# Are College Football Coaches Efficiently Compensated?

Senior Honors Thesis

Student: Aaron Grau

Faculty Advisor: John Vrooman

May 1, 2015

In 2014 College football generated over a billion dollars of revenue [1]. As a result coaches who are responsible for both recruiting athletes as well as the performance on the field are well compensated. This paper will address how and why coaches are compensated. It will attempt to determine what makes a coach a superior coach and how to identify these coaches. A coach's ability to coach is most easily judged by casual fans as their ability to win games on the field. However, there is obviously a big difference between a coach who is capable of going 9-4 at Vanderbilt and a coach who goes 9-4 at the University of Michigan. The coach who does this at Vanderbilt will be considered far superior because of the lower funding, inferior facilities, and lesser tradition or program prestige. As a result it is harder to recruit high quality talent. Without high quality talent it is extremely difficult for a coach to compete for conference and national championships. Coaches that have access to teams with higher talent, whether it is because of their own recruiting or simply because of how attractive the school is for recruits, are expected to and should be able to win more games.

This paper will attempt to evaluate the effectiveness of a major conference coach by determining what portion of a coach's recruiting ability is due to the school he is at and what portion is due to him. The paper will then look to evaluate how coaches perform and should be expected to perform based on their team's talent level. Using this information I will be able to quantify the effectiveness of a coach regardless of whether he is at a traditional power or not. Using this information we can look at coach salaries, firings, and new contracts to see how compensation

---

[1] http://www.usatoday.com/sports/college/schools/finances/

decisions are being made. By finding quantitative values for these factors university athletic departments could better evaluate hiring and firing decisions, as well as whether to give current coaches extensions.

## I. Recruiting

### A. Recruiting Model

Every year high school football players that are graduating are recruited across the country by college coaches. Each coach tries to recruit the players he deems most talented and the best fit for his team. The recruitment is a long process, which can often start during freshman or sophomore year of high school for extremely talented players.  The recruiting process involves many visits to different schools and meetings and conversations with various coaches. Because of the large amount of personal interaction between a coach and a recruit it is clear that a coach's ability to make a recruit feel comfortable and sell the benefits of the school the coach is at will play a significant role in a recruit's decision on whether to attend that particular school. In addition to this connection formed with the coaches other factors specific to the school will attract recruits. Schools with superior tradition, football program prestige, facilities, stadium, and recent success will have an edge in recruiting the best talent. For these reasons, as was mentioned in the introduction, we would expect the University of Michigan to consistently have better recruiting classes than Vanderbilt regardless of how good a recruiter the coach is.

The following model is used to determine where we would expect each school to finish in recruiting if there was no human element of a coach, and instead recruits were simply allocated based on the school's football programs themselves.

$$Y = \beta_1 * \omega + \beta_2 * \rho + \beta_3 * \varepsilon + \mu \quad (1)$$

Where Y is defined as a recruiting class value[2] where a recruiting class is defined as the collection of all players recruited to a given school in a given year, where the most talented classes will be rated with the most points. $\omega$ is defined as the average win percentage for the school's football team in the prior five years. The win percentage over the prior five years is used to determine the team's recent success level. This is important because even if a school has been very successful historically if it has been completely incompetent on the field in recent years recruits will see this and be less likely to want to play there as they do not want to be part of a losing team. The same can be said for the reverse; a school that has traditionally been poor but has recently seen on field success will likely see a boost in recruiting.

The next term $\rho$ is a school's program historical prestige ranking based on ESPN.com program prestige rankings[3]. This prestige ranking was selected for the model because it incorporates several factors which are specifically important to recruiting. Most obviously it incorporates on field success for the school, awarding ranking points for winning seasons, bowl wins, and national championships. However, it also includes bonuses for a school producing Heisman winners, All-

---

[2] Recruiting class values are based of Rivals.com annual rankings
[3] ESPN.com 2009 research piece: *Prestige Rankings* By Chris Fallica, Nick Loucks and Harold Shelton

Americans, and first round NFL draft picks. These individual accolades are particularly important to include when considering a program's attractiveness to recruits because players often grow up looking up to a specific super star player. As a result, having a significant pool of superstar talents come from a school will increase attractiveness to potential student-athletes. This term captures the historical success of a school's football program's success in attracting recruits.

The final term $\varepsilon$ is the school's stadium capacity[4]. This was chosen for the model for two reasons. The first and the most direct is that players like to play in front of large crowds and are attracted to schools with larger stadiums. The second reason for its inclusion is that it acts as a proxy for a school's investment in facilities. Especially in the past decade schools have been investing hundreds of millions of dollars in practice facilities, locker rooms, stadium improvements, and lounges for student athletes with the intention of attracting superior talent. In general schools that have invested in building extremely large stadiums have also invested heavily in facilities. This is true of many major programs including Alabama, Michigan, LSU, Ohio State, and many others. The end result of these superior facilities is better recruiting classes.

One notable exemption from the model is any kind of academic ranking of the school because there is no significant relationship between better academics leading to better recruiting classes. This makes sense intuitively because many of the top high school football players are more concerned with making the NFL and improving their football ability than any academics the school they attend offers.

---

[4] Stadium Capacities are taken from Wikipedia.com list of NCAA stadium capacities http://en.wikipedia.org/wiki/List_of_American_football_stadiums_by_capacity

When the model is regressed using an OLS regression, with the recruiting class value, prestige rank, 5 year win %, and stadium capacity, using data from 2009 through 2012 for major conference teams the following values are solved for the beta coefficients and intercept:

*Table 1.*

| R^2 = 0.577 | *Coefficients* | *Standard Error* | *t Stat* |
|---|---|---|---|
| Intercept | -351.79439 | 226.28099 | -1.55468 |
| $\beta_1$ (Prestige Ranking) | -4.00092 | 1.51287 | -2.64458 |
| $\beta_2$ (5 year trailing win %) | 707.44428 | 215.62274 | 3.28094 |
| $\beta_3$ (Stadium capacity) | 0.01891 | 0.0021 | 8.9957 |

As can be seen from the t-statistic values all the coefficients are highly statistically significant. Additionally all coefficients have the expected sign. We would expect that teams with larger stadium capacities and higher five-year win percentages would have higher recruiting class values. And teams with higher (worse as it is a ranking from 1-130) prestige ranks would have lower recruiting class values. The regressions gives an R^2 value of 0.577. This seems fairly good, as we would expect a good amount of unexplained variance because the model does not contain the coach's recruiting ability.

In order to determine the coach's ability to recruit we can rearrange equation

to solve for the error term ($\mu_R$), which can be defined as the specific coach's
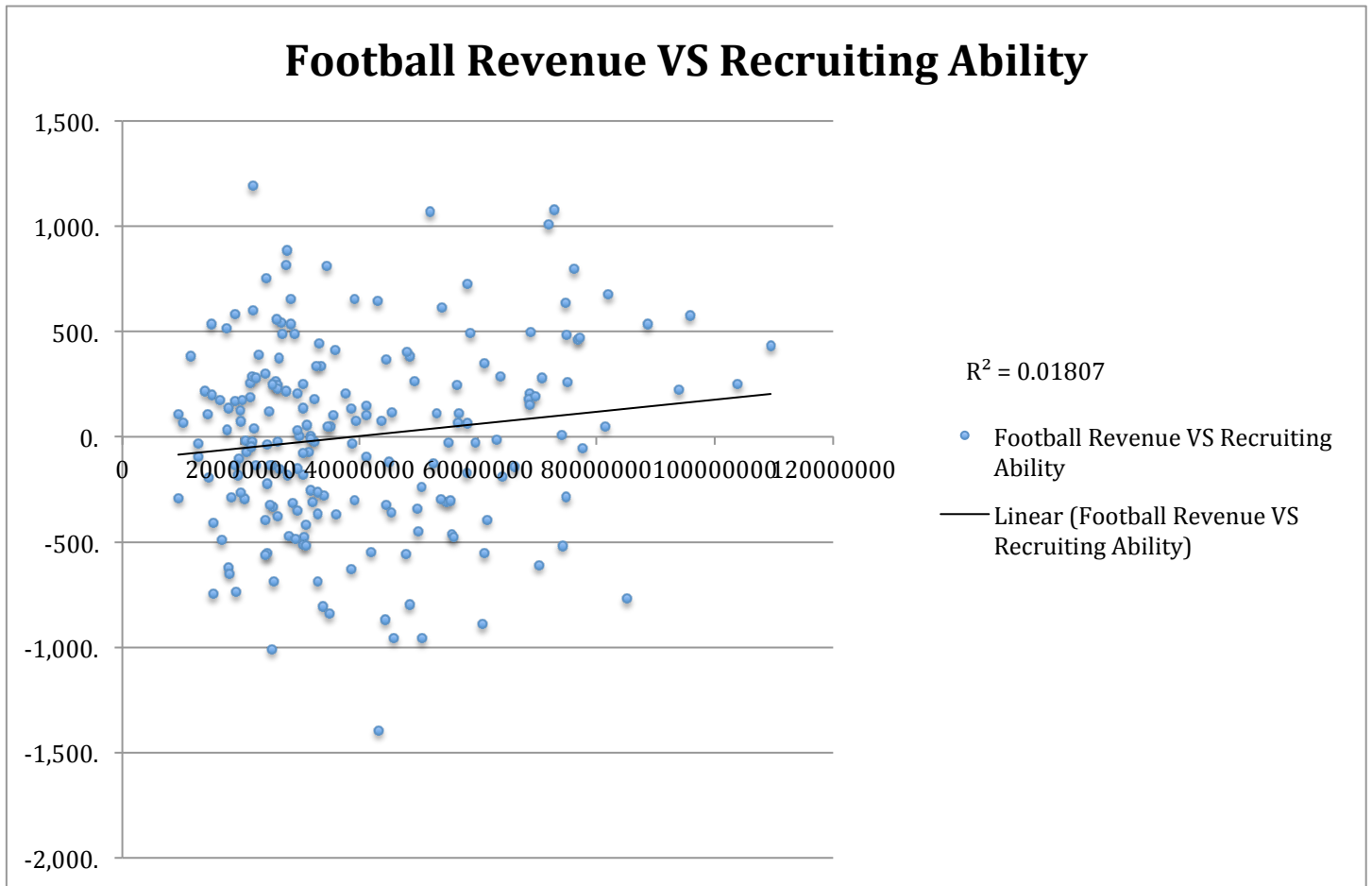
recruiting ability:

$$\mu_R = \beta_1 * \omega + \beta_2 * \rho + \beta_3 * \varepsilon - Y \quad (2)$$

This means that when regressed the residual of each observation represents

the coach's ability to recruit. This assumes that all differences in recruiting class

rankings not explained by recent success, program prestige, and stadium size are

due to the coach and his staff's ability to recruit. While this is not always the case

due to NCAA sanctions and other factors not accounted for in the model, such as the

rare case of a school like Oregon, which has a small stadium but top end facilities, in

general I deem this as an acceptable assumption for the model.


*B. Allocation*

If coaches were optimally allocated we would expect that coaches at the

schools with the higher revenues would be more likely to be better recruiters than

coaches at smaller programs. By looking at the following graph of the residuals

versus the school's revenue the coach is at, we can see if coaches who are good

recruiters tend to end up at schools with higher revenues. We would expect for this

to be a naturally occurring process as there is high turnover -- approximately 25%

of jobs are turned over in division I football each year and many of these are a result

of a double hit as top programs fire a coach and steal away a coach from a lesser

division I program. Additionally we would expect that better coaches would

increase the revenue of their schools and perhaps see some reflection of that.

*Figure 1.*



**Football Revenue VS Recruiting Ability**

$R^2 = 0.01807$

• Football Revenue VS Recruiting Ability

—— Linear (Football Revenue VS Recruiting Ability)

As can be seen in the graph there is not a significant pattern besides for the very high revenue schools tending to have superior recruiters. There are two explanations for this. The first is a function of the data and could be an econometric issue in the sense that we would expect schools with higher revenue to have larger stadiums and better traditions (bigger fan bases). As a result school revenue will be correlated with dependent variables of the model and may already be accounted for when considering the coach's ability to recruit. The other explanation could be that massive powerhouse programs with huge revenues tend to do well in recruiting regardless of the coach. These programs instead may need coaches who excel at on

the field coaching and developing talent. Later in the paper I will explore this hypothesis and see if it is the case that "game day" coaches gravitate towards these elite schools.

*C. Coach Salary and Recruiting Ability*

Because coaches are compensated on the basis of their performance we would expect coaches that are superior recruiters according to the model to be compensated at a higher rate than other coaches. By graphing coach salaries versus their recruiting performance in the given year we can see if this is the case[56]. We would also expect the most highly paid coaches to recruit above average.

*Figure 2.*
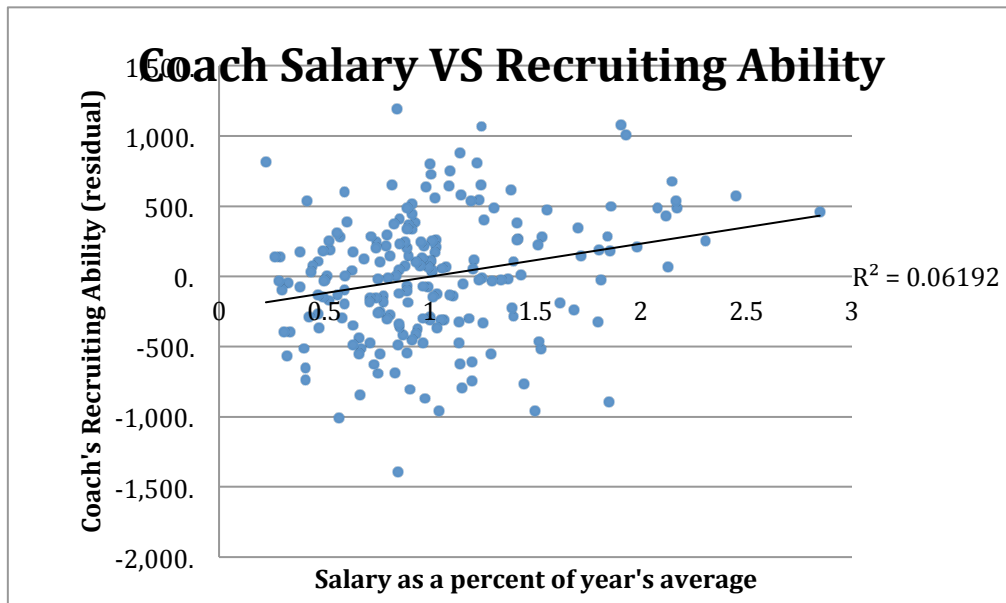
Both hypotheses are confirmed by this graph. The upward slope of the plot shows that as coaches are paid more they tend to recruit better. Additionally, we see

[5] Coach Salaries Taken from USAToday.com Coach salary database: http://sports.usatoday.com/ncaa/salaries/
[6] Almost all private schools do not make salary data public and are excluded as a result

that all coaches earning over two times the average salary in the given year recruit at a better than average level.

This model shows a coach's ability to recruit or their performance in each year. In the next section I will evaluate on field coaching performances. By combining both models I will be able to evaluate a coach's overall coaching ability and determine if the allocation and payment of coaches resembles what would be considered efficient.

## II. On field Coaching

*A. On Field Coaching Ability Model*

While the talent of the players on the team is obviously very important to a team's winning percentage the ability of a coach to maximize his talent through his scheme and talent development ability is arguably even more important. This section of the paper will attempt to quantify a coach's ability to generate wins with a given talent level.

In any year a coach will have access to some portion of each of the 5 prior recruiting classes at the given school (redshirts allow for 5th year seniors). As a result the total talent level of a given team in any given year will be some function of the recruiting class value for the five prior seasons.  By finding the expected conference win percentage of a team with a given talent level, I can calculate the effect the coach had on the given season.

The reason I chose conference win percentage instead of overall win percentage for the season is because I wanted to have similar strengths of schedule

for all teams. Because many power 5 conference teams schedule very weak out of conference opponents the win percentage could be higher for an inferior team simply due to the fact that they played a few much easier games. While this problem will also exist between different conferences to some extent, better power 5 conferences will require higher levels of talent to reach the same result as in a poorer conference; it should be much less drastic than the issue of teams playing FSC teams and other cupcake opponents.

My initial model to solve for the expected wins of a team with a given talent level was:

$$\gamma_t = \beta_1 * Y_{t-1} + \beta_2 * Y_{t-2} + \beta_3 * Y_{t-3} + \beta_4 * Y_{t-4} + \beta_5 * Y_{t-5} \ (3)$$

Where $\gamma_t$ is a team's conference win percentage in year t. $Y_{t-1}$ is a teams recruiting class value for year t-1. Each subsequent term represents another year prior's recruiting class.

When this model was regressed using 2009-2012 seasons it gave R^2 value of .1 indicating there was some level of explanation. However, there were highly insignificant p-values (p>.15) for all the coefficients. While at first this result was somewhat surprising as I expected a fairly significant regression in which the recruiting classes from 3-4 years prior would be valued the most due to the fact that those players would be expected to be developed as well still in school and as a result have the biggest impact on a team's win percentage (many fifth year high talent players, while fully developed may have graduated or moved on to the NFL). However this was not the case. In order to ensure that colinearity issues were not causing the lack of significance, as there is good reason to believe that prior year

recruiting class values would be highly correlated with each other, the recruiting class values were regressed against each other ($Y_{t-1}$ against $Y_{t-2} ... Y_{t-5}$) . While there was high a r^2 for these regressions it was only around .65 on average. This is not considered high enough to be a problem of multi-colinearity.  As a result the lack of significant results from the model must be attributed to a flaw in the model.

I concluded that the model was flawed because it assumed that all teams would benefit from each class equally. That is to say it assumes the t-3 recruiting class value for all schools should have an equal impact on the team. However, in practice this is probably not the case. For example a team with a very highly valued freshman class but a very low rated junior class will rely more on the freshman class and play more freshman than a school with an equally rated junior and freshman class. As a result the freshman class in this situation would likely add more wins than at a school with similar freshman and junior classes. In many ways this is the responsibility of a coach, to place the best players on the field in the best position to succeed and win the most games. As a result of this I switched to the following model:

$$\gamma_t = \beta_1 * \log\left(\frac{(Y_{t-1}+Y_{t-2}+ Y_{t-3}+ Y_{t-4}+ Y_{t-5})}{5}\right) + \mu_f \quad (4)$$

In this model the team's conference win percentage in a given year is instead simply described as function of a team's average recruiting class over the past 5 years. This was chosen to allow flexibility to what age of players a coach decides to use as was previously described as opposed to equation 3 which assumes all teams will place

equal importance on specific recruiting classes. When equation 4 was regressed in

an OLS regression using the 2009 through 2012 seasons the following calculated:

*Table 2.*

| R^2 = 0.155 | Coefficients | Standard Error | t Stat |
|---|---|---|---|
| **Intercept** | 0.30614 | 0.03795 | 8.06635 |
| **(trailing recruiting class 5 years)** $\beta_1$ | 0.00017 | 0.00003 | 6.05915 |

This gives a t-stat of 6.05, which shows high levels of significance. However,

the R^2 value is only 0.155, which we would not expect it to be extremely high due

to the coach being left out of the model intentionally. A possible improvement to the

model to increase the R^2 value would be:

$$\gamma_t = \beta_1 * \log\left(\frac{(Y_{t-1}+Y_{t-2}+Y_{t-3}+Y_{t-4}+Y_{t-5})}{5}\right) + \beta_2 * \alpha \quad (5)$$

Where $\alpha$ represents the average age, or playing experience, of the roster in a given

year. I do not currently have this data so I was not able to regress it. For now I will

accept the limitations of my model and preform analysis with it in a similar fashion

as was done for recruiting. By solving for the residuals once again:

$$\mu_f = \beta_1 * \log\left(\frac{(Y_{t-1}+Y_{t-2}+Y_{t-3}+Y_{t-4}+Y_{t-5})}{5}\right) - \gamma_t \quad (6)$$

We solve for the unexplained portion of the win loss record which for the purpose of

this paper I will assume that this is all due to the effect of the coach, defined as $\mu_f$.

In reality there are probably other unexplained factors, such as age, which has

already been addressed.

*B. Coach Salary and On Field Coaching Ability*

As was discussed in the recruiting section we would expect the best on field coaches to end up at the highest paying jobs through turnover in the market. We would expect that the more the coach is paid, the more likely he would be to coach at an above average level for his given talent level. Figure 3 displays this relationship.

*Figure 3.*
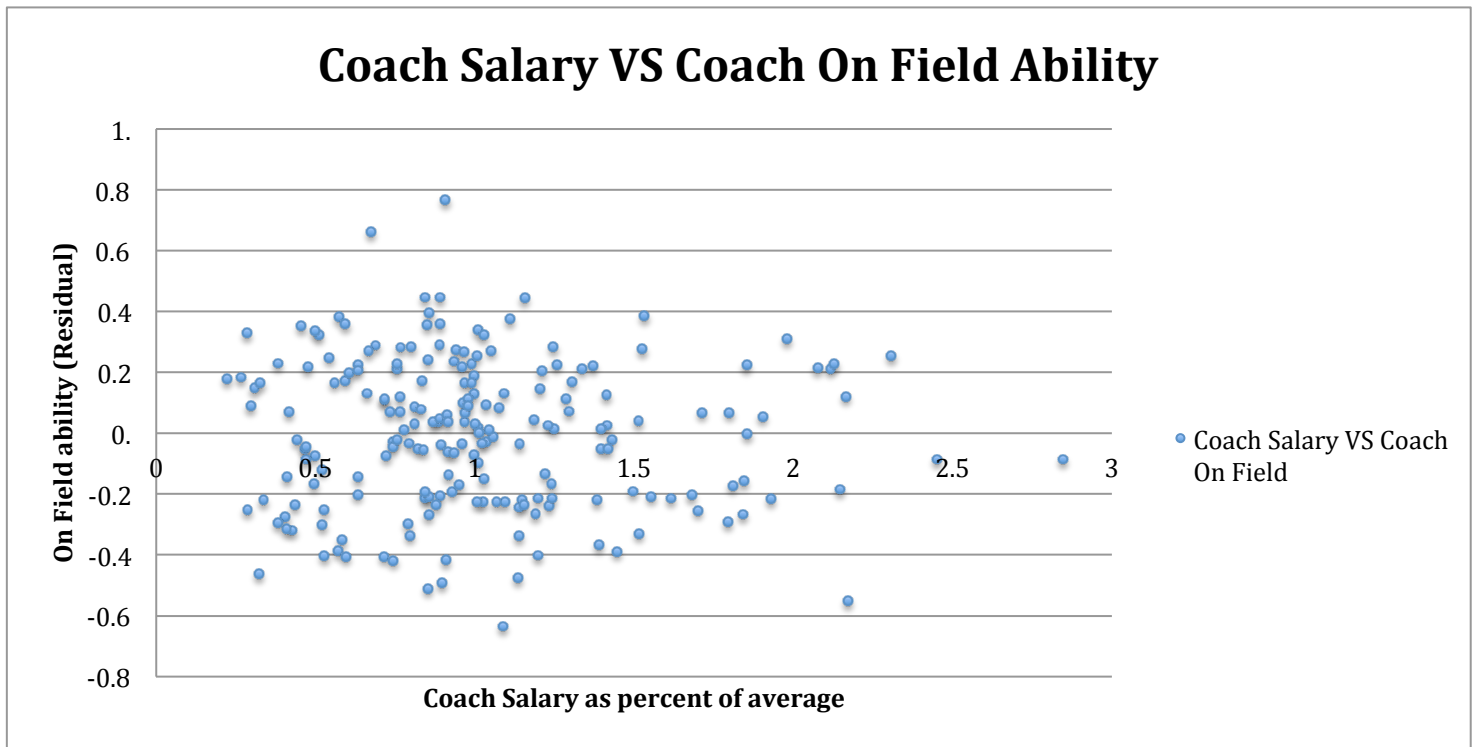


In contrast to Figure 2, which is for coach salary VS recruiting ability, we do not see a trend where higher paid coaches are better on the field coaches. This could simply be a function of the model not describing a team's winning percentage as well as it should so there is too much noise in the residuals. It could also be the case that coaches at the top schools recruit very well so it is hard for them to exceed the expectations of the model given their high talent levels since the model already

assumes they should have very high win percentages. In order to confirm which of these hypotheses is correct the model must be improved, as described above.

Using this model along with the recruiting model we can formulate an expected win percentage for each coach if they were all coaching at the same school.

### III. Overall Coaching Ability

*A. Combining the Models*

In the long run a coach's win loss record will be determined by his ability to recruit and his ability to coach with the given talent level. This is because in the long run all talent a coach has will be based on his own recruiting. As a result a coach's long run win percentage at any given school can be calculated as follows by combining equations 1 and equation 4 which represent the recruiting and on field models:

$$\bar{Y} = \beta_1 * \omega + \beta_2 * \rho + \beta_3 * \varepsilon + \mu_R \quad (1)$$

$\bar{Y}$ Is the average talent value of the coach's classes at any given school.

We can substitute this into equation 4 to replace the average talent term as shown:

$$\gamma_t = \beta_1 * \log\left(\frac{(Y_{t-1} + Y_{t-2} + Y_{t-3} + Y_{t-4} + Y_{t-5})}{5}\right) + \mu_f \quad (4)$$

$$\bar{Y} = \log\left(\frac{(Y_{t-1} + Y_{t-2} + Y_{t-3} + Y_{t-4} + Y_{t-5})}{5}\right) \quad (7)$$

$$\gamma_t = \alpha + \beta_1 * \bar{Y} + \mu_f \quad (8)$$

Using the intercept and beta value solved for in the prior section of the paper:

$$\varpi_t = 0.30614 + 0.00017 * \bar{Y} + \mu_f \quad (9)$$

Equation 9 allows us to define a coach's long-term win percentage at any given school if we have his on field coaching ability and recruiting ability $(\mu_f, \mu_R)$.

*B. Calculated Coach Win Rates*

In this section I display the predicted long term win rates of each coach observation (treating each year as an independent performance even if a coach repeats) if they were placed in a school where the expected recruiting class value is 1,223 but still played against power 5 conference schedule. This was chosen as it is the average for major conference teams and it is large enough to avoid the issue of having any coaches recruiting negative value classes, which in practice is impossible. A long run win rate for each coaching performance from 2009 through 2012 was calculated. A few values were below 0 and a few were above 1 these were changed to 0 and 1 respectively. The average of all the calculated expected long-term conference win rates came out to .512. This coming out to near .500 is good as we would expect that at an average school (expected recruiting class is 1,223) the average of all coaches' long term performances should be close .500 if the model is functioning properly and unbiased. A histogram representation is shown below of the distribution of expected win percentages:

*Histogram of expected long term results from coaching performances 2009-2012:*

*Figure 4.*
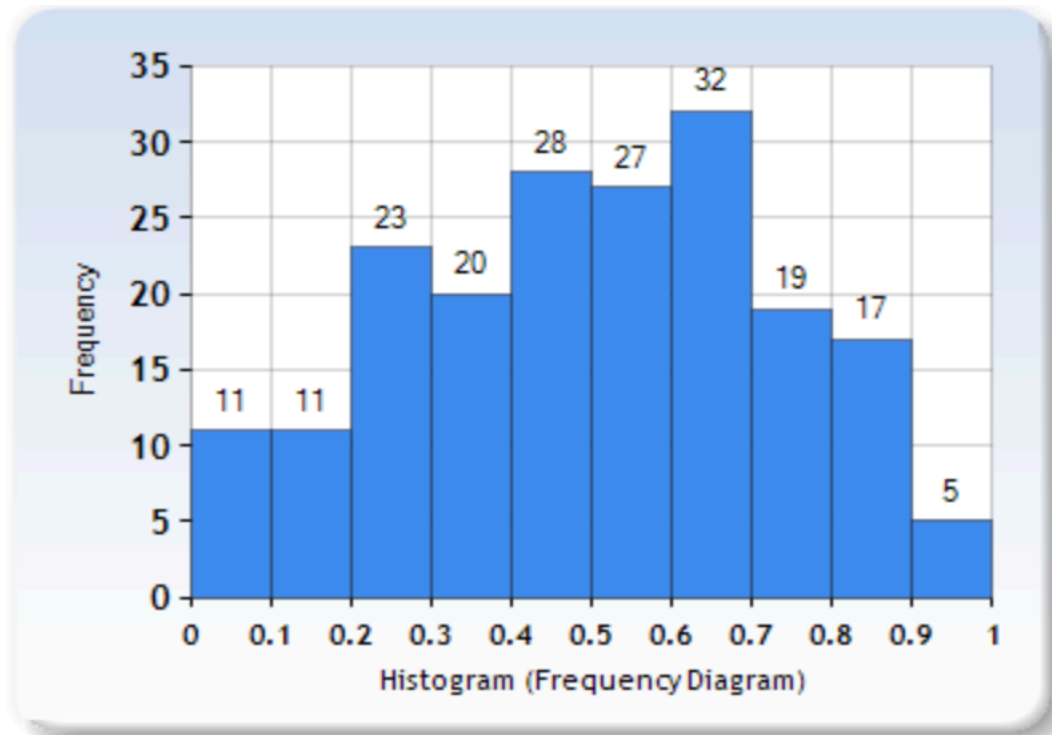


As we would expect to be the case the model produces a distribution that resembles a normal distribution. The majority of coaches are within 20 points of win percentage of 500. The coaches who have had seasons with long term expected win values of 100% during these years are: Nick Saban (Alabama, 2009), Chip Kelly (Oregon, 2010, 2011).

The model has identified two of the best coaches in college football as its favorites. Nick Saban is able to consistently recruit at an above average level and coach at an above average level and this is seen by the model and his undefeated 2009 performance while recruiting the number 1 class is given 100% long term win percentage despite the natural advantage of recruiting at Alabama. Chip Kelly, whose obvious talent has been recognized and has since moved on to the NFL, was

the only coach given 2 long term perfect coaching performances both in 2010 and 2011, where he went 12-1 and 12-2 respectively. He also recruited the #13 class in 2010 and #9 in 2011. While Kelly is a good recruiter his ability is likely overstated by the model because Oregon has great facilities but a smaller stadium, something the recruiting model fails to account for. However, the model has done a great job of identifying two of the best coaches as owners of the best coaching seasons.

C.  *Calculated Coaching win Rates and Compensations*

As I have done in each of the prior sections I will see if it is the case that the coaches with the best expected long term win rates are also the most highly compensated. The following is a graph of this relationship.

*Figure 5*
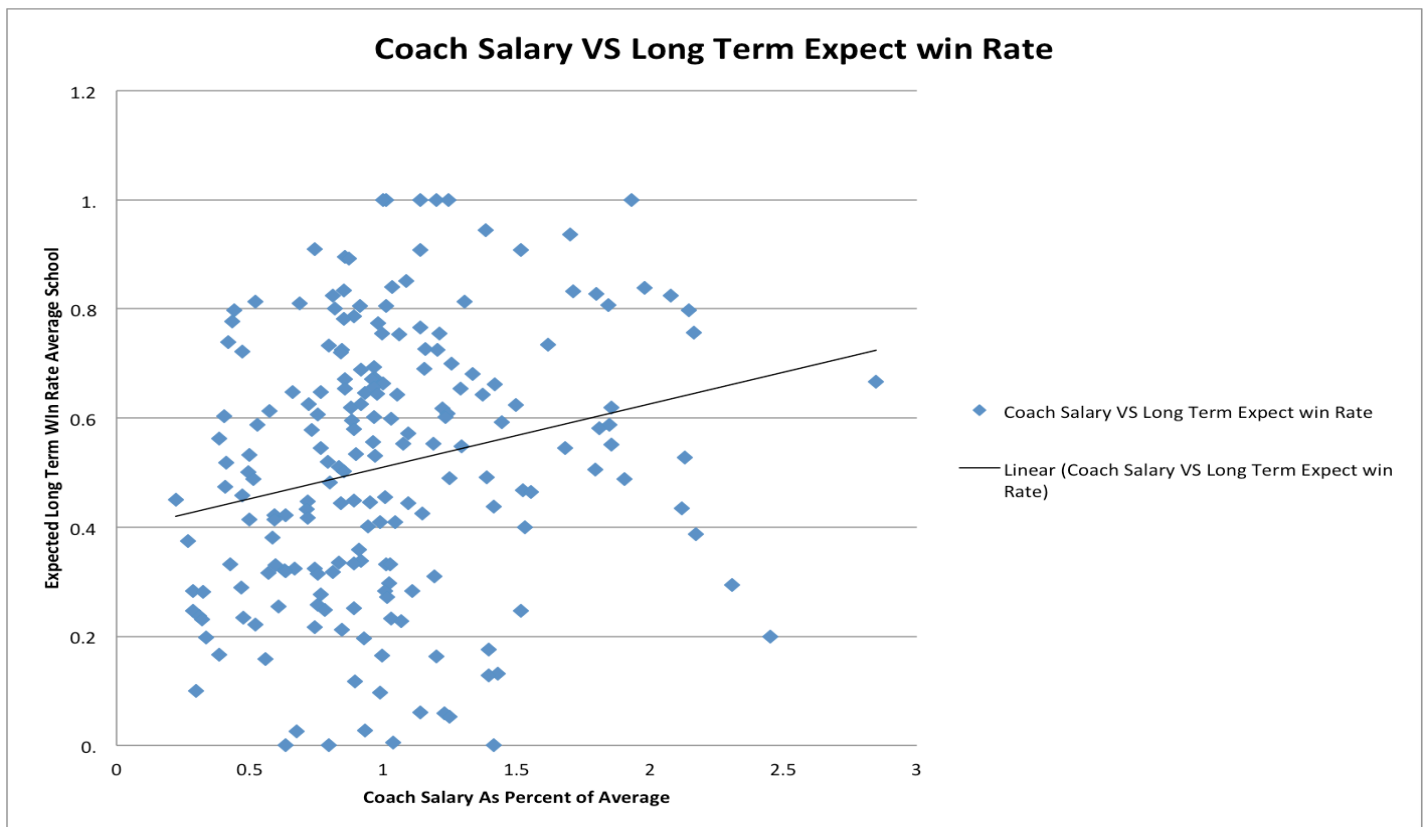
As hypothesized there is an increasing trend in performance as salaries increase. While there is noise and there are good and bad coaching performances at all levels of salary it is clear that coaches who are being paid exceptionally well and have salaries over 150% of the average tend to perform better than .500. On the other hand coaches making below 50% of average tend to perform at a long run win rate of less than .500.

## IV. Firings and Contract Extensions

At the end of every season athletic directors evaluate their coaches performance and decide whether to give a contract extension, fire the coach, or keep the current contract. In this section I will address the likelihood a coach gets a new (more lucrative) contract or gets fired based on their performance.

*A. Likelihood to be fired*

Using a binary logistic model I will determine the likelihood of a coach being fired based on the calculated long-term win percentage of a given season. The equation to be regressed is:

$$Fired = \frac{1}{1 + e^{-(\beta_0 + \beta_1 * \varpi_t)}} \quad (12)$$

Where Fired is simply a binary variable with 1 representing a coach getting fired and 0 not fired for each season coached. $\varpi_t$ is the calculated long term winning percentage of a coach for the corresponding season. Regressing this equation with the data from 2009-2012 gives the following significant result:

*Table 3.*

| Variable | Beta | Standard Error | p-level |
|---|---|---|---|
| Long term win rate (#1) $\beta_1$ | -3.965 | 1.02285 | 0.001 |
| Intercept $\beta_0$ | -0.192 | | |

Figure 6 graphs the probability level for a coach to be fired for each single year observations.

*Figure 6.*
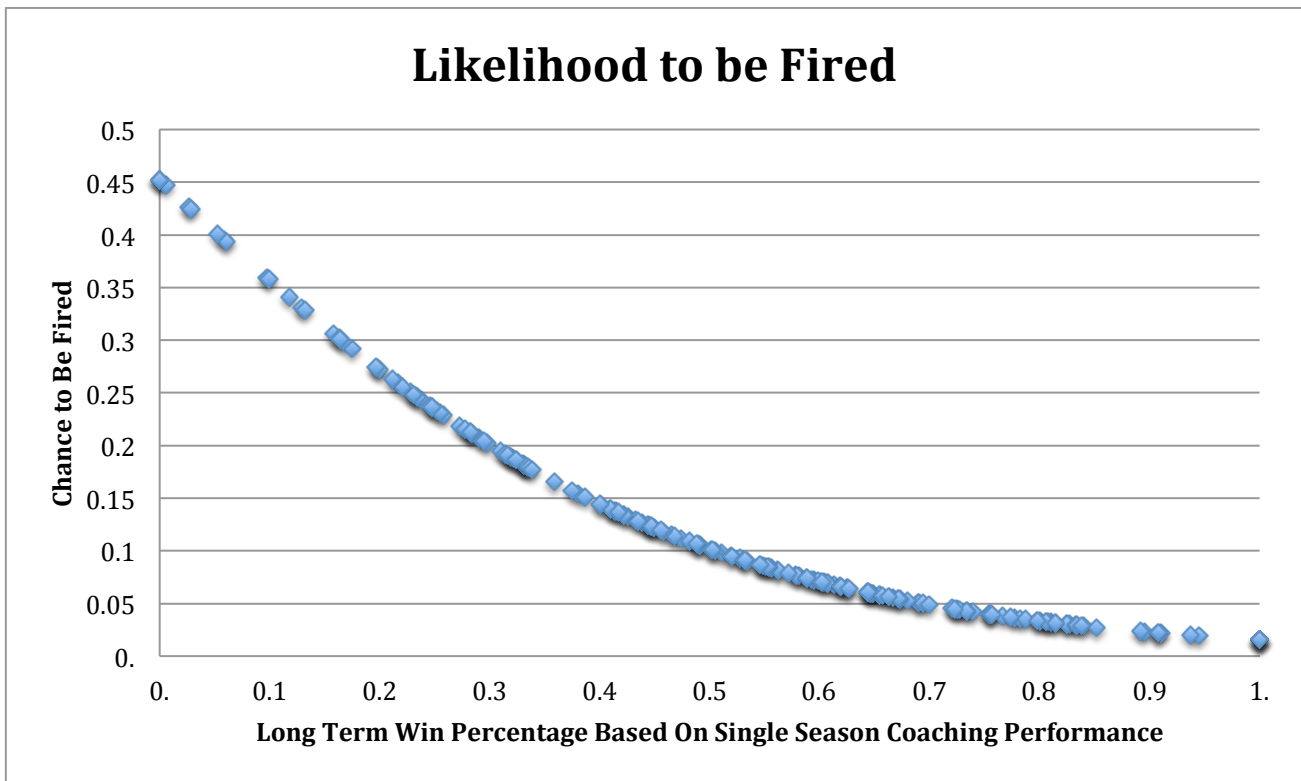


As expected the chance to be fired goes up significantly at the lower end of the win percentages. Coaches with expected long term win rate above .500 have a very low chance to be fired which is logical. Below .500 the chance to be fired begins to increase fairly rapidly and eventually reaches 45% for coaching performances with a long-term value of 0 wins.

In order to see if there is a significant difference in the probability of being fired based on my model and the simple win percentage model, the simple win percentage values were also used in the same regressions.

While these results are significant and logical it is reasonable to assume that firing decisions are not made based on isolated seasons. One would expect that coaches who perform poorly two years in a row would have a much greater chance of being fired than those who have had only one bad season amongst other good seasons. To test this, I regressed the same equation (12), but this time used only data points where the coach had a calculated long term win percentage below .500 in the previous season the following results.

*Table 4.*

| Variables | Beta | Standard Error | p-level |
|---|---|---|---|
| Long term win rate (#1) $(\beta_1)$ | -4.03698 | 1.55705 | 0.00952 |
| Intercept $\beta_0$ | 1.1268 | | |

Graphing the probability results of this regression on top of the prior regression, which used all data points, allows for comparison of the likelihood of being fired

*Figure 7.*
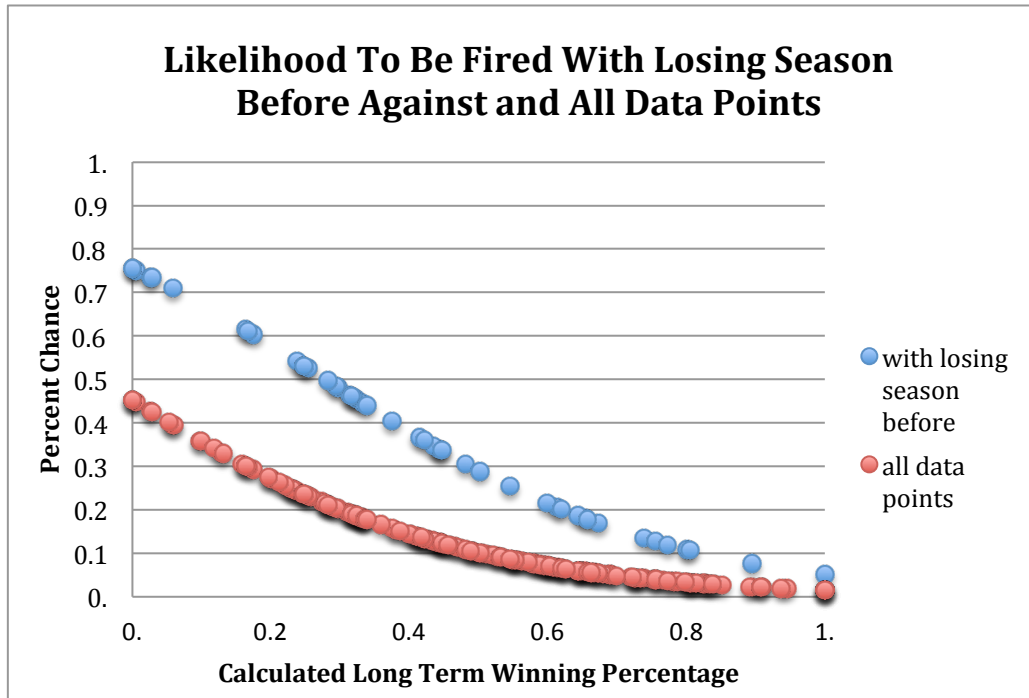


As expected the likelihood of being fired is far greater in years following a losing season. There are several factors that logically suggest that this is the case. One reason is the fact that the Athletic Director who hired the coach has an incentive to not admit his mistake and stick with the coach he hired. Another reason is that discontinuity in the program is generally bad and can damage the program long term. Finally the most direct and obvious reason is that the contracts contain large buyouts that the school must pay to terminate the coach.

*B. Contract Extensions*

This section will use the same method as for likelihood to be fired. The only change is that the binary value will be replaced with a binary representing the coach getting a new contract rather than being fired.

$$New\ Contract = \frac{1}{1 + e^{-(\beta_0 + \beta_1 * \varpi_t)}}\ (13)$$

Where new contract is simply a binary value, 1 for a new contract and 0 for no new

contract after a given season, and $\varpi_t$ is long term win percentage calculated for a

given coach in a given year. Regressing this equation gives the following results.

*Table 5.*

| Variable | Beta | Standard Error | p-level |
|---|---|---|---|
| Long term win rate (#1) $\beta_1$ | 3.91695 | 0.94103 | 0.00003 |
| Intercept $\beta_0$ | -3.77162 | | |

Figure 8 below is a graph of the probability level of a raise for each observation.

*Figure 8.*



**Likelihood to get a raise**
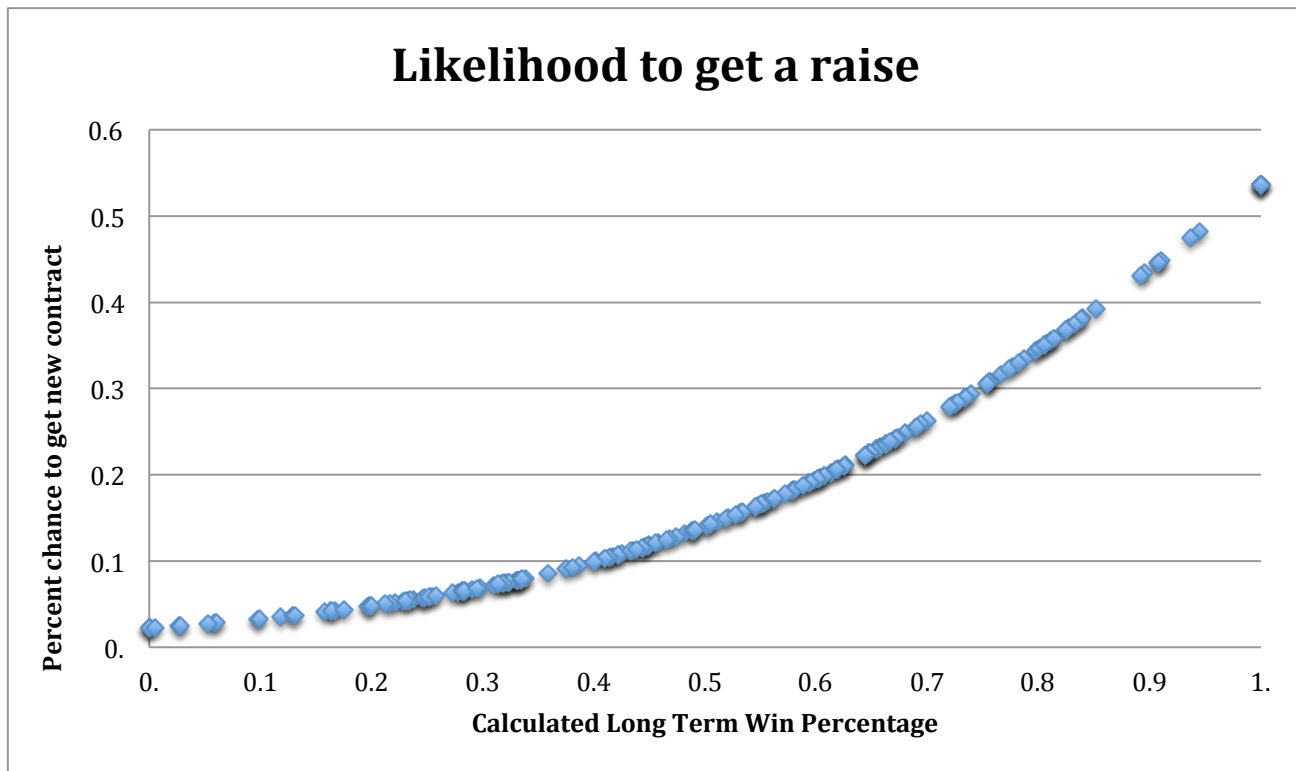
As expected, in seasons where coaches score higher they are far more likely
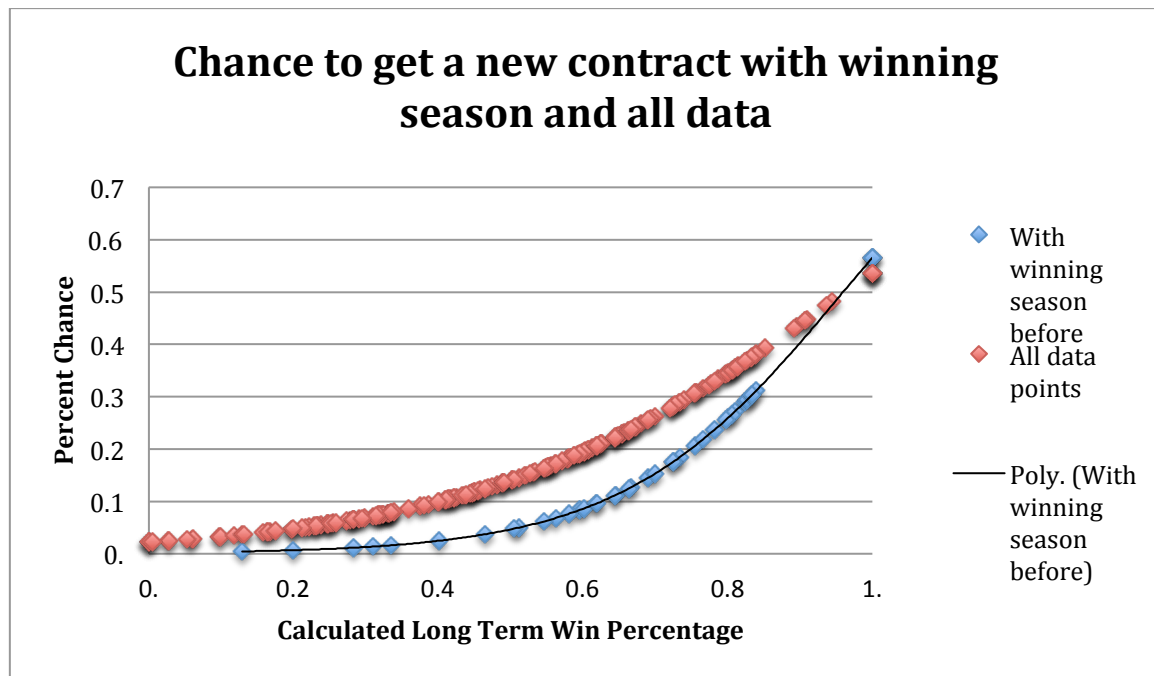
to receive a new contract.

Once again it seems logical that coaches who perform well two years in a row are more likely to get raises. To test this, I used the same method for coaches being fired, but data points were limited to where the coaching performance was over .500 in the prior season. Regressing this gives the following significant result.

*Table 6.*

| Variable | Beta | Standard Error | p-level |
|---|---|---|---|
| Long term win rate (#1) $\beta_1$ | 6.56795 | 3.26145 | 0.04403 |
| Intercept $\beta_0$ | -6.30629 | | |

Graphing these probability values along side ones for all data points produces the following graph

*Figure 9.*



Interestingly, the effect is not the same as with the firings. Having a winning season the prior year actually slightly lowers the chance of recieving a new contract the

next season. This suggests that many coaches tend to get extensions after one quality season, but having another one the following year will not get them a second new contract. This is likely because coaches are very good at floating their names for other jobs and making the school feel that it is necessary to increase their salary. Additionally, it is easier for an athletic director to give a new contract than fire a coach, since there is no need to admit mistake or pay a buyout.

Figure 10 shows the chance of being fired and the chance of receiving a raise on the same graph. They intersect at .450, which is logical as it is the point where a coach starts to move from being average to being below average.

*Figure 10.*



Chance to be fired or get a raise all observations
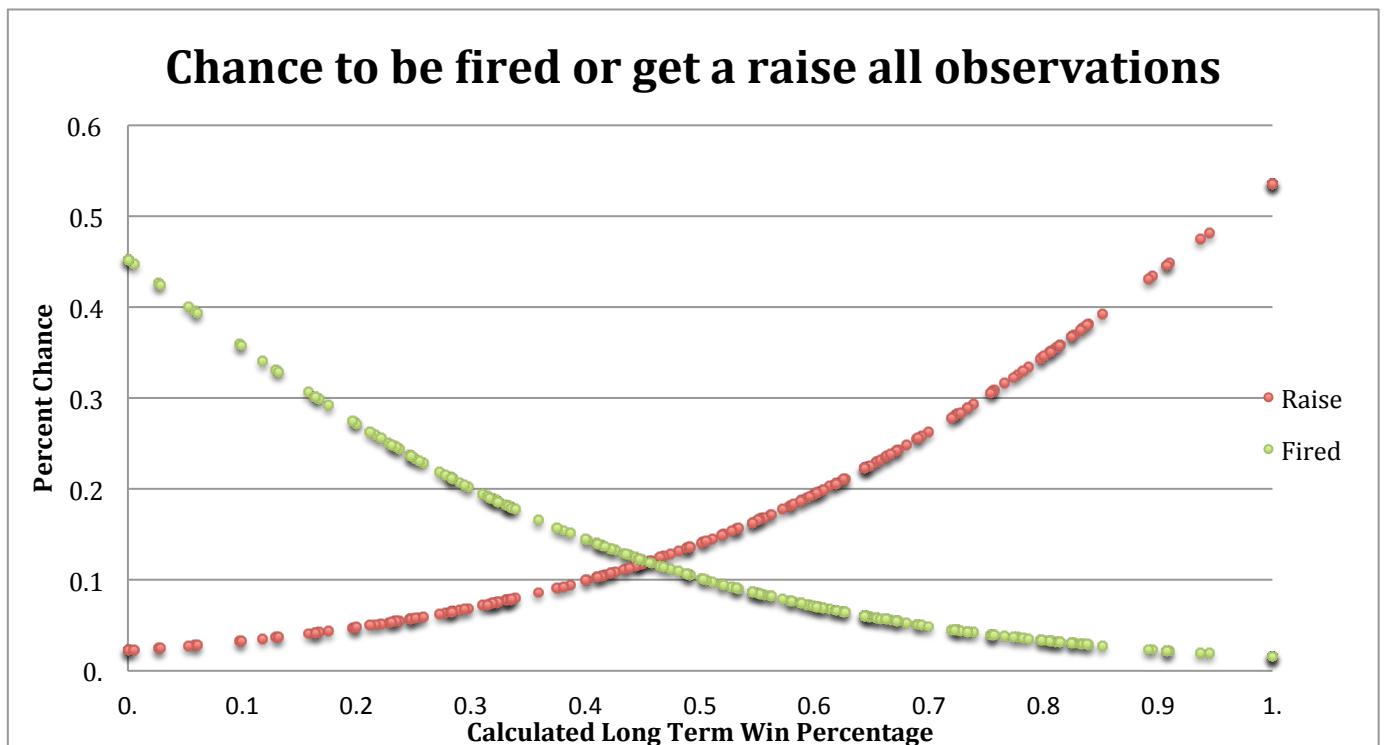
*C. Annual Change in Coach Salary*

In this section I explore how well the long term win rates generated from the model predict changes in a coach's salary in the following season. I regressed the calculated long-term win rates for each coach at an "average" institution, which has an expected recruiting class value of 1,223 against the changes in salary each year. The aim of this regression was to see if, on average a successful season according to the model lead to an increase in salary from a new contract.

$$\kappa_{t+1} - \kappa_t = \beta_1 * \varpi_t \ (14)$$

Where $\kappa_t$ is the coach's salary as a percent of the average in year t and $\varpi_t$ is the coach's long term expected win rate at the average school expresses this relationship. Regressing this equation using the data from 2009-2012 gives the following significant results.

*Table 7. (Long term calculated win percentage)*

| R^2 = 0.05652 | Coefficients | Standard Error | t-Stat |
|---|---|---|---|
| **Intercept** | -0.0934 | 0.0354 | -2.6323 |
| (Long term win rate) $\beta_1$ | 0.2159 | 0.0629 | 3.435 |

The coefficient of 0.2159 means that for each 10% increase in long term win percentage based on the model the coach salary is expected to increase by 2.16% of the average salary in the next season.

While the model does explain changes in salary significantly, the R^2 value is only 0.0562. I expect part of the reason for this is that schools tend to be overly focused on the headline number, conference win percentage, when deciding whether to give coaches new contracts. Instead schools should be evaluating

coaches based on more complex comparatives such as this model. To test this

hypothesis I will regress the following equation:

$$\kappa_{t+1} - \kappa_t = \beta_1 * \gamma_t \text{ (15)}$$

Once again $\kappa_{t+1} - \kappa_t$ represents the change in salary as a percent of average. $\gamma_t$ is

simply the conference win percentage of a coach's team in the given year. This gives

the following results:


*Table 8. (simple win percentage)*

| R^2 = 0.0797 | Coefficients | Standard Error | t-Stat |
|---|---|---|---|
| **Intercept** | -0.144 | .0417 | -3.452 |
| (win percentage) $\beta_1$ | 0.2859 | 0.0692 | 4.13 |


The results are once again significant. Although we cannot directly compare R^2

values it is interesting to note that the R^2 for this regression is slightly higher. This

suggests that my hypothesis that schools look too closely at win percentage when

deciding to give contract extensions could be correct. It is also important to note

that we would expect any regression that is significant for the model's long term win

percentage for a given coaching performance to also be significant for the win

percentage of the coach in the given year since these two values are highly

correlated[7].

While both results are statistically significant there is not much more that

can be determined from this data about what specifically motivates schools to

---

[7] R^2 of .62 when regressed against each other

change a coach's salary after a given season. It is unclear that my model describes

the changes any differently than the naïve win percentage model.


**IV. Conclusion and Future Improvements**

*A. Conclusion*

By separating coaching performance from the advantages of coaching at a

powerhouse football school, the model has shown that college coaches appear to be

at least paid somewhat efficiently by performance and not just by the luck of the

school they end up at. Currently the model does a far better job at explaining

recruiting than the on field performance model. As a result I believe this is why the

recruiting model shows coaches compensated efficiently, where better recruiters

are paid more, while the on field model does not show any trend. When the models

are combined to solve for a long-term win rate for each coaching performance, it

shows that the best coaches are more highly compensated than the rest. The model

successfully identified two of the top coaches as owners of its favorite seasons,

which shows that in terms of identifying the best coaches it does quite well.

The model does a fairly good job of explaining extensions, firings, and

changes in coach salary. However, it is unclear that it explains these things

significantly differently from simply looking at win percentage as a casual fan would

do. Below is a graph illustrating this for likelihood of be fired.

*Figure 11.*



**Chance to be fired using Complex and Naive model**
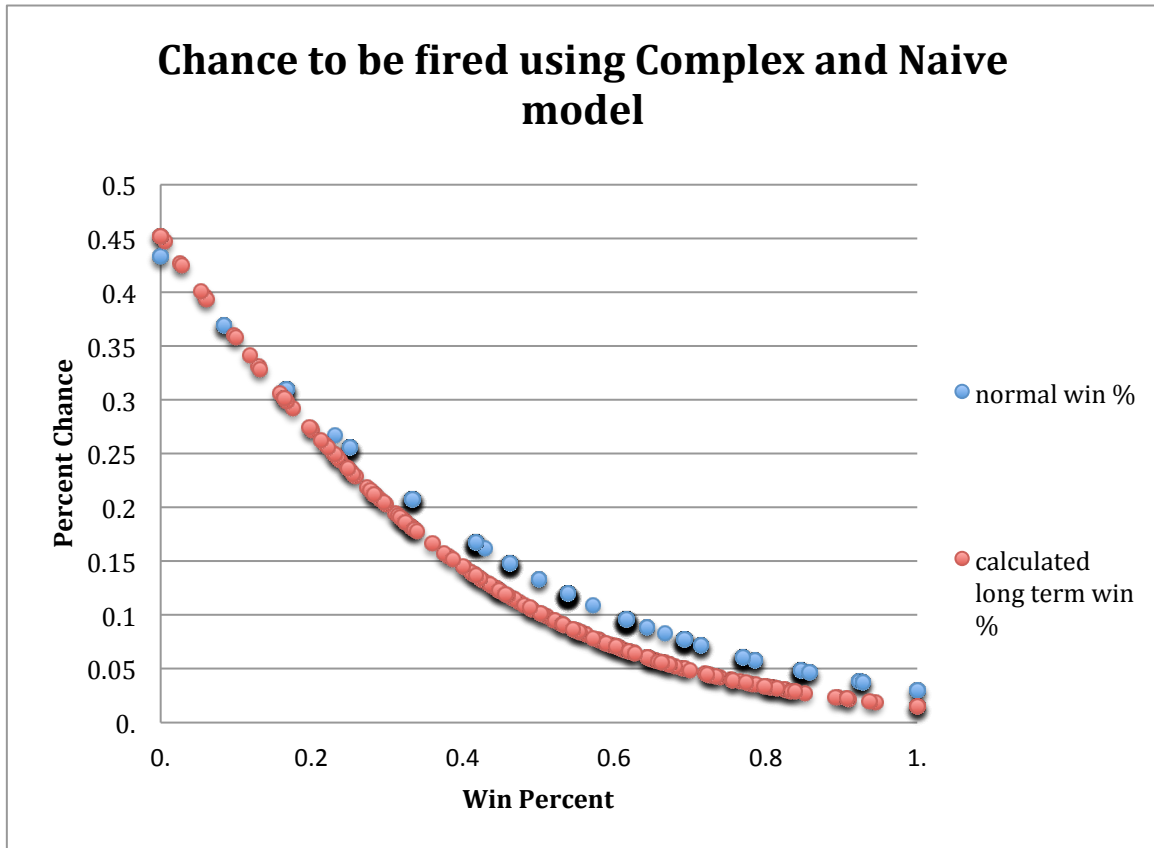
The difference between the probability levels for the two are clearly very small. While normal win percentage is a significant part of the model for the calculated long term win percent, it would be better, if there was a clearer divergence between the two when explaining coach salary decisions. This would allow a conclusion to be reached about what information athletic directors are looking at when making decisions.

*B. Further Questions and Considerations*

The most obvious question is does the complex model actually predict anything different than simply using the naïve model. With the current data it is hard to determine whether this is the case.

Another interesting point to consider is that some coaches may be inherently better at identifying highly talented players who are not rated as such by Rivals.com. In the long term this would show up as the coach outperforming his expected win total with his given talent pool but in the short term it may not. If a coach has only been at a school for one or two years, the players he has recruited will not yet be making a significant impact for the team since many of them will be redshirting and developing. Thus the model will not consider this important advantage that the coach has over other coaches who do not evaluate talent as well as him and the coach will, therefore, be underrated.

Another interesting factor to explore is the effect of variance and outlier seasons on individual coach's contracts. For example if a coach generally coaches at around a .600 level in his career, according to the model, does one .050 season or one .950 have a larger effect on his salary than the string of .600 seasons he has shown to be his norm? Given a data set spanning more years this would be possible to calculate.

An additional use of the recruiting model is to measure the effects a school investing in increasing facilities. School will be able to predict how increasing stadium size and other facilities would improve recruiting classes and in turn how many wins this will get them.

**Sources and References:**

1. John Vrooman

Theory of the Perfect Game: Competitive Balance in Monopoly Sports Leagues

Review of Industrial Organization (2009) 34:5-44


2. Robert Brown

Estimates of College Football Player Rents

Journal of Sports Economics April 2011 12: 200-212,


3. Robert Brown

Do NFL Player Earnings Compensate for Monopsony Exploitation in College?

Journal of Sports Economics August 2012 13: 393-405,


4. Robert Brown

An estimate of the rent generated by a premium college football player

Economic Inquiry 31.4 (Oct 1993): 671.


5. Rodney K. Smith, A Brief History of the National Collegiate Athletic

Association's Role in Regulating Intercollegiate Athletics, 11 Marquette. Sports

L. Rev. 9 (2000)

Available at: http://scholarship.law.marquette.edu/sportslaw/vol11/iss1/5

6. Chris Fallica, Nick Loucks, and Harold Shelton. "Counting down College Football's Most Prestigious Programs (1-5)." *ESPN.com*. N.p., 23 Jan. 2009.

http://sports.espn.go.com/ncf/news/story?id=3849468

7. "USA TODAY Sports." USA TODAY Sports NCAA Coaches Salaries.

http://sports.usatoday.com/ncaa/salaries/

8. Rivals.com recruiting rankings 2002-2015:

http://sports.yahoo.com/footballrecruiting/football/recruiting/teamrank/2012/all/all

9. Allen R. Sanderson and John J. Siegfried. "The Case for Paying College Athletes." *Journal of Economic Perspectives* Winter 29.1 (2015): 115-38.

10. Wikipedia, list of American Football Stadium Capacities

http://en.wikipedia.org/wiki/List_of_American_football_stadiums_by_capacity

11.

Joel G. Maxcy, *Journal of Sports Economics, August 2013; vol. 14, 4: pp. 368-388*

Efficiency and Managerial Performance in FBS College Football: To the Employment and Succession Decisions, Which Matters the Most, Coaching or Recruiting?