What Makes a Winner in the n+1 Year?

Jarrett Markman

November 13, 2022

Abstract

The Giants are currently 6-2, with a 6-1 record in one score games. On the other side of the coin, the Jacksonville Jaguars are currently 3-6, with a 1-6 record in one score games. The Giants actually beat the Jaguars in one of these games. While the Giants may have a 6-2 record, their point differential is a lowly 6, and the Jaguars with a 3-6 record have a better point differential of 21. Fans and analysts may attribute the Giants' success to Brian Daboll and the lack of success for the Jaguars just due to the team's "grit" and "ability to win," but, how sustainable is this success and failure? What factors in a year are important for success, and how important are those factors in the following year?

Data Collection

With the library "nflverse", we are able to get team game-by-game and game play-by-play data, that we can use to clean and manipulate team standings and team statistics data. With all the data collected, the next step was to try to find which variables in a year could have an impact on overall production in the following year. The Giants and Jaguars both have different records in one score games, is record in one score games correlated with future success? What about things like offensive and defensive production? Does point differential play a role? All the data that was collected was from 2016-2021.

Modeling Process

Variable Selection Process

With all the data that was collected I decided to evaluate win percentage and point differential in one score games, win percentage and point differential over the course of a season, and the expected points added per play on both offense and defense to measure offensive and defensive talent in the prior season. These are all variables that should have an impact on the following year's win percentage. In plots 1 through 6 (see below), you can see all the data points for every variable and the Y variable (Win Percentage in the n+1 Year), as well as the line of best fit for every data point. The outer bands measure the standard error for each set of variables. The summary stats for each variable can also be seen in the summary stats table below.

Summary Stats in the n Year					
	Min	Median	Mean	Max	Standard Deviation
Win Percentage in One-Score Games	0	0.551	0.53	1	0.181
Win Percentage	0.25	0.563	0.55	0.813	0.114
Offensive EPA	-0.237	0.01	0.008	0.236	0.087
Defensive EPA	-0.217	0.014	0.006	0.207	0.072
Point Differential in One-Score Games	-36	5.5	3.163	36	14.756
Point Differential	-116	28.5	22.94	149	51.169

Creating the Model

With all the data collected, the last step is to create a regression model for what actually makes a winner in the NFL. For the model, I just did a basic linear regression. The results for the model for win percentage in the n+1 year can be seen below (Model).

Regression Equation

Win Percentage in the n+1 Year = .688 -.123(Record in One Score Games in the n Year) - .181(Record in the n Year) - .062(Offensive EPA) + .008(Defensive EPA) + .002(Point Differential in One Score Games in the n Year) + .0004(Point Differential in the n Year)

What do the results mean?

- For every 1 unit increase in win percentage in one score games, win percentage in the n+1 year decreases by .123.
- For every 1 unit increase in win percentage in the n year, win percentage in the n+1 year decreases by .181.
- For every 1 unit increase in offensive epa, win percentage in the n+1 year decreases by .062.
- For every 1 unit increase in defensive epa, win percentage in the n+1 year increases by .008.
- For every 1 unit increase in point differential in one score games, win percentage in the n+1 year increases by .002.
- For every 1 unit increase in point differential for the entire season, win percentage in the n+1 year increases by .0004.
- The constant of .695 or 69.5% indicates the win percentage in the following year if all the variables were 0.

Based on the model, defensive epa (.008), point differential in one score games (.002), and point differential over the course of the year (.0004) all have a positive relationship with win percentage in the following year. Despite these positive relationships however, the betas are very low; this is due in part to the fact that point differential is scaled a lot differently than all the other variables. A point differential of 30, is much larger than a team's win percentage or epa that could be 0.2. Additionally, a team's given point differential will most likely be larger over the course of a season than over the course of uniquely one score games, so the coefficient over the course of the year is larger than that in only one score games.

In contrast, win percentage in one score games (-.123), win percentage over the course of the year (-.181), and offensive epa (-.062) all have a negative relationship with win percentage in the following year. The beta coefficients for these variables are a lot higher, meaning that in the case when there's a higher win percentage in one score games in the n year, the win percentage in the following year will be smaller. There is definitely a flaw within the model, especially given the fact that win percentage in the n year carries such a low coefficient. In the NFL if a team goes 17-0, the chances that they have a low win percentage in the following year are extremely low.

Model Projections

With my model, I was able to predict team records for this current season based on how they performed last year (Expected Record in 2022 Plot). The 4 best teams this year according to the model would be the Patriots, Texans, Packers, and Bears. Additionally, the Minnesota Vikings were the only team to have a win percentage below .500. While my model can detect the relationship between variables, it is not very effective for predicting team win percentage in the following year, which can explain the low R squared of .038.

Conclusion

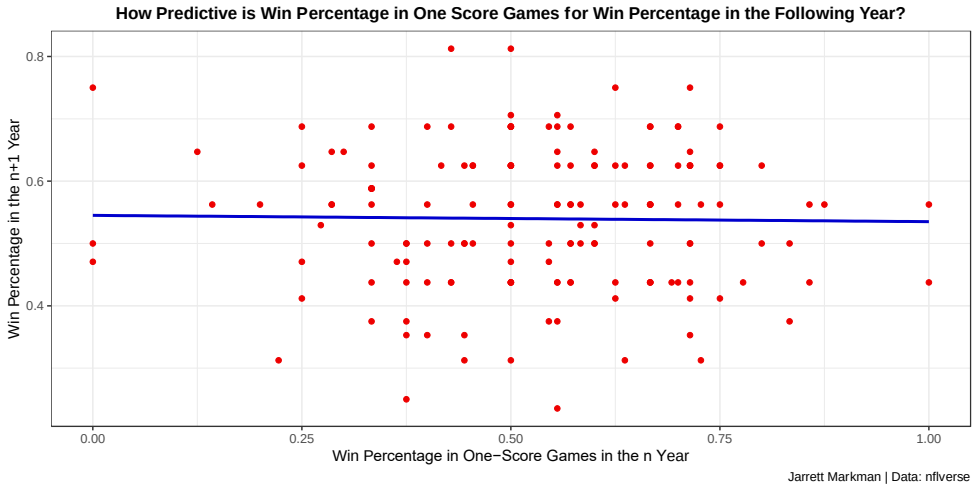
While the Giants may be finding a lot of success this year, there is no evidence that they will continue to thrive in one score situations. Based on the model, the better they do in one score games this year, the more likely they will have a lower win percentage next year. The model agrees with my initial thoughts, that their close wins are not sustainable. Additionally, if the Jaguars are able to continue their trend of success via point differential, they are more likely to

have a high win percentage in the following year. Point differential has a positive correlation with win percentage in the following year and win percentage in close games has a negative correlation with win percentage in the following year.

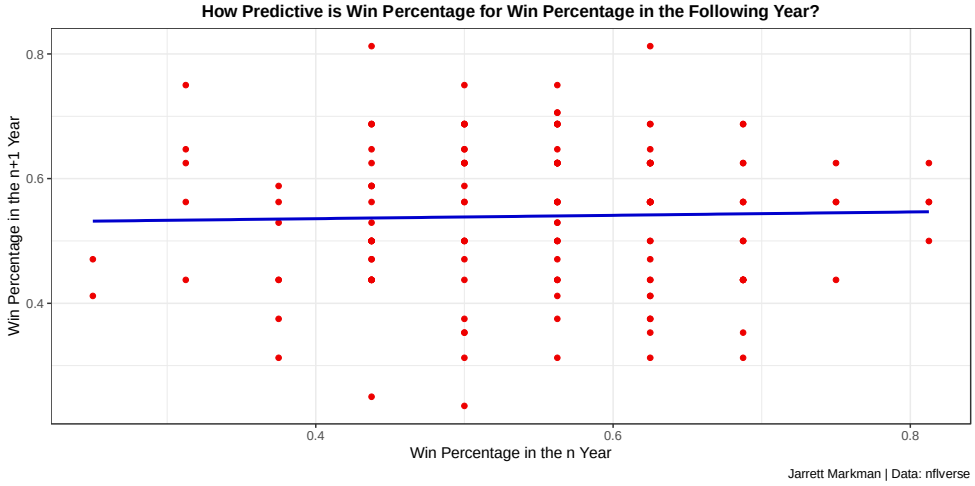
Predicting a winner in the NFL is extremely difficult. Every year is different than the next, and certain teams come out of nowhere and have great years. Based on the model created, only 3.8% of the variation in win percentage in the n+1 year can be explained by all the variables, meaning that the model is not very effective in predicting a winner. It's possible that the statistics used are not the best for predicting year-to-year success, however, it is more likely that it's difficult to predict year-to-year success in the NFL.

Visualizations

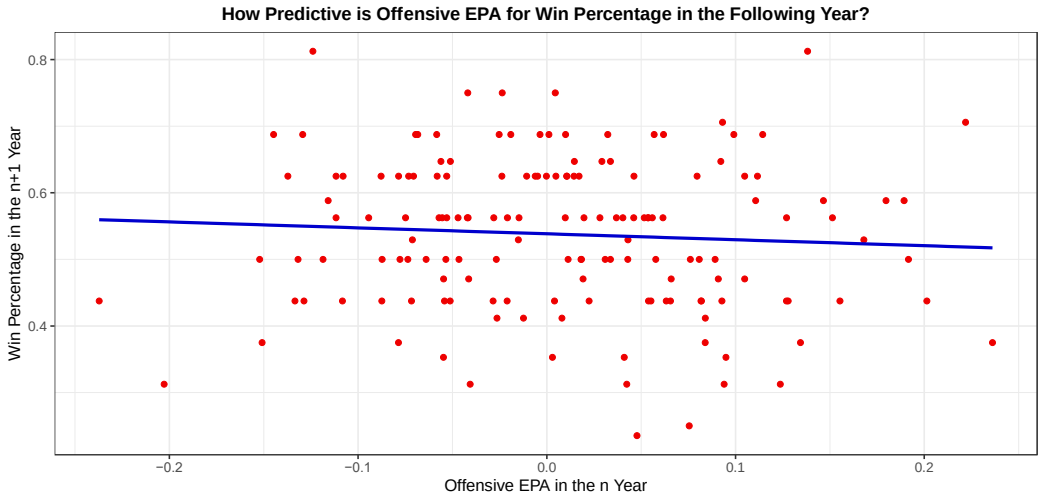
Plot 1



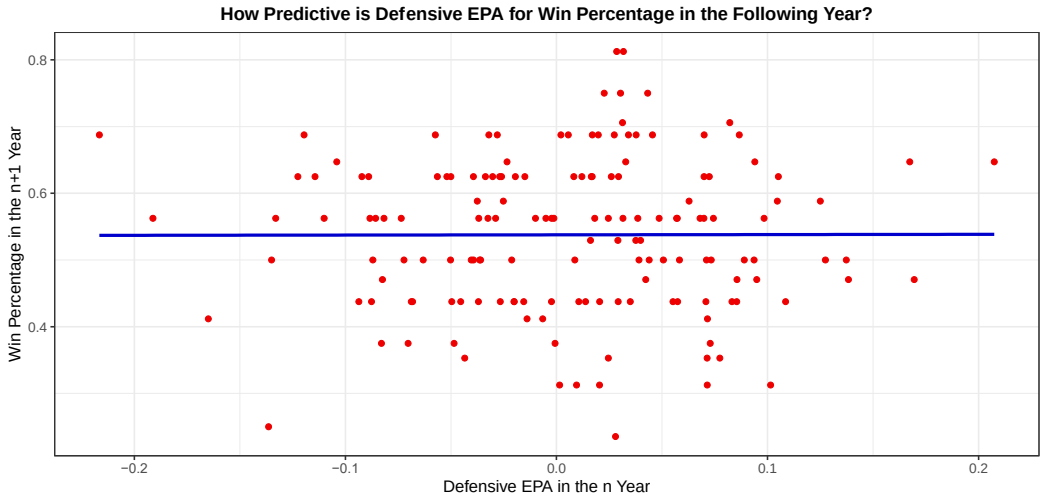
Plot 2



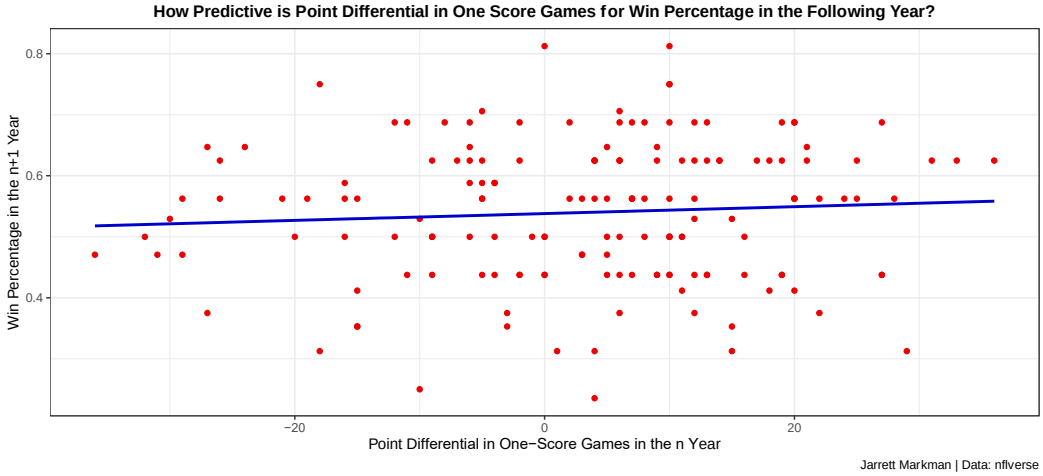
Plot 3



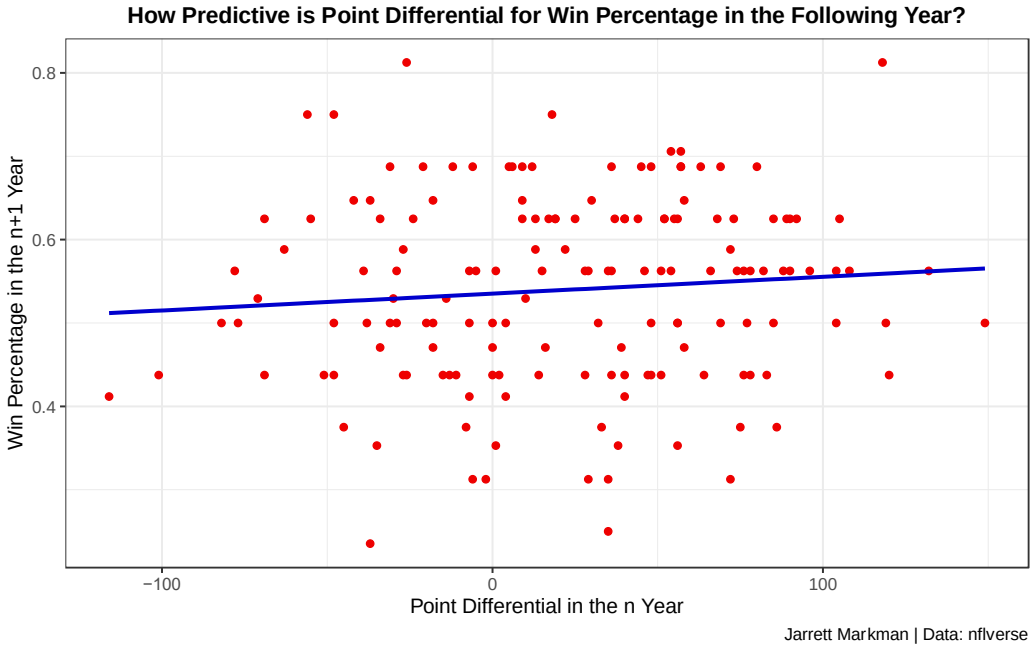
Plot 4



Plot 5



Plot 6


How Predictive is Point Differential for Win Percentage in the Following Year?
Win Percentage in the n+1 Year
0.8
0.6
0.4
−100
0
100
Point Differential in the n Year
Jarrett Markman | Data: nflverse

Model

What Makes a Winner in the NFL?

	Dependent variable:
	Win Percentage in the n+1 Year
Win Percentage in One Score Games in the n Year	-0.123
	(0.132)
Win Percentage in the n Year	-0.181
	(0.238)
Offensive EPA in the n Year	-0.062
	(0.107)
Defensive EPA in the n Year	0.008
	(0.130)
Point Differential in One Score Games in the n Year	0.002*
	(0.001)
Point Differential in the n Year	0.0004
	(0.0004)
Constant	0.688***
	(0.094)
Observations	155
R^2	0.038
Adjusted R^2	-0.001
Residual Std. Error	0.114 (df = 148)
F Statistic	0.986 (df = 6; 148)
Note:	*p<0.1; **p<0.05; ***p<0.01

Expected Record in 2022 Plot