wildcard win. The resulting equation is as follows:

***WinIndex = .8 RegWin% + .029 Wildcard + .029 Div + .044 Conf + .101 Superbowl***

We compute the coefficient for playoff games using the monetary bonus players on each team receive for a win. These values range from $40k for Wild Card wins to $150k for a Superbowl victory.

It is also worth noting that the variables we show below are percentages, but are represented in decimal form; hence our regression interpretations require division by 100 (an additional unit in the model corresponds to 100% increase in a player's pay out of the total salary cap). To interpret this correctly, we focus on the effect of a 1% increase in positional spending.

| Variable Name | Number of Observations | Mean | Standard Deviation | Minimum | Maximum | Skewness |
|---|---|---|---|---|---|---|
| QB | 224 | 9.611% | 4.902% | 0.659% | 23.814% | 0.08833 |
| RB/FB | 224 | 4.342% | 2.030% | 0.956% | 12.144% | 0.80983 |
| WR | 224 | 9.754% | 3.996% | 1.749% | 20.877% | 0.34390 |
| TE | 224 | 4.421% | 2.117% | 1.096% | 11.446% | 0.65940 |
| OL | 224 | 16.004% | 4.369% | 4.057% | 28.055% | 0.23877 |
| DL | 224 | 14.321% | 5.857% | 3.556% | 29.802% | 0.38608 |
| LB | 224 | 11.230% | 5.313% | 2.082% | 27.288% | 0.55844 |
| DB | 224 | 15.046% | 5.247% | 33.156% | 5.088% | 0.41551 |
| K/P/LS | 224 | 26.519% | 1.149% | 0.773% | 6.569% | 0.47151 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Players | 224 | 65.88 | 4.625 | 56 | 79 | 0.21889 |
| WinIndex | 224 | 0.41606 | 0.177 | 0.000 | 0.900 | 0.31046 |

Table 2: Descriptive Statistics

# IV. Assumption Check

1. MLR. 1

   Our regression equation is linear, so our model satisfies MLR. 1.

2. MLR. 2

   Though we did not employ random sampling in our data collection, our population consists solely of our dataset. Any conclusions drawn from this data, therefore, still satisfy MLR. 2.

3. MLR. 3

   No perfect collinearity exists in our data.

4. MLR. 4

   Because we are primarily interested in the financial determinants of team-winning, our regressions do not include other variables that could impact in-season success. These include positional skill indices and certain coaching metrics, both of which were difficult to obtain data on and could potentially contribute to omitted variable bias (OVB) in our models.

5. MLR. 5

   Using the Breush-Pagan and White tests, we evaluated our models for heteroskedasticity. However, it did not factor into any one of them.

6. MLR. 6

   All of our variables appeared normally distributed, and therefore did not require any transformations for normality.

# V. Regressions

## A. Simple Regression

In our simple regression, we regress ***WinIndex*** on ***QB***. The resulting equation is as follows:

$$\boldsymbol{WinIndex = \beta_0 + \beta_1 QB}.$$

We expect a positive coefficient on ***QB*** because assuming quarterbacks are compensated correctly for their performance, the better quarterbacks with a larger impact on winning would be paid more, illustrating $\beta_1 > 0$. However, because other positional salary cap variables are omitted in this model, we expect OVB. First, the portion of a team's cap allotted to quarterbacks is somewhat inversely related to the portion of the cap allotted to other positions on the team (due to a fixed amount of cap). However, the correlation between a general form of positional pay is less clear—it is unclear whether higher paid players consistently perform better. Therefore, though the direction of our OVB is likely negative, the strength is somewhat ambiguous.  It is also worth noting that constant $\beta_0$ has a meaningless interpretation because every team needs a quarterback who is paid simply for participating in professional competition.

```
Model 1: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       0.380231     0.0260747    14.58     3.12e-34 ***
  QB          0.372743     0.233876      1.594    0.1124

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    6.894359   S.E. of regression   0.176226
R-squared            0.010684   Adjusted R-squared   0.006227
F(1, 222)            2.540086   P-value(F)           0.112411
Log-likelihood       72.02333   Akaike criterion    -140.0467
Schwarz criterion   -133.2234   Hannan-Quinn        -137.2924
```

Our model demonstrates that a 1% increase in ***QB*** (percentage of total salary cap paid towards QB)  corresponds to a .0037 increase in ***WinIndex*** (***QB*** data points are decimal values rather than percentages). However, this coefficient is not significant at any level (10%, 5%, 1%) with a ***p***-value of 0.1124. Therefore, we cannot reject the null hypothesis and $\beta_1$ has an essentially meaningless interpretation. The $R^2$ value of 0.011 and the **adjusted** $R^2$ value of 0.006 demonstrate that very little of the variation in team winnings can be explained by ***QB***.

## B. Multiple Regressions

In our multiple regression, we regress ***WinIndex*** on all other variables (***QB, RB/FB, WR, TE, OL, DL, LB, DB, K/P/S, Players***). We account for Multicollinearity through our choice of independent variables (Section V. Part F). We also account for heteroskedasticity by using Robust Standard Error (this is further discussed in Section YY). The variable ***Players*** controls for differences in the total number of players on a team's roster. Our final equation is shown below:

$$\boldsymbol{WinIndex = \beta_0 + \beta_1 QB + \beta_2 RB/FB + \beta_3 WR + \beta_4 TE + \beta_5 OL + \beta_6 DL + \beta_7 LB + \beta_8 DB + \beta_9 K/P/LS + \beta_{10} Players}$$

We expect a positive coefficient for $\beta_1$ through $\beta_9$—under the assumption all players in their respective positions are compensated correctly for their on-field performance. However, we expect a negative coefficient for $\beta_{10}$ for two reasons. First, a greater number of players on a roster could signify an increase in injuries to key players. Second, it could represent a team's diminished confidence in its starting players. As a result, we expect the coefficient on $\beta_{10}$ to be at least moderately negative

Model 2: OLS, using observations 1–224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

| | coefficient | std. error | t-ratio | p-value | |
|---|---|---|---|---|---|
| const | 0.959625 | 0.205628 | 4.667 | 5.41e-06 | *** |
| QB | 0.623183 | 0.235635 | 2.645 | 0.0088 | *** |
| RBFB | 1.06003 | 0.549947 | 1.928 | 0.0552 | * |
| WR | 0.412859 | 0.260175 | 1.587 | 0.1140 | |
| TE | 1.23882 | 0.503237 | 2.462 | 0.0146 | ** |
| OL | 0.526251 | 0.239358 | 2.199 | 0.0290 | ** |
| DL | 0.319354 | 0.213315 | 1.497 | 0.1358 | |
| LB | 0.466262 | 0.229331 | 2.033 | 0.0433 | ** |
| CB | 0.542038 | 0.222064 | 2.441 | 0.0155 | ** |
| KPLS | 1.81940 | 0.802070 | 2.268 | 0.0243 | ** |
| Players | −0.0160369 | 0.00219157 | −7.318 | 5.08e-12 | *** |

| | | | |
|---|---|---|---|
| Mean dependent var | 0.416056 | S.D. dependent var | 0.176777 |
| Sum squared resid | 4.856188 | S.E. of regression | 0.150993 |
| R-squared | 0.303154 | Adjusted R-squared | 0.270438 |
| F(10, 213) | 10.77389 | P-value(F) | 9.42e-15 |
| Log-likelihood | 111.2737 | Akaike criterion | −200.5474 |
| Schwarz criterion | −163.0193 | Hannan-Quinn | −185.3992 |

Excluding the constant, p-value was highest for variable 8 (DL)

Our model demonstrates largely positive associations between positional cap allocations and teams' winIndex. For instance, a 1% increase in ***QB*** (percentage of total salary cap paid towards QB) corresponds to a .0062 increase in ***WinIndex*** (***QB*** data points are decimal values rather than percentages). The variables ***QB, RB/FB, TE, OL, LB, DB, K/P/S,*** and

***Players*** demonstrate some degree of significance. The two variables that did not show significance to the 10%, 5%, or 1% level were ***WR*** and ***DL.*** The variables with the most significant were ***QB*** and ***Players***, with ***p***-values of 0.0088 and $5.08 \cdot 10^{-12}$ respectively. According to our model, the coefficients $\beta_2$ and $\beta_4$ were the largest in magnitude at 1.239 and 1.060, respectively, while $\beta_{10}$ was the smallest at -.016. This finding aligns with our initial expectation that the number of players on a roster demonstrates a weak correlation with a team's winnings, but seems to contradict our initial hypothesis that the percent of the salary cap a team allocates to its quarterback is most vital to their overall success. Additionally, with an $R^2$ value of 0.303, and an **adjusted** $R^2$ value was 0.270, our regression explains 27% of the variation in our independent variables.

To extract as much significance from our model as possible, we ran three F-Tests on groupings of individually insignificant variables. Our results are detailed in the sections below.

## C. F-Test 1: Omitted WR and DL Model

First, we omit the two variables that were insignificant in our MLR. We find that the variables were not jointly significant at any of the 10%, 5%, or 1% levels. We also note that the **adjusted** $R^2$ decreases from .270 to .262, and the ***P-value*** corresponding to our F-Test is 0.110572. This demonstrates that variables WR and DL are jointly insignificant. Furthermore, the variable ***LB*** became insignificant with these omissions, with a p-value of 0.1426.

```
Test on Model 6:

  Null hypothesis: the regression parameters are zero for the variables
    WR, DL
  Test statistic: Robust F(2, 213) = 2.22501, p-value 0.110572
  Omitting variables improved 2 of 3 information criteria.

Model 7: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       1.21094      0.162953      7.431    2.51e-12 ***
  QB          0.482853     0.222255      2.173    0.0309   **
  RBFB        0.912018     0.534853      1.705    0.0896   *
  TE          1.24948      0.496114      2.519    0.0125   **
  OL          0.483049     0.243477      1.984    0.0485   **
  LB          0.273403     0.185806      1.471    0.1426
  CB          0.490306     0.218800      2.241    0.0261   **
  KPLS        1.44260      0.821070      1.757    0.0803   *
  Players    -0.0175474    0.00206750   -8.487    3.46e-15 ***

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    4.959896   S.E. of regression   0.151886
R-squared            0.288272   Adjusted R-squared   0.261789
F(8, 215)            12.95978   P-value(F)           3.38e-15
Log-likelihood       108.9070   Akaike criterion    -199.8141
Schwarz criterion   -169.1093   Hannan-Quinn        -187.4201

Excluding the constant, p-value was highest for variable 9 (LB)
```

## D. F-Test 2: Omitted WR, DL, LB Model

Our second F-Test excludes variables ***WR, DL,*** and ***LB*** (we include ***LB*** because it demonstrates individual insignificance in our last test). After our omissions, we find that the insignificant variables are not jointly significant at any of the 10%, 5%, or 1% levels. We note that the model's **adjusted** $R^2$ decreases from .262 to .258, and the test's p-value is 0.100349, indicating joint insignificance. Moreover, the variable ***RB/FB*** also demonstrates individual insignificance. Thus, we run a final F-Test with the omissions of ***WR, DL, LB,*** and ***RB/FB***—all the variables that show individual insignificance from our previous regressions.

```
Test on Model 6:

  Null hypothesis: the regression parameters are zero for the variables
    WR, DL, LB
  Test statistic: Robust F(3, 213) = 2.10695, p-value 0.100349
  Omitting variables improved 2 of 3 information criteria.

Model 11: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       1.26542      0.157016      8.059    5.18e-14 ***
  QB          0.446510     0.216210      2.065    0.0401   **
  RBFB        0.823264     0.526301      1.564    0.1192
  TE          1.24773      0.507355      2.459    0.0147   **
  OL          0.442374     0.241008      1.836    0.0678   *
  CB          0.442556     0.214179      2.066    0.0400   **
  KPLS        1.53763      0.817102      1.882    0.0612   *
  Players    -0.0176260    0.00207363   -8.500    3.12e-15 ***

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    5.004256   S.E. of regression   0.152210
R-squared            0.281907   Adjusted R-squared   0.258635
F(7, 216)            14.51820   P-value(F)           1.89e-15
Log-likelihood       107.9098   Akaike criterion    -199.8196
Schwarz criterion   -172.5264   Hannan-Quinn        -188.8027

Excluding the constant, p-value was highest for variable 4 (RBFB)
```

## E. F-Test 3: Omitted WR, DL, LB, RB/FB

Omitting ***WR, DL, LB,*** and ***RB/FB***, we find that they are not jointly significant at any of the 10%, 5%, or 1% levels. We also note that the **adjusted** $R^2$ decreases from .258 to .253, and the corresponding ***P-value (F)*** is 0.116053.

To start, all variables demonstrate some level of individual significance. In terms of overall explanatory fit, we note a decrease in the **adjusted** $R^2$ from .270 (in the initial regression) to

.253 (in the final one). Also, due to the aforementioned OVB in the simple regression and the variables acting as controls, our best model is still the initial MLR that includes all variables.

```
Test on Model 6:

  Null hypothesis: the regression parameters are zero for the variables
    RBFB, WR, DL, LB
  Test statistic: Robust F(4, 213) = 1.87444, p-value 0.116053
  Omitting variables improved 2 of 3 information criteria.

Model 10: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const       1.29698      0.153537      8.447    4.30e-15 ***
  QB          0.398329     0.208404      1.911    0.0573   *
  TE          1.20694      0.506935      2.381    0.0181   **
  OL          0.439971     0.241773      1.820    0.0702   *
  CB          0.439342     0.215076      2.043    0.0423   **
  KPLS        1.65185      0.811939      2.034    0.0431   **
  Players    -0.0174976    0.00206025   -8.493    3.21e-15 ***

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    5.064804   S.E. of regression   0.152775
R-squared            0.273218   Adjusted R-squared   0.253123
F(6, 217)            16.77222   P-value(F)           7.04e-16
Log-likelihood       106.5628   Akaike criterion    -199.1256
Schwarz criterion   -175.2441   Hannan-Quinn        -189.4858
```

## F. Multicollinearity

The independent variables in our model do not demonstrate any multicollinearity; no factor's Variance Inflation Factor (VIF) exceeds the generally accepted 10.0 threshold. In other words, the amount a particular position is paid is not a significant determinant of another position's pay.

```
Variance Inflation Factors
Minimum possible value = 1.0
Values > 10.0 may indicate a collinearity problem

         QB    1.214
       RBFB    1.101
         WR    1.139
         TE    1.047
         OL    1.106
         DL    1.469
         LB    1.407
         CB    1.200
       KPLS    1.126
    Players    1.119

VIF(j) = 1/(1 - R(j)^2), where R(j) is the multiple correlation coefficient
between variable j and the other independent variables
```

## G. Collated Regression Results

The table below contains the results from each of the regressions we discuss above.

**Simple Regression**

```
Model 1: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       0.380231     0.0260747    14.58     3.12e-34 ***
  QB          0.372743     0.233876      1.594    0.1124

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    6.894359   S.E. of regression   0.176226
R-squared            0.010684   Adjusted R-squared   0.006227
F(1, 222)            2.540086   P-value(F)           0.112411
Log-likelihood       72.02333   Akaike criterion    -140.0467
Schwarz criterion   -133.2234   Hannan-Quinn        -137.2924
```

**Multiple Regression (Best Model Still)**

```
Model 2: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       0.959625     0.205628      4.667    5.41e-06 ***
  QB          0.623183     0.235635      2.645    0.0088   ***
  RBFB        1.06003      0.549947      1.928    0.0552   *
  WR          0.412859     0.260175      1.587    0.1140
  TE          1.23882      0.503237      2.462    0.0146   **
  OL          0.526251     0.239358      2.199    0.0290   **
  DL          0.319354     0.213315      1.497    0.1358
  LB          0.466262     0.229331      2.033    0.0433   **
  CB          0.542038     0.222064      2.441    0.0155   **
  KPLS        1.81940      0.802070      2.268    0.0243   **
  Players    -0.0160369    0.00219157   -7.318    5.08e-12 ***

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    4.856188   S.E. of regression   0.150993
R-squared            0.303154   Adjusted R-squared   0.270438
F(10, 213)           10.77389   P-value(F)           9.42e-15
Log-likelihood       111.2737   Akaike criterion    -200.5474
Schwarz criterion   -163.0193   Hannan-Quinn        -185.3992

Excluding the constant, p-value was highest for variable 8 (DL)
```

**Omitted Offense**

```
Test on Model 2:

  Null hypothesis: the regression parameters are zero for the variables
    QB, RBFB, WR, TE, OL
  Test statistic: Robust F(5, 213) = 3.75535, p-value 0.00279941
  Omitting variables improved 1 of 3 information criteria.

Model 3: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       1.37212      0.179554      7.642    6.72e-13 ***
  DL          0.227494     0.211957      1.073    0.2843
  LB          0.271422     0.214172      1.267    0.2064
  CB          0.248062     0.210494      1.178    0.2399
  KPLS        1.90656      0.808704      2.358    0.0193   **
  Players    -0.0168013    0.00222211   -7.561    1.10e-12 ***

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    5.295668   S.E. of regression   0.155859
R-squared            0.240090   Adjusted R-squared   0.222661
F(5, 218)            18.61124   P-value(F)           2.19e-15
Log-likelihood       101.5705   Akaike criterion    -191.1411
Schwarz criterion   -170.6712   Hannan-Quinn        -182.8785

Excluding the constant, p-value was highest for variable 8 (DL)
```

Omitted Defense

```
Test on Model 2:

  Null hypothesis: the regression parameters are zero for the variables
    DL, LB, CB
  Test statistic: Robust F(3, 213) = 2.83532, p-value 0.0391294
  Omitting variables improved 2 of 3 information criteria.

Model 4: OLS, using observations 1-224
Dependent variable: WinIndex
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const       1.25154      0.171154      7.312    5.05e-12 ***
  QB          0.391341     0.218435      1.792    0.0746   *
  RBFB        0.971072     0.555786      1.747    0.0820   *
  WR          0.378550     0.274377      1.380    0.1691
  TE          1.12991      0.504370      2.240    0.0261   **
  OL          0.357457     0.242574      1.474    0.1420
  KPLS        1.82984      0.865353      2.115    0.0356   **
  Players    -0.0168143    0.00212404   -7.916    1.27e-13 ***

Mean dependent var   0.416056   S.D. dependent var   0.176777
Sum squared resid    5.058351   S.E. of regression   0.153030
R-squared            0.274144   Adjusted R-squared   0.250621
F(7, 216)            14.15221   P-value(F)           4.35e-15
Log-likelihood       106.7056   Akaike criterion    -197.4112
Schwarz criterion   -170.1180   Hannan-Quinn        -186.3943

Excluding the constant, p-value was highest for variable 5 (WR)
```

# VI. Heteroskedasticity

We test our MLR model for Heteroskedasticity using the Breusch-Pagan test. As a whole, heteroskedasticity is not present in our model; our ***p***-value of 0.188 does not pass any significance test at (10%, 5%, 1%) thereby not rejecting the null hypothesis that heteroskedasticity is not present in the regression. The variables ***TE, DL,*** and ***LB***, have some degree of heteroskedasticity at the 10% and 5% significance levels. We account for this in our models by using the Robust Standard Error.

```
Breusch-Pagan test for heteroskedasticity
OLS, using observations 1-224
Dependent variable: scaled uhat^2

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const       3.46617      1.82753       1.897    0.0592   *
  QB         -2.63299      2.01485      -1.307    0.1927
  RBFB        0.612364     4.63499       0.1321   0.8950
  WR          0.947172     2.39456       0.3956   0.6928
  TE          8.45587      4.33375       1.951    0.0523   *
  OL          2.05756      2.15801       0.9535   0.3414
  DL         -3.22057      1.85497      -1.736    0.0840   *
  LB         -5.08297      2.00117      -2.540    0.0118   **
  CB         -2.61622      1.92963      -1.356    0.1766
  KPLS        3.78963      8.27810       0.4578   0.6476
  Players    -0.0259536    0.0205013    -1.266    0.2069

  Explained sum of squares = 27.3437

Test statistic: LM = 13.671840,
with p-value = P(Chi-square(10) > 13.671840) = 0.188493
```

# VIII. Conclusion

In conclusion, as we hypothesized, NFL Front Offices are justified in their large spendings on the Quarterback position. However, our findings are not so straightforward. Due to differences in the magnitude of their coefficients, some positions merit a deeper dive into understanding positional value. For instance, TE has the largest coefficient at 1.24 but teams only allocate about 4.4% of the total salary cap to this position. By contrast, the LB position may need to be scaled back in terms of pay with a smaller coefficient of .31 in conjunction with the high salary percentage of 14.3%. According to our model, since all positions have positive coefficients, teams should spend to the total cap limit. However, our data only spans between 2015 through 2021 and our findings should not be extrapolated beyond this timeframe. Our model attempts to simplify the exceedingly complex task of understanding winning in the National Football League.

There are two primary sources of error we note in this project. The first is that by designating cap spending as a proxy for player skill, we assume a relatively direct relationship between the two; however, this may not always be the case. In other words, greater amounts of a team's cap alloted on players do not always correspond with higher levels of skill or greater levels of success. Players in "prove-it years" (contract expiration years) and rookies (who usually do not have max contracts) may perform just as well if not better than tenured veterans. The second (mentioned briefly in Section III as well) is that the variables we include in our specifications are not the only factors that affect a team's winnings. Though we discuss OVB and potential collinearity in our analysis, it is worth reiterating that factors like coach salary and positional skill index could also play roles in determining team success.

# IX. References

1. https://www.spotrac.com/nfl/positional/breakdown/{year}/
2. https://www.spotrac.com/nfl/cap/
3. https://calbizjournal.com/nfl-paychecks-how-do-nfl-players-get-paid/
4. https://www.pro-football-reference.com/years/{year}/