Isolating the Effects of Rushing and Passing Yardage in the NFL

Andrew Katbi, Ohio Northern University
Kurt Kleman, Ohio Northern University

Abstract

Football is a game composed of two primary actions – running and passing the football. As a result, many General Managers in the National Football League center their teams around players who either directly or indirectly are able to help run, pass, or catch the football. However, General Managers also need players who are efficiently able to defend these actions. As a result of the need for such players, teams in the NFL spend millions of dollars annually in order to attract and retain top athletes who are able to contribute to the team and increase the probability of winning games. Also, several studies have indicated a positive association between wins and teams with such players. In this paper, we test the hypothesis that teams who are more effective at running and passing the ball, as well as in possession of a defense which is able to limit the rushing and passing yardage of opponents, win more games. Using time-series panel data regression, we find positive correlations between wins and effective rushing and passing offenses and defenses.

Key Words: Rushing, passing, football, NFL

Introduction

"There are three things that can happen when you pass, and two of them aren't good." These words were once spoken by the legendary college football coach Wayne Woodrow "Woody" Hayes – a man remembered for winning five national titles and 13 Big Ten championships at Ohio State University.

As suggested by Hayes and conventional football wisdom, one would expect that the key to winning games is creating a successful run-offense. After all, such an offense is believed to be more consistent and reliable. Rushing yards are often seen as a safe bet. Passing, on the other hand, can be seen as less attractive due to its large variance of outcomes. For instance, the mean rushing yards per attempt in the NFL between 1990 and 2007 is 4.0 yards per attempt. If a team could run for that amount on every play it would score a touchdown on every drive. However, the mean passing yards per attempt during the same time period was 5.92 yards per attempt. A team with the capability to put up that number would also score a touchdown on every drive and would score more points that the opponent who only ran the football.

The above goes against over a half century of conventional wisdom. Since the beginning of the sport both teams and fans have believed that the key to winning games involved trotting into the end zone on the ground. As a result, franchises have spent enormous sums of money in order to obtain players with the necessary speed and toughness and capability to run around defenders or straight through them. Passing has only been seen as a last resort which was only necessary in third-down long yardage situations.

One example which demonstrated just how important passing truly is occurred during week 6 of this previous 2008 NFL season when the sub-par Cleveland Browns defeated the defending Super Bowl champions known as the New York Giants. The Giants dominated on the ground by rushing for 181 yards on 25 carries for a 7.2 yards per attempt average while the

Browns averaged a respectable 4.8 yards per attempt. Despite being outperformed on the ground Cleveland still beat the defending champions quite handily with a score of 35-14. Although the aforementioned stats are significant, they proved to be irrelevant due to the fact that the Browns won the air battle. Cleveland averaged 10.7 yards per attempted pass while New York only 6.6 yards per each attempt to throw the ball. That previous statement destroys conventional wisdom.

Offense only counts for half of the action on the football field. A team simply cannot be successful without a stellar defense. After all, defense has always been known as the key to winning championships. Due to the fact that it has always been believed that running the ball was the key to success, it is no wonder that teams have traditionally built defenses focusing on stopping the run.

Between 1990 and 2007 the mean number of rushing yards given up per attempt in the NFL was 4.01 while also giving up 5.93 yards per attempt through the air. Based upon this fact one could conclude that the teams should be more pass-defense-oriented while selecting plays for their roster. General managers should consider defensive linemen with superior pass-rushing skills and focus on obtaining the best corners and safeties as well as linebackers who are capable of stopping the pass.

A Brief Review of the Literature

Alamar (2006) states that teams have more to gain by throwing the ball more. Also, the passing game has become more efficient and less risky over time, but these changes have had no effect in passing frequency. For instance, the completion rate has increased nearly 20% between 1960 and 2005. During this same time period the interception rate dropped 50%. This data stands as evidence that all risk involved with passing the football has dropped significantly over time. Alamar also goes on to state that despite these risk deflations, passing usage has changed very little. In 1960, 52% of all plays were pass attempts and 54% of all 1990 plays were pass attempts. Alamar suggests that the passing game, for obvious reasons, is underutilized in the NFL.

Rockerbie (2008) took Alamar's suggestion one step further. Rockerbie developed a portfolio model which could be used to help a coach determine the optimal number of both passing and running plays. He then compared this portfolio with actual statistics from the 2006 NFL regular season. Rockerbie's results led him to believe that most NFL teams actually pass the ball too often and as a result a "running premium" exists. Also, a tradeoff was found between expected net yardage and the risk associated for both running and passing plays.

Burke (2009) suggests that turnovers are the most important factor in winning a game. He found a correlation of 0.70 between turnovers and wins while offensive passing and rushing has correlations of 0.59 and 0.20. Burke also found run defense to be statistically insignificant.

Alamar and Weinstein-Gould (2008) agree with the statement that passing is statistically significant and broke their analysis down one step further by isolating the effect of individual linemen on the passing game. Their conclusion states that the effect which each linemen has on the passing attack will vary by team due to the fact that quarterbacks have varying degrees of sensitivity to pressure. A team with a quarterback who is highly sensitive to pressure would want to invest in a better-quality offensive line. However, a team who was less pressure-sensitive could spend more money and resources on players such as wide-receivers.

A study by *Cold Hard Football Facts* (2008) concludes that, in general, passing efficiency is still the key factor in winning games and that rushing yardage is particularly important in just one game – the Super Bowl. However, the actual rushing yardage is not what is

considered to be significant. It is the number of attempts which truly matters in the Super Bowl. The study examined data from 42 Super Bowls and discovered that the champions of these games may not have ran more effectively than your average opponent, but averaged a significant number of more attempts. The average Super Bowl champ averaged 37.3 attempts. The losers, on the other hand, only averaged 22.7 attempts. However, this should not overshadow the importance of effective passing. Teams which passed more effectively have won 36 out of the 42 Super Bowls.

This paper examines the relationship among wins and offensive passing and rushing as well as defensive passing and rushing yards given up. Our empirical works are different from past works in several ways. For starters, our data set is unlike any other used in previous literature. No other source has examined data from every NFL team for such an extended time. Our research is unique due to the fact that it focuses on which of the four basic components of a football team has the highest correlation with wins.

The rest of the paper is organized as follows: our next section presents a graphical presentation of our NFL analysis which is followed by our empirical findings on the relationships and a conclusion.

Graphical Presentation

This section will describe all the figures and tables while giving a brief overview of results.

Figure 1 shows the regression of offensive passing yards per attempt against the wins per season. There is a clear and obvious, statistically significant positive result. An increase of one yard per offensive passing attempt will increase the wins in a season by over two, all other factors held constant.

Figures 2, 3, and 4 show the same graphical equations as figure one, regressing defensive passing yards per attempt, offensive rushing yards per attempt, and defensive rushing yards per attempt, respectively.

Figure 5 regresses the offensive pass rate on wins. There is a negative relationship that is statistically significant. However, we believe this is due to previously mentioned reasons and is not a causal statistic; rather, it is an effect of being in and maintaining the lead.

Figure 6 compares offensive passing yards per attempt to offensive pass rate. While there is a positive relationship, it is not statistically significant. We found that passing the ball more or less often has no relationship to the offensive passing yards per attempt.

Figure 7 compares offensive rushing yards per attempt to offensive rushing rate. This finds a positive, statistically significant relationship. The more an offense runs the ball, the more offensive rushing yards per attempt they earn.

Table 1 shows the correlations of our four variables to wins.

Table 2 shows the statistics on our main regression equation. A sample size of 577, an F-value of 287.97 ($P < .0001$), a dependent mean of 7.99, a root MSE of 1.75, a coefficient of variation of 21.92, and an adjusted R^2 of .6659 are all the listed statistics.

Table 3 shows the results of our main regression equation. All variables are statistically significant, and the standard errors and t-values are heteroscedastic consistent values.

Table 4 shows the descriptive statistics of our main regression variables.

Table 5 shows our other regression results: offensive passing yards per attempt *offensive passing rate, offensive rushing yards per attempt*offensive rushing rate, and wins*offensive passing rate. All values are adjusted for heteroscedastic consistent time series data.

Data

We assembled panel data involving all NFL teams from the past 19 years (1990-2008). This totals 577 data entries. We used one data source, Pro Football Reference (http://www.pro-football-reference.com). The stats used are total team stats for the entire regular season only.

Methodology

In our estimation, we attempted to isolate football competence to four statistics – running and passing efficiency on offense and defense. The defensive statistics are yards given up per attempt, so an increase in this statistic would be considered a negative outcome, as giving up five yards per attempt is worse than giving up four. In order to account for other important variables that may possibly influence wins and also consider individual team characteristics, we attempted to use panel data fixed effects. We tried to use time-fixed effects, entity-fixed effects, and a combination of time- and entity-fixed effects for each team and every year. However, including any of these effects or a combination of the two only muddled the results. F-values fell significantly, adjusted R-squared was improved negligibly, and the accuracy of our estimation was not improved. Hence, we decided to use the simple model:

$$W = \beta_0 + \beta_1(PYPAO) + \beta_2(PYPAD) + \beta_3(RYPAO) + \beta_4(RYPAD) + \varepsilon$$

Where W represents wins in a season, PYPAO represents offensive passing yards per attempt, PYPAD represents defensive passing yards per attempt, RYPAO represents offensive rushing yards per attempt, and RYPAD represents defensive rushing yards per attempt.

These statistics are very descriptive and independent. There is no issue of multicollinearity. The variance inflation rate for each of the variables is between 1.0474 and 1.12291. Certainly total numbers of rushing or passing yards are some of the most noticed and analyzed statistics. However, these statistics are not causal to wins; they can be a product of wins. Consider a game in which one team leads. The leading team will attempt to run the ball because running allows the clock to run continuously, bringing them ever closer to the end of the game. Hence, a team that wins consistently will tend to have higher rushing attempts and rushing yards. On the other hand, a losing team will be attempting to pass the ball down the field, trying to catch up. This team will post higher passing attempts and yards. But these statistics are derived from the success (or failure) of each team; they do not necessarily cause that success. However, rushing or passing yards per attempt quantify a team's competence in each category, providing for a descriptive variables that are independent of one another (Burke 2007).

Almost all other statistics can be derived by a team's ability to run and pass (and defend against) the football. First downs are a result of consistent successful runs and passes. Scoring,

whether by touchdowns or field goals, come from a series of successful runs and passes. Attempts, yards, and completion percentage are all considered when determining yards per attempt. Penalties and turnovers are important; however, these statistics are obvious deterrents to winning and are largely influenced by luck.

Therefore, we are left with two positively influencing statistics – offensive passing and rushing yards per attempt – and two negatively influencing statistics – defensive passing and rushing yards per attempt.

Conclusion

Our time series empirical results clearly show that all four measures have statistically significant causal relationships with wins per season. The results we obtained are in contrast to the majority of "conventional wisdom" and in line with much of the scientific previous literature we examined. While conventional wisdom hinges on run offensive and defense, scientific literature, including our study has proclaimed passing efficiency to be much more important to teams' ability to win.

This has significant implications for general managers and coaches. When structuring the team, general managers should focus on what a player can bring to the team in terms of passing offense and passing defense capabilities. An increase in passing production by 1% will result in a 1.63% increase in wins. An equal increase in rushing production, however, will only result in a 0.44% increase in wins (Table 4). Therefore, all players must be evaluated on their contribution towards the passing game. Offensive lineman should be selected based on their pass blocking abilities; running backs that can catch the ball will be more valuable than those who are run specialists; a good stable of receivers is very necessary; receiving tight ends will be more valuable than blocking tight ends. On the defensive side of the ball, pass rushing defensive lineman will contribute more than run stopping lineman; linebackers who contribute many sacks or can cover tight ends and wide receivers are more valuable that run stoppers; safeties must be pass stoppers, not run stoppers; and a good stable of cornerbacks and other defensive backs is a major necessity.

For coaches designing a game plan, determine which plays to call, and how to attack, clearly passing should take on a more significant role. We find that the percent of plays that are passes (pass attempts divided by total plays) has no negative effect on the passing yards per attempt (Figure 6, Table 5). In fact, we found no statistical significance between offensive passing attempt rate and offensive passing yards per attempt. And while there is a negative, statistically significant relationship between offensive pass rate and wins, this is most likely because of the previously discussed reasons, including the fact that running the ball allows for a continuously ticking clock, which is advantageous for teams in the lead looking to proceed quickly to the end of the game. Therefore, we propose that not only should teams look to optimize their passing offense and defense, they should also attempt to pass the ball more often to increase their scoring.

References

"A Victory for Conventional Wisdom." Cold Hard Football Facts. 29 Jan 2009. 18 Feb 2009 <http://www.coldhardfootballfacts.com/Articles/11_2659_A_victory_for_conventional_wisdom.html>.

Alamar, Benjamine. "The Passing Premium Puzzle." Journal of Quantitative Analysis in Sports Vol 2 Issue 4(2006) 18 Feb 2009 <http://www.bepress.com/cgi/viewcontent.cgi?context=jqas&article=1051&date=&mt=MTIzNDk4NDA0MA==&access_ok_form=Continue>.

Alamar, Benjamine and Weinstein-Gould, Jesse. "Isolating the Effect of Individual Linemen in the National Football League." The Journal of Quantitative Analysis in Sports Vol 4 Issue 2(2008) 18 Feb 2009 <http://www.bepress.com/cgi/viewcontent.cgi?context=jqas&article=1113&date=&mt=MTIzNDk4NDUyNQ==&access_ok_form=Continue>.

Burke, Brian. Advanced NFL Stats. 2007-2009. 18 Feb 2009 <http://www.advancednflstats.com/>.

Byrne, Kerry. "Cold Hard Football Facts: Running Dominance Doesn't Lead to Wins." Sports Illustrated. 15 Oct 2008. Sports Illustrated. 18 Feb 2009 <http://sportsillustrated.cnn.com/2008/football/nfl/10/15/chff/>.

Rockerbie, Duane. "The Passing Premium Revisited." The Journal of Quantitative Analysis in Sports Vol 4 Issue 2(2008) 18 Feb 2009 <http://www.bepress.com/cgi/viewcontent.cgi?context=jqas&article=1093&date=&mt=MTIzNDk4NDM3MA==&access_ok_form=Continue>.

Stuart, Chase. "Why do teams run the ball?." Pro-football-reference.com. 12 Feb 2008. 18 Feb 2009 <http://www.pro-football-reference.com/blog/?p=482>.

Figure 1: Wins*PassYPAO

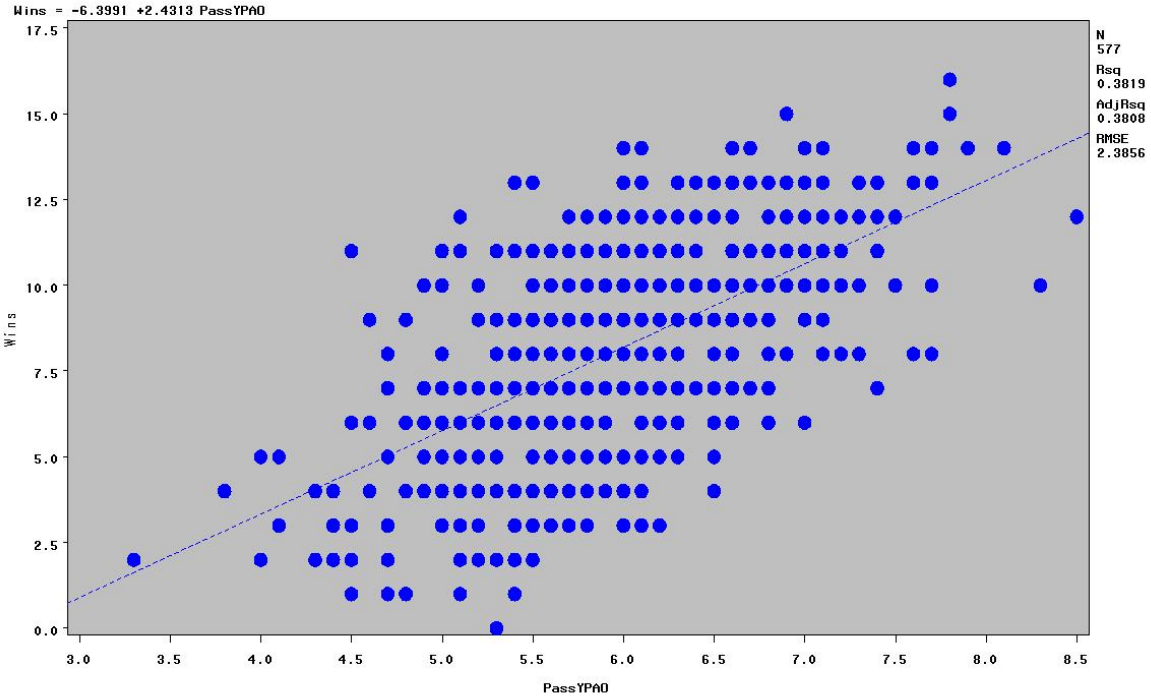

Figure 2: Wins*PassYPAD

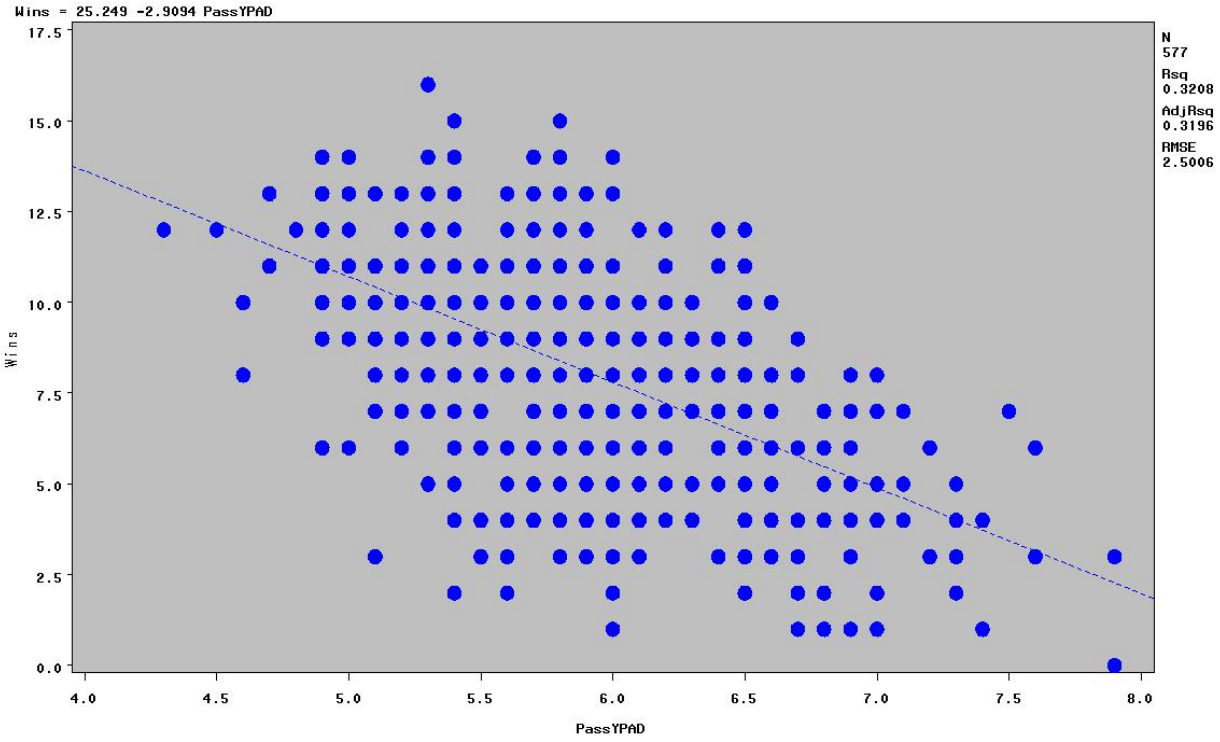

Figure 3: Wins*RushYPAO

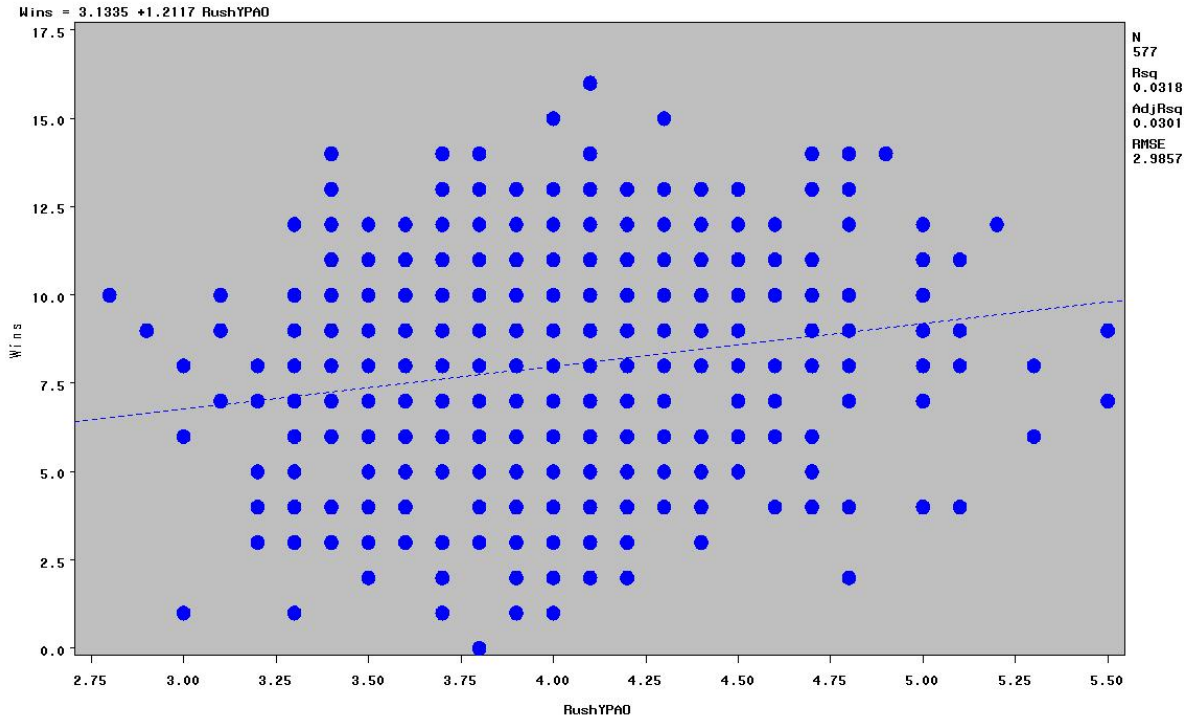

Figure 4: Winds*RushYPAD

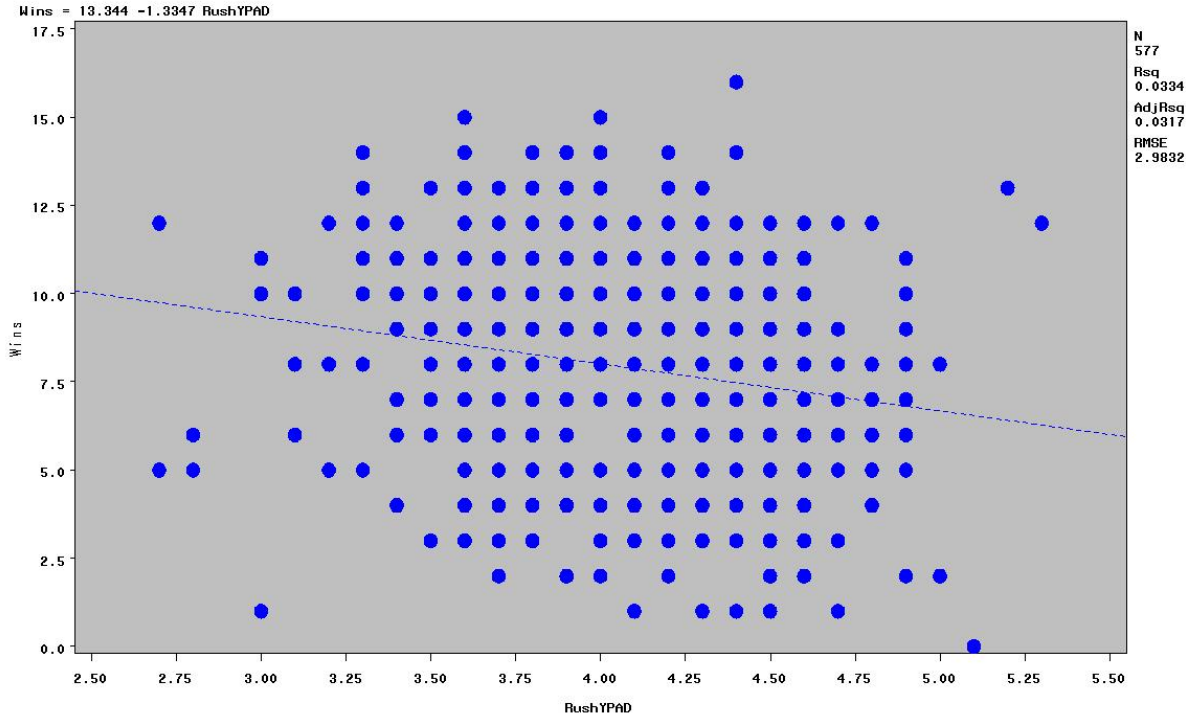

Figure 5: Wins*PassAttRO

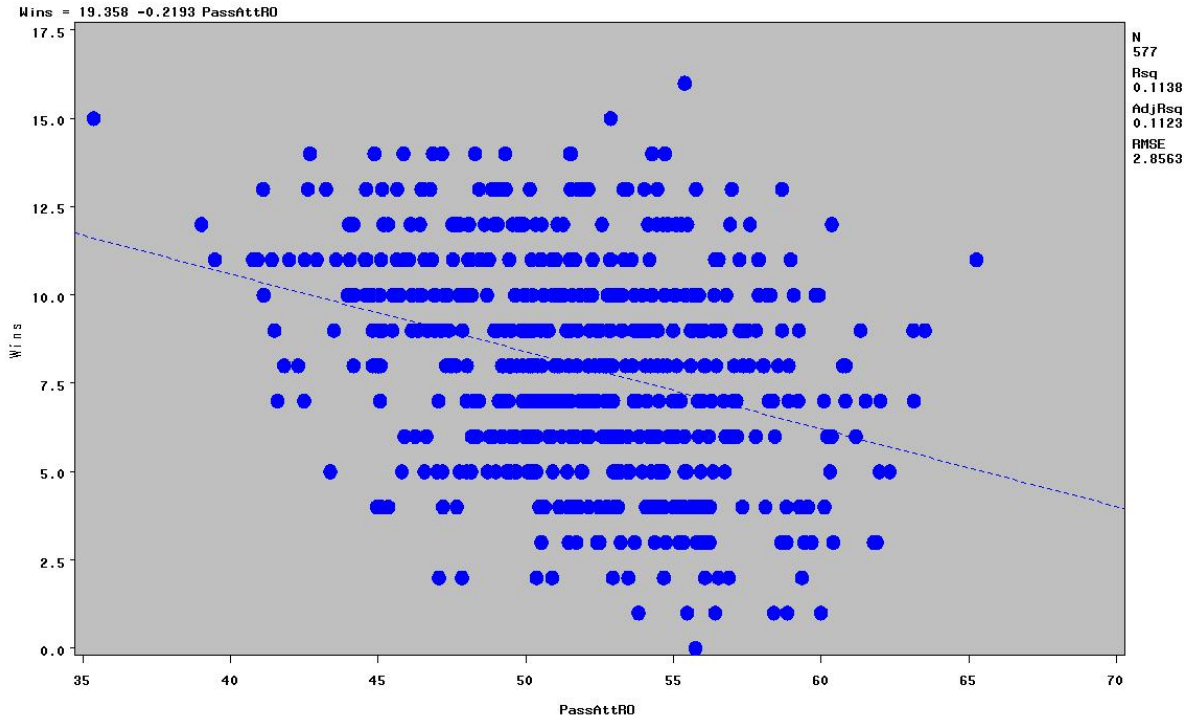

Figure 6: PassYPAO*PassAttRO

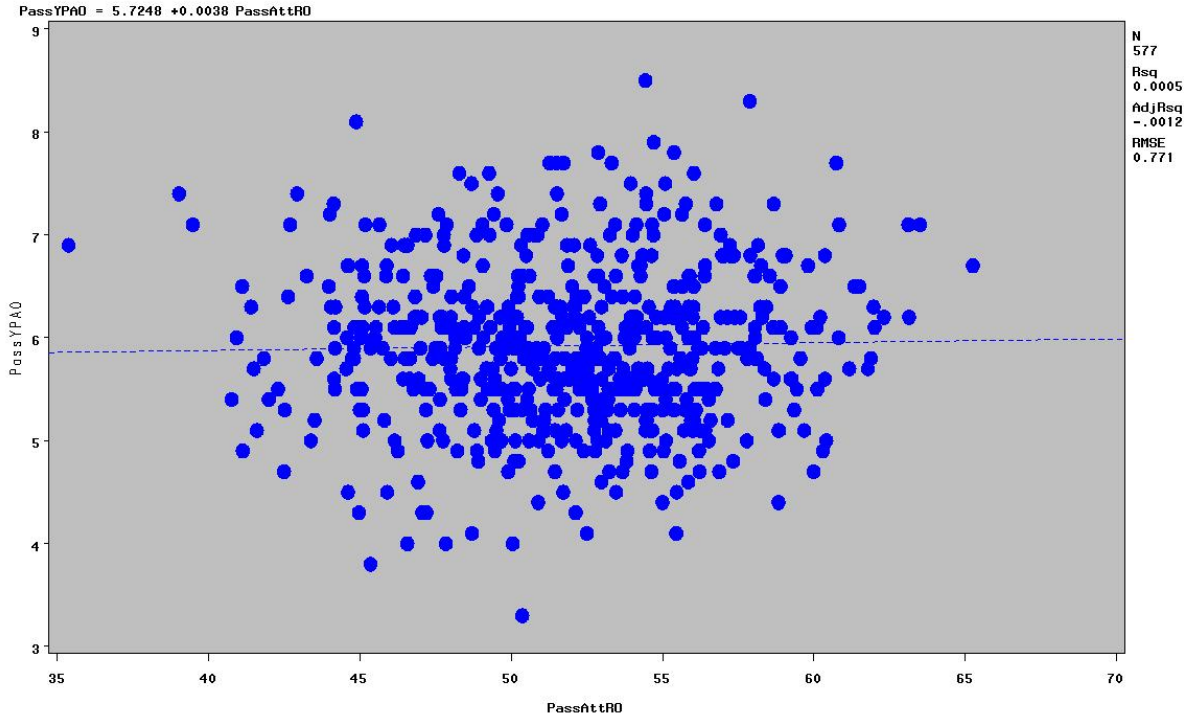

Figure 7: RushYPAO*RushAttRO

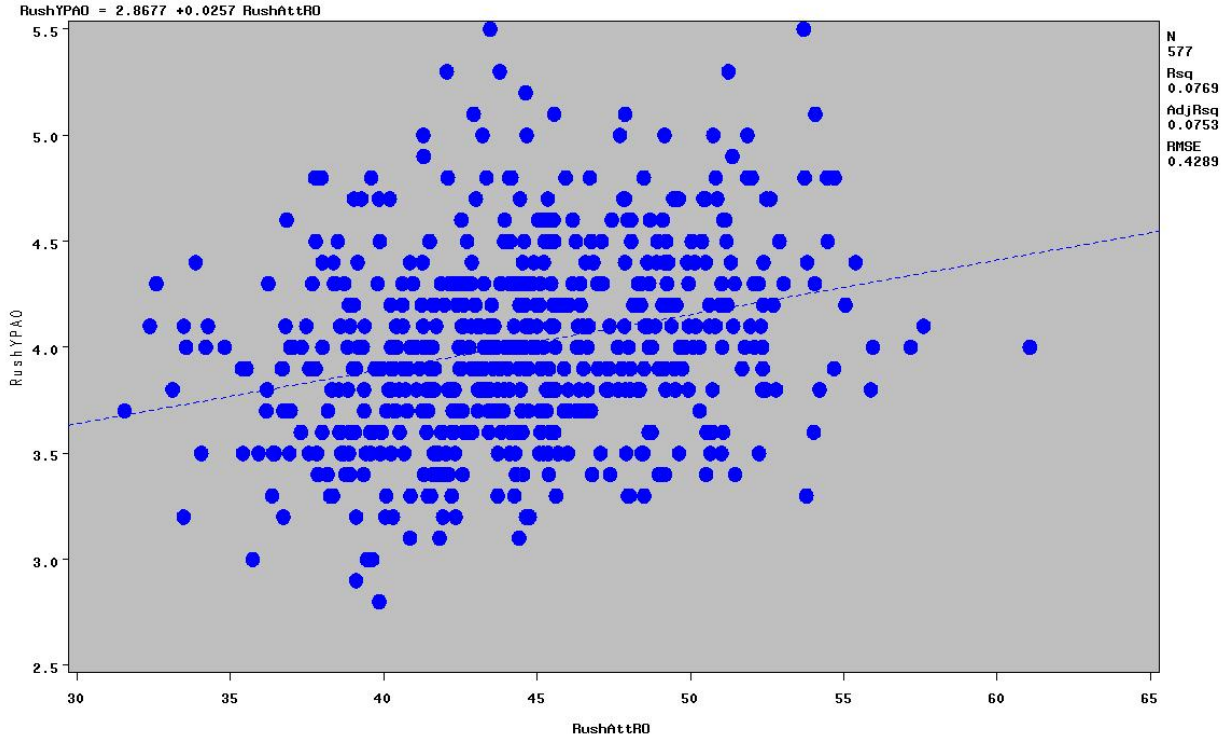

Table 1: Correlation to Wins

	Wins
Wins	1
PassYPAO	0.617956
RushYPAO	0.178265
PassYPAD	-0.56641
RushYPAD	-0.18268

Table 2: Main Regression Statistics

N	577
F Value	287.97
Root MSE	1.75242
Dependent Mean	7.99307
Coeff Var	21.92429
Adj R^2	0.6659

Table 3: Main Regression Results

Variable	Paramater Estimate	HC Standard Error	HC T-Value	P-Value
Intercept (Wins)	9.48703	1.182455201	8.023162307	P < .0001
RushYPAO	0.89277	0.169357346	5.271516254	P < .0001
RushYPAD	-0.77327	0.189962255	-4.070650768	P < .0001
PassYPAO	2.20402	0.095116135	23.17188343	P < .0001
PassYPAD	-2.5326	0.122993674	-20.59130289	P < .0001

Table 4: Descriptive Statistics

Variable	Mean	STD	Max	Min	Elasticity
Wins	7.993068	3.0316537	16	0	1
RushYPAO	4.010572	0.4460193	5.5	2.8	0.447953
RushYPAD	4.009012	0.4149405	5.3	2.7	-0.38784
PassYPAO	5.919411	0.7705321	8.5	3.3	1.632226
PassYPAD	5.931023	0.5902193	7.9	4.3	-1.87924

Table 5: Other Regression Results

Variable	Parameter Estimate	HC Standard Error	HC T-Value	P-Value
Intercept (PassYPAO)	5.72483	0.365611237	15.65824413	P < .0001
PassAttRO	0.00375	0.007011269	0.53485321	P = .5930
Intercept (RushYPAO)	2.86772	0.165645659	17.31237643	P < .0001
RushAttRO	0.02571	0.003713893	6.92265532	P < .0001
Intercept (Wins)	19.35776	1.266833124	15.28043405	P < .0001
PassAttRO	-0.2193	0.024679945	-8.885757133	P < .0001