Machine Learning and the NFL Field Goal

By James LeDoux

1. ABSTRACT

The new NFL extra point rule requires a kicker to attempt his extra point with the ball snapped from the 15 yard line. This attempt is 13 yards longer than under the previous rule, stretching an extra point to the equivalent of a 32-yard field goal attempt from its original 19. Though a 32-yard attempt is still a chip shot to any professional kicker, many NFL analysts were surprised to see the number of extra points that were missed at the onset of the 2015 NFL season, when the rule change was first implemented. Should this really have been a surprise, though? Field goals and extra points, being the relatively isolated events that they are, often have a disproportionately large impact on the outcome of games, as has been analyzed in Clark et al. (2013)[i]. Beginning with a replication of their study, I aim to explore the world of NFL kicking from a statistical perspective, applying econometric and machine learning models to gain a deeper perspective on what exactly makes some field goal attempts more difficult than others. Ultimately, my goal is to go beyond the previous research on this topic and provide an improved predictive model of field goal success.

2. DATA EXPLORATION

The data used in this experiment comes from Armchair Analysis, covering every field goal attempted from the 2000 through 2013 seasons, roughly 14,000 attempts in total[ii]. It is broken down into the binary outcomes of field goal attempts, atmospheric conditions at kickoff time (temperature, precipitation, humidity, wind speed), and various situational variables such as kick distance, time left in the game, the game's score, and whether the kicker was iced. While there is no way to predict probability with perfect precision, the robustness of this data set allows certain models to form predictions that are reasonably close to the true value on this 0-to-1 scale of make or miss.

Here is an overview of the most significant information in the data set:

Distance:

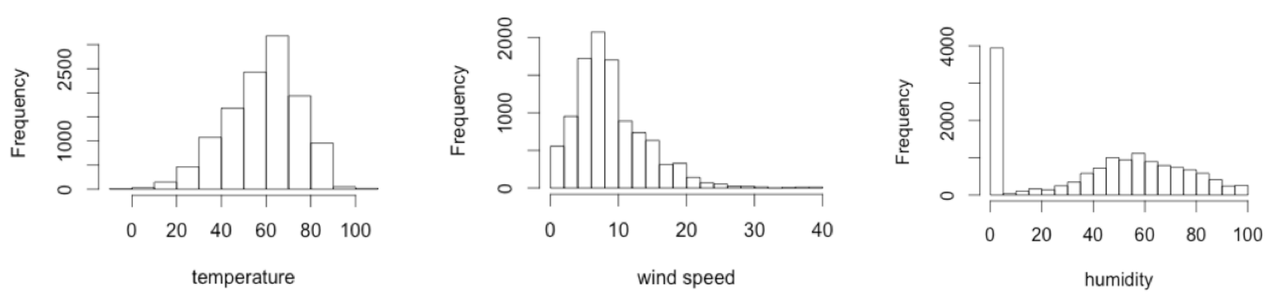
Distance is the most important variable in this data set. The distances in the data range from 18 to 76 yards, with the inner quartile range being from 25 to 45 yards, and the median distance being 37 yards. The distribution of distances in the data set is more or less equally distributed from 20 to 50 yards, becoming sparse beyond 55.

Stadium Factors:

The data also includes the stadium in which the kick took place, whether it took place on turf or natural grass, and whether this happened at home or on the road. Most of these situational factors proved insignificant, and specific stadiums were not included in the final models. There was, however, significant proof to support the much-discussed "Mile-High effect"[iii], under which kickers experience significantly increased range capabilities while playing in the high altitude of the Denver Broncos' Mile High Stadium.

Climate Factors:

Among the climate factors measured are temperature, humidity, wind speed, and precipitation. Here are the distributions for temperature, humidity, and wind speed:



Figures 1, 2, and 3: Temperature, wind speed, and humidity distributions in the data

Given these distributions, I carried out similar feature engineering to that done in the Clark paper, creating a binary "cold" dummy variable for temperatures under 50°F, a "windy" variable for wind speeds greater than 10 mph, and a "humid" one for humidities greater than 60%. The cold and windy dummy variables are useful because both wind speeds and temperatures vary throughout games, while the measurements in this data set are only those recorded at kickoff time. Converting these measurements to categorical variables limits the number of cases that are misclassified due to this limitation in the data set. It should be noted that I do, however, revert to using the continuous temperature variable in certain later models, with temperatures typically being stable within a reasonable range to make the continuous variable useful despite its time-of-measurement flaw. Humidity was not significant in any of the models, and was dropped.

Also noteworthy, the data set contained a variety of classifications for weather conditions, including terms such as "rain", "snow", "light rain", "stormy", "foggy", "clear", and so on. To simplify the ambiguity between these, I lumped all rain and snow-containing conditions into a single dummy variable called "precipitation." Of this, the set has 847 instances with precipitation, and 13,144 without.

There were many other variables to test in the data set, but these were the ones that proved most important to the end-models. The individual effects of each of these variables on the final predictive models can be seen in the appendix, exhibit B.

3. BINOMIAL LOGISTIC REGRESSION MODELS

The first model I test is a logistic regression model, replicating that used by Clark et al. and experimenting with some slight adjustments.

Logistic regression is a non-linear translation of the standard MLR function, measuring probability between the only two possible field goal outcomes of a make and a miss (1 and 0). This probability is calculated by the following function:

$$P(make = 1) = \frac{1}{1 + e^{-(\beta_\theta + \beta_{dist}X_1 + \beta_{cold}X_2 + \beta_{precip}X_3 + \beta_{windy}X_4 + \beta_{turf}X_5 + \beta_{altitude}X_6 + \beta_{iced}X_7)}}$$

The Clark paper leaves in only what is statistically significant at the 5% level (P value < .05). This leads it to only include distance and a series of environmental factors, omitting all situational ones such as point spread, the kicker being iced, the home field, and so on. I test three models, the first being a perfect replication of the Clark model, the second including the icing of the kicker, and the third including statistically significant home-field advantages.

Model 1: Clark Replication
My first task was to replicate the Clark study with more recent data. Clark's model uses data from 2000-2011, so I replicated this with the 2012 and 2013 seasons included to see if there were any significant changes. The coefficients were all close to those from the original study, and each of the variables either became more significant or remained the same. The outputs are as follows on the next page, and can be compared to the originals in exhibit C.

Variable	Coefficient	Std. Error	P-Value
Intercept	5.5232	0.1215	0.0000
Distance (yards)	-0.1024	0.0027	0.0000
Cold (<50°F)	-0.2430	0.0690	0.0004
Precipitation	-0.3205	0.0944	0.0006
Windy (≥10 mph)	-0.2040	0.0509	0.0000
Turf	0.2809	0.0487	0.0000
Altitude (≥4000 ft.)	0.6068	0.2709	0.0250

Clark Replication Regression Outputs

Model 2: Bring in Icing the Kicker

After replicating the study, I brought in the variable for icing the kicker. This is a binary variable equal to one when the opposing team had called a timeout just before the field goal was attempted in a potentially game winning or tying situation, and equal to zero in all other cases. This was excluded in the original paper because it did not meet the significance level of 5%. Although this is still the case in my model, I use it regardless. Given the importance of icing kickers to the outcomes of games, I argue that a 9.1% significance level is good enough for a binary variable such as this one that offers such great explanatory power in crucial scenarios. This model is discussed in greater detail below.

Model 3: Home Field Advantages

For model 3 I regressed all the previously mentioned variables plus the individual stadium dummies on field goal outcomes in order to see which home fields were the most significant. I then iteratively removed the least significant stadiums until I arrived at a list that met the 10% significance level, relaxing my significance levels slightly in order to get a better view of home field advantages in the data set.

The stadiums that were significant at the 10% level and their coefficients are:
Adelphia Coliseum (-.3567), Giants Stadium (-0.5444), Network Associates Coliseum (-0.6063), Paul Brown Stadium (-0.2335), Ralph Wilson Stadium (-0.3736), Texas Stadium (-0.3096), and Veterans Stadium (-0.1536)

The reason why most stadiums were insignificant appears to be the high correlation between many stadiums and their associated weather conditions. Lambeau Field, Soldier Field, and Gillette Stadium, for example, are all storied stadiums with reportedly difficult crowds to play against. These stadiums also, however, have high levels of precipitation and historically cold temperatures, which are highly correlated with these particular stadiums' dummy variables. This problem with collinearity led me to discard this model altogether, with the intuition that each stadium dummy is going to be highly correlated with its home-city's climate conditions and the type of surface of its playing field.

Of the three logistic regression models tested, I find the second model, with icing the kicker included, to be the most useful. Though it does not quite meet the same significance standard as the Clark model, the importance of icing the kicker and the explanatory power it offers in game-winning and tying scenarios – one of the most enticing times to use such a model – makes this a worthwhile sacrifice. Additionally, one sub-10% significance level is certainly not problematic in a model such as this one, where everything else meets a high standard of proof.

With this established as my model of choice, the regression output is as follows:

Variable	Coefficient	Std. Error	P-Value
Intercept	5.5089	0.1361	0.0000
Distance (yards)	-0.1026	0.0030	0.0000
Cold (<50°F)	-0.2170	0.0780	0.0053
Precipitation	-0.3264	0.1062	0.0021
Windy (≥ 10mph)	-0.1730	0.0573	0.0025
Turf	0.3301	0.0548	0.0000
Altitude (≥ 4000ft)	0.8225	0.3213	0.0104
Iced	-0.1730	0.1024	0.0913

When tested against data it had not seen before, here is the layout of the model's predictions measured against their distances:

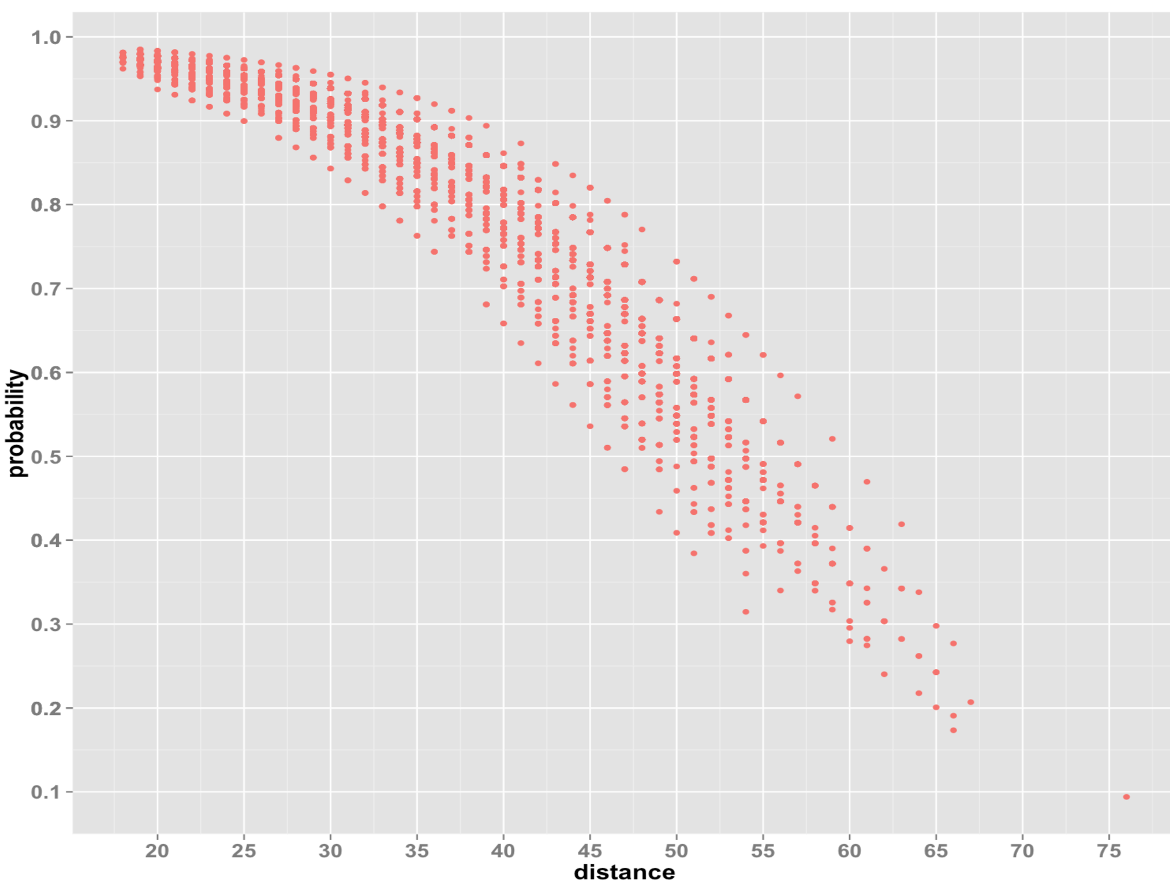


Figure 4: Logistic Regression Model predictions (probability that a field goal is successful). For a larger view, see ***exhibit D.***

Figure 4 shows that the presence of the various environmental factors and the icing of the kicker creates a considerable spread in field goal difficulties, even within the same distance-levels. A 50-yard field goal, for example, can have anywhere between a 41 and 72% probability of success given this wide array of non-distance factors.

4. CONSIDERING FIELD GOAL DIFFICULTY WHEN RATING KICKER ABILITY

The most common metric used to evaluate kicker performance is make percentage. While this is an acceptable surface-level measurement, it makes the false assumption that all kicks are created equal, which is simply not the case. Based on team strategy and quality, some kickers may attempt a greater portion of their field goals from longer distances than others. Also, due to climate factors, a kicker in New England will typically have a more difficult job than one in Oakland, all else held equal. Because of factors such as these, make percentage is a lacking metric; a statistic taking into account field goal difficulty would tell a much deeper story.

Such a metric can be formed using the logistic regression model created above. To do this, one simply takes the difference between each field goal's outcome (0 for miss, 1 for make) and its predicted probability of success, as determined by the model, and multiplies this by 3, the number of points that the field goal is worth if successful. This creates a "points added" metric, showing the number of points that the kicker has generated for his team above the expected value, or, in other words, the value that the kicker has created above that which would be generated by a perfectly average replacement[iv]. This standardized points measure punishes kickers greater amounts for missing "easy" field goals than it does difficult ones, overcoming the shortcomings of using the make percentage statistic alone.

Points Added = (field goal success – probability of success) x 3

This metric, measuring field goal outcomes against their difficulty level, will provide an unbiased assessment of a kicker's skill.

So: who are the best kickers of the modern era?
Using this statistic, one can determine not only who have been the best kickers of the modern era, but also the most underrated, taking the difference between kickers' rankings according to the points-added statistic and those determined by make percent.

Best Kickers by Points Added, FG%, and Degree Underrated:

Kicker	Points-Added Per Attempt	Pct. Made	Points-Added Rank	Percent Rank (FG%)	Degree Underrated	Attempts (2000-2013)
Justin Tucker	0.372615842	0.923	1	1	0	78
Dan Bailey	0.304864407	0.908	2	2	0	98
Kai Forbath	0.30107944	0.875	3	4	1	40
Blair Walsh	0.282920419	0.898	4	3	-1	69
Greg Zuerlein	0.198175934	0.830	5	24	19	59
Rob Bironas	0.19545974	0.855	6	10	4	283
Connor Barth	0.187260008	0.841	7	19	12	120
Robbie Gould	0.166262695	0.863	8	7	-1	278
Steven Hauschka	0.15511815	0.872	9	5	-4	125
Dan Carpenter	0.145006194	0.838	10	21	11	192

This measurements can also tell you who is most overrated (ranked higher by FG% than by points added):

Kicker	Points-Added Per Attempt	Pct. Made	Points-Added Rank	Percent Rank	Degree Underrated	Attempts
John Carney	-0.031961265	0.829861111	42	25	-17	288
Garrett Hartley	-0.046479034	0.825688073	46	29	-17	109
Mike Vanderjagt	0.032939242	0.850961538	28	12	-16	208
Stephen Gostkowski	0.047780284	0.859848485	23	8	-15	264
Gary Anderson	0.013872664	0.844262295	31	16	-15	122
Morten Andersen	-0.021497815	0.82967033	41	26	-15	182
Matt Bryant	0.034029428	0.848387097	27	14	-13	310
Matt Stover	0.069996328	0.868656716	17	6	-11	335
Mike Nugent	-0.061526922	0.8125	48	37	-11	208
Lawrence Tynes	-0.092412104	0.805555556	53	42	-11	252

And most underrated (ranked higher by points added than by FG%):

Kicker	Points-Added Per Attempt	Pct. Made	Points-Added Rank	Percent Rank	Degree Underrated	Attempts
Sebastian Janikowski	0.094237196	0.801339286	14	44	30	448
Randy Bullock	-0.021045149	0.742857143	40	63	23	35
Caleb Sturgis	0.01259131	0.764705882	32	54	22	34
Josh Scobee	0.064687163	0.808118081	19	40	21	271
Greg Zuerlein	0.198175934	0.830508475	5	24	19	59
Connor Barth	0.187260008	0.841666667	7	19	12	120
Dan Carpenter	0.145006194	0.838541667	10	21	11	192
John Kasay	0.122759211	0.838509317	11	22	11	322
Ryan Succop	0.043994456	0.813333333	25	36	11	150
Mason Crosby	0.001348368	0.794354839	36	47	11	248

Sebastian Janikowski and Stephen Gostkowski are two of the great stories of this metric.

Gostkowski is widely touted as the best kicker in the game today. Much of his praise is well earned: he is consistently among the most accurate kickers in the NFL as measured by field goal percent (in the two most recent seasons, not in this data set, he has scored on 94.6 and 91.7%[v] of his attempts). Gostkowski also, however, plays for the high-powered New England offense, and part of his high make percentage appears to result from him being set up for large amounts of easier field goals. Because of this, while he is ranked 8th in field goal percentage, he is only 23rd in points-added, making him overrated by 15 positions on the list. Granted, given his successful 2014 and 2015 seasons, it is likely that these ranks have improved since 2013.

Janikowski, on the other hand, plays for the woeful Oakland Raiders. The Raiders have struggled offensively throughout the past several seasons, and relied greatly on Janikowski's powerful leg during the period tested. As a result, he was set up with a disproportionately high amount of long field goal attempts, and his 80% field goal accuracy from 2000-2013 (ranked 44th) does not tell the full story of his impact. Due to his ability to convert from 50+ yards, his points-added score is quite high at .09/attempt, making him underrated by 30 positions on the list.

Who had the best seasons?
Finally, one can also use this statistic to measure which kickers have had the greatest impacts on individual seasons. Taking the sum of a kicker's points-added by year gives his total points added for each season, showing exactly how valuable a kicker was to his team each year. Here are the greatest single-season performances according to this measurement:

Kicker	**Season**	**Team**	**Points Added**	**Season Attempts**
Neil Rackers	2005	ARI	19.684	42
Sebastian Janikowski	2009	OAK	18.926	29
Sebastian Janikowski	2011	OAK	17.371	35
Mike Vanderjagt	2003	IND	16.595	40
Steven Hauschka	2013	SEA	15.781	43
Justin Tucker	2013	BAL	15.726	41
Stephen Gostkowski	2013	NE	15.591	42
Rob Bironas	2011	TEN	15.340	32
Phil Dawson	2013	SF	15.161	42
Blair Walsh	2012	MIN	14.574	39

All of the above kickers were crucial to their teams' offenses these seasons. As you can see, one reason many people appear on this list is high kick quantity, which is to be expected – kickers who are relied upon more heavily have increased chances of accumulating points-added for their teams. This makes it all the more impressive that Sebastian Janikowski's 2009 season makes number two on this list with only 29 attempts. Although the raiders were a losing team this season, they would have been far worse off had it not been for Janikowski bailing out their poor offense from beyond the 50 yard line multiple times over, netting an additional 19 points that the Raiders would not have secured with a replacement kicker.

5. IMPROVING PREDICTIONS WITH MACHINE LEARNING
Finally, while there is a great deal to be learned from the logistic regression model and the measurements that follow from its predictions, I could not resist the temptation of taking this a step further to see what additional insight machine learning models could provide in this situation. Using the same data, I tested random forest, neural network, multiple linear regression, and hybrid models to see how they fared against the logistic regression benchmark.

Random Forests
A random forest is an ensemble learning algorithm that takes a large number of deep decision trees, and uses the output of each tree as a "vote" toward what the forest's output should be. By taking the average of the many decision trees' votes, the random forest model avoids some of the overfitting and rigidity problems associated with using individual trees, and often produces accurate predictions. In this case, however, the random forest underperformed the other models tested, and did not prove useful for prediction improvement.

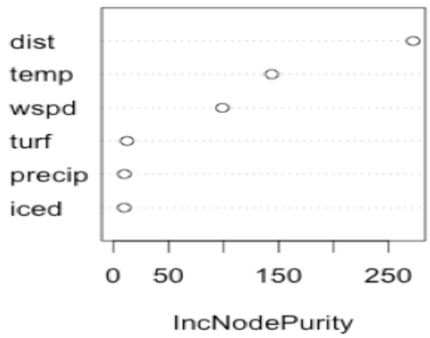


Figure 5: Random Forest Feature Importances

Neural Networks

Neural Networks were the next type of model tested. This class of model loosely imitates the structure of neurons in the brain, beginning with one node per variable in the outer layer, a number of "hidden" nodes in the middle layer, and then an end node for the model's outputs. The network uses a backpropogation algorithm to determine the optimal weights between its nodes, which allows it to improve the accuracy of its outputs[vi]. These models have gained a great deal of support in machine learning for their ability to gain deep insights on the patterns within data.

Although I tested deeper networks, this simple 7x3x1 network proved most effective:

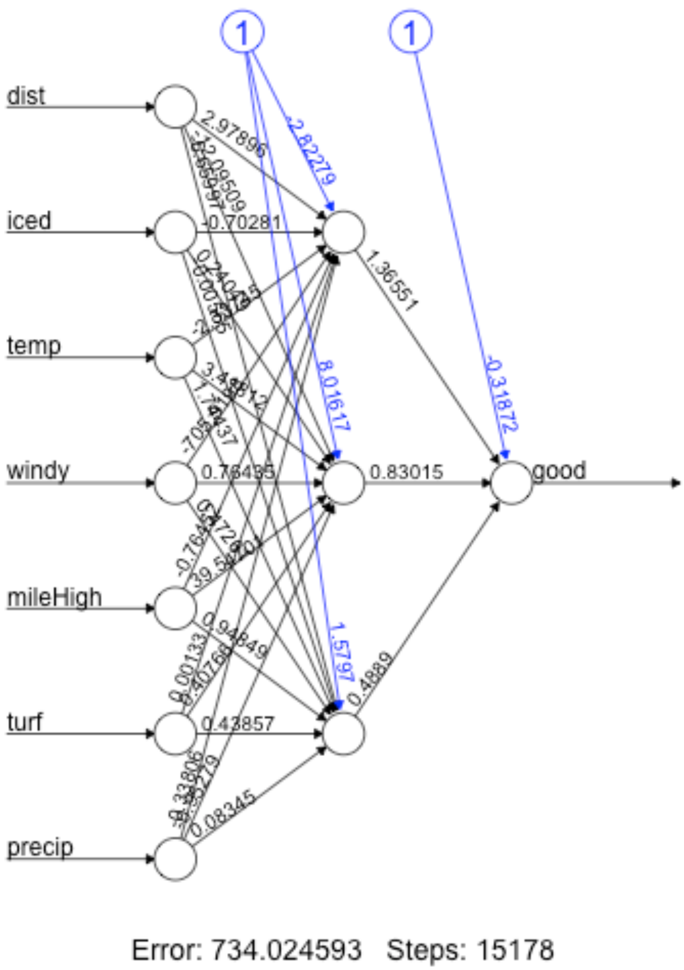


Figure 6: 7x3x1 neural network architecture

While this was by far the most time consuming model to train and implement, it also yielded the best results of any individual model tested.

Multiple Linear Regression

Last, I tested a multiple linear regression model to see how that would fare. Because of the logistic nature of the relationship between distance and field goal difficulty (the difference between a 55 and 65-yard field goal's difficulty is much larger than that between a 25 and 35-yard attempt), simply throwing all the above variables into a standard MLR model will not fit the data well. Adding the log of the distance into a model with all of the previously used variables, however, yielded quality results, with an SSR measurement beating that of the random forest model and the original Clark logistic model. This MLR model allows for the non-distance variables to be linear intercept variables, and also makes for the highest ease of interpretation of all the models used, using the standard MLR model of:

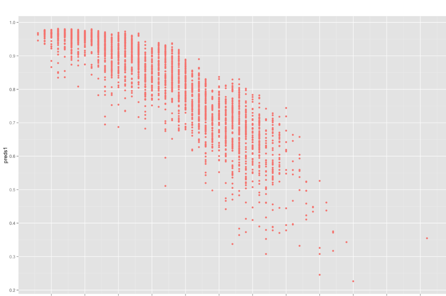
$$P(make) = \beta_{\theta} + \beta_{dist}X_1 + \beta_{\ln(dist)}X_2 + \beta_{cold}X_3 + \beta_{precip}X_4 + \beta_{windy}X_5 + \beta_{turf}X_6 + \beta_{alt}X_7 + \beta_{iced}X_8$$

With regression outputs:

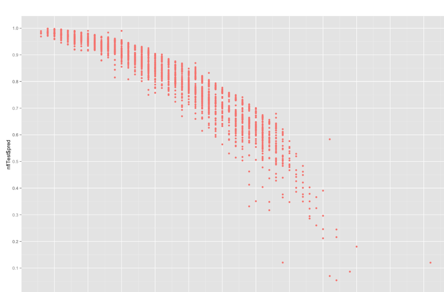
Variable	**Coefficient**	**Std. Error**	**P-Value**
Intercept	-0.3806	0.2146	0.0762
Distance (yards)	-0.0318	0.0024	0.0000
ln(distance)	0.6473	0.0850	0.0000
Temp (°F)	0.0008	0.0002	0.0001
Precipitation	-0.0398	0.0147	0.0069
Windy (≥ 10mph)	-0.0222	0.0076	0.0037
Turf	0.0417	0.0070	0.0000
Iced	-0.0272	0.0146	0.0629

Different Models, Different Predictions

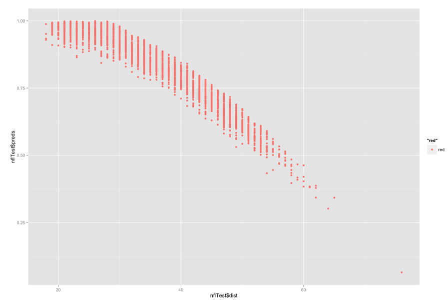
It should not come as a surprise that different models yield slightly different predictions. The regression outputs and error measurements mentioned up until this point in the paper, however, do not display model patterns well. Because distance is such a strong predictive factor in this particular set of models, simply showing probabilities graphed against distance seems to do the trick for illustrating probability patterns. Here's how the models turned out:



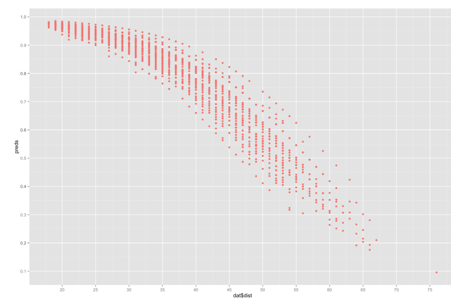
Random Forest



Neural Network



Multiple Linear Regression



Logistic Regression

Figures 7, 8, 9, and 10: model results

The most important takeaway from these figures is that certain models are more flexible than others. The random forest and neural network, being the most fluid, show a wider variance in potential outcomes than the more rigidly modeled logistic and multiple linear regression functions. In many cases this leads to a higher predictive power, since real-world scenarios such as field goal attempts may not perfectly fit mathematical functions.

And the winner is. . .

While logistic regression is the standard for probabilistic measurements such as this, it was not necessarily the most precise of those tested. The best performing model in this situation is actually one that has not been mentioned until this point. While the neural network performed best of all the models already discussed, the best performer of all was actually one that combined the results of this and the second logistic regression model, taking the average of the two's predictions. Here is how all the models compared in terms of SSRs (sums of squared residuals):

Rank	Model	SSR
1	Neural Net and Logistic #2 Combined	369.78
2	Neural Network	370.19
3	Logistic #2 (icing included)	370.34
4	Multiple Linear Regression	370.42
5	Logistic #1 (Clark replication)	370.67
6	Random Forest	374.40

For those who do not use this measurement often, the sum of squared residuals is an error measurement taking the sum of the squared differences between the predicted values (from the model) and their true values (1 if made, 0 if missed) across all predictions made. The training and test data were the same across all tested models, so this is a fair test across models. The smaller the SSR score, the more accurate the model.

While the neural network, MLR model, and logistic regression model with icing included were able to beat the logistic regression benchmark set by Clark, the best model of all was that which combined the predictions of the top two performers.

6. SO DOES ANY OF THIS REALLY MATTER?

To begin to answer, I present two scenarios.

The first took place on January 19, 2002. The New England Patriots were playing the Oakland Raiders in the AFC divisional playoff game. The Patriots were down 13-10 when Adam Vinatieri lined up for a 45-yard attempt with time running low. In 19-degree weather and heavy snow, Vinatieri's kick went through the uprights with only 27 seconds left, and the Patriots went on to win the game in overtime on another Vinatieri field goal.

This snowy field goal has gone down in NFL history as one of the greatest clutch kicks of all time. It also happens to be an excellent example of one of the worst possible conditions to kick a field goal in, showing the need for models such as those discussed in this paper. The probabilities of Vinatieri's kick going in had the game been in Oakland vs. in New England are illustrated below.

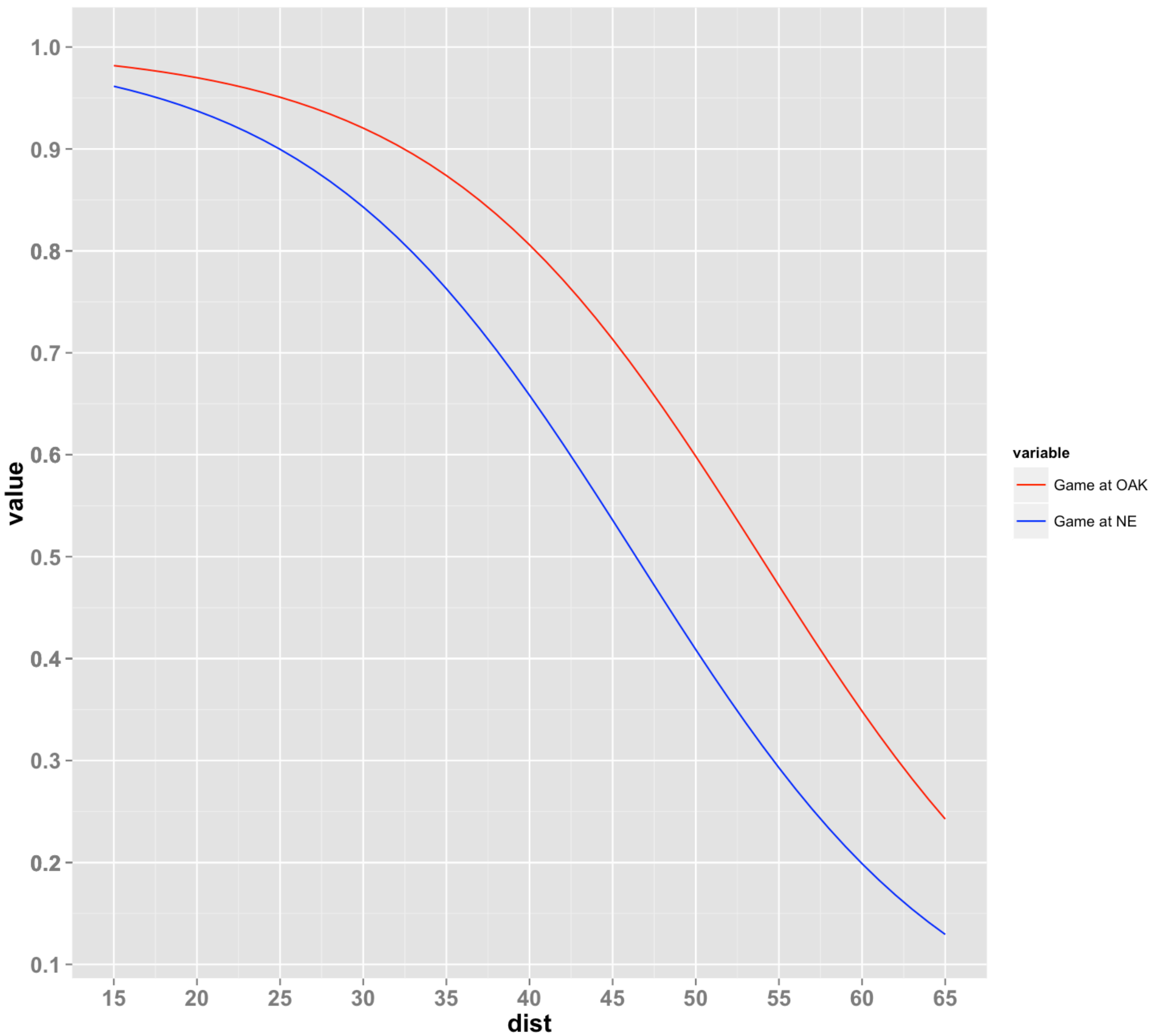


Figure 11: Impact of Weather Conditions on the Snow Bowl Game

Had the game taken place in Oakland with its temperate climate, Vinatieri's 45-yard attempt would have had a very reasonable 71% likelihood of succeeding. In the New England blizzard, however, it had much riskier odds, at 54%. Though the kick was ultimately good, this model goes to show just how high the stakes were that Sunday afternoon, and the amount of pressure that Vinatieri, the coaching staff, and the New England fan base must have felt as they waited for the result.

The second situation is one that hasn't happened yet, but one that every NFL coach should be concerned about: a kicker's missed extra point costing his team the game. Moving the extra point back 13 yards may still seem like an easy conversion, being a 32-yard attempt, but as these models show, there is a noticeable decline in the probability of success in going from a 19 to a 32-yard kick. The result of this decreased likelihood has been seen across the league this season, with extra point conversion percentages decreasing from 99.6 to roughly 95 percent with the implementation of the new rule. While nobody has lost a game due to this rule change just yet, this new area of risk is a perfect example of the importance of understanding field goal likelihoods for making effective coaching decisions. This increased risk will make two-point conversion attempts increasingly appropriate in the NFL.

As is seen in these two scenarios, understanding field goal conversion likelihoods is crucial to a coach's decision-making process. Understanding not only that distance, climate factors, and situational pressure have tangible impacts on field goal success likelihoods, but also to what extent they do, could make all the difference in weighing the risk and potential reward of a crucial coaching decision.

To conclude, this study teaches two main things:

1. First, that field goals follow predictable patterns, allowing their probabilities of success to be modeled to a useful degree. This can and should be taken into consideration by coaching staffs during their evaluations of game time decisions.
2. Second, that good kickers really do matter. Finding a great one like Rob Bironas or Sebastian Janikowski can make all the difference in a 16 game season where playoff outcomes are not determined until the very last week, or in a playoff game where a single point-added above average replacement is all that stands between being a runner-up and a super bowl champion.

While baseball has been thoroughly transformed by sabermetrics and the Moneyball revolution, football, having smaller sample sizes with its 16-game seasons, still has a way to go in its acceptance of statistics as the governing dynamics of certain aspects of play. Field goal kicking, however, being the isolated component of play that it is, has the potential to be one of the first great leaps in the quantification of the ways that coaches, scouts and fans view the game. Through the acceptance of methods and models such as those discussed in this paper, NFL organizations can make great strides in improving their abilities to make data-informed decisions and maximize the potential values of their rosters.

APPENDIX

EXHIBIT A: League-Wide FG% by Distance, 2000-2013

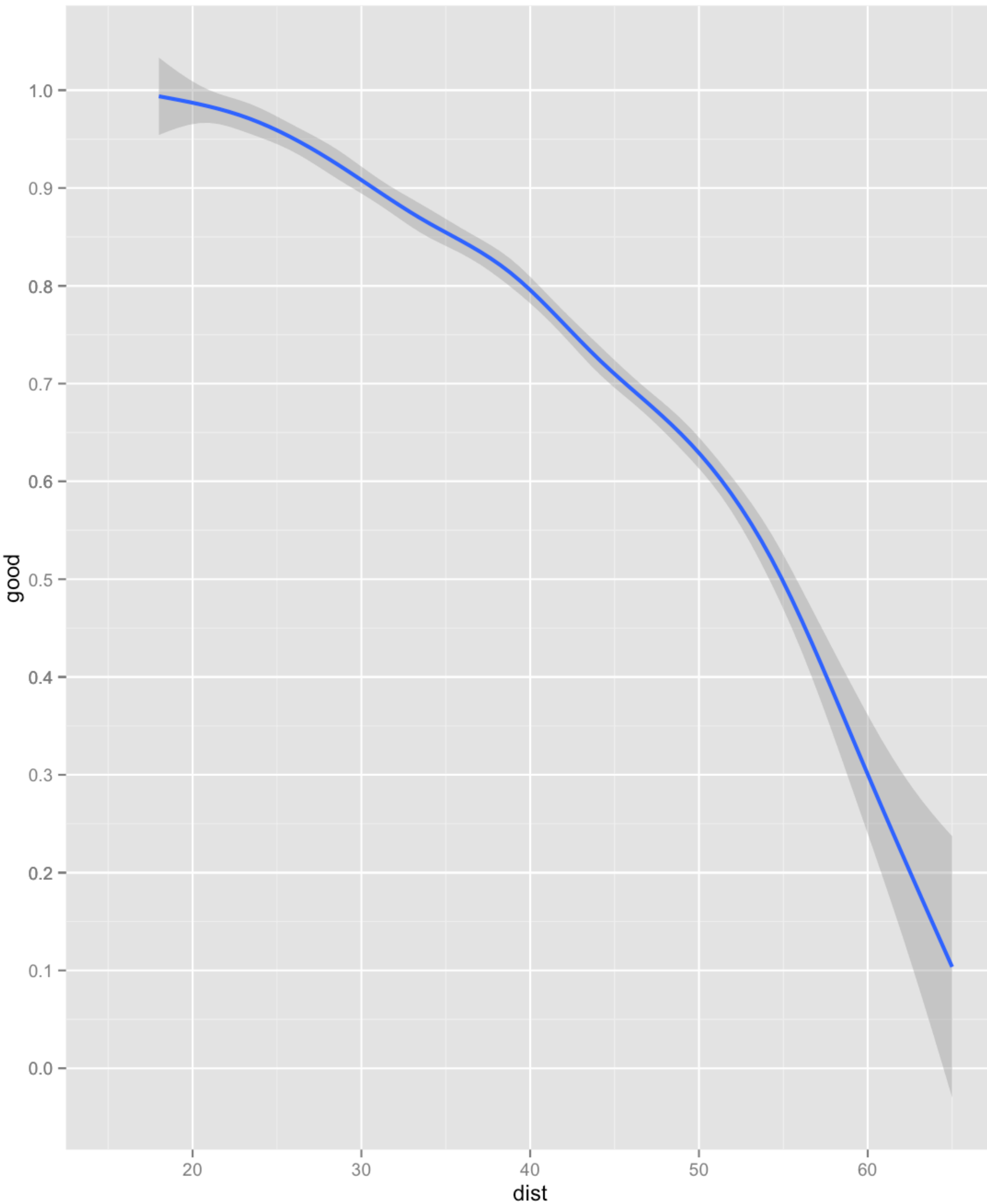
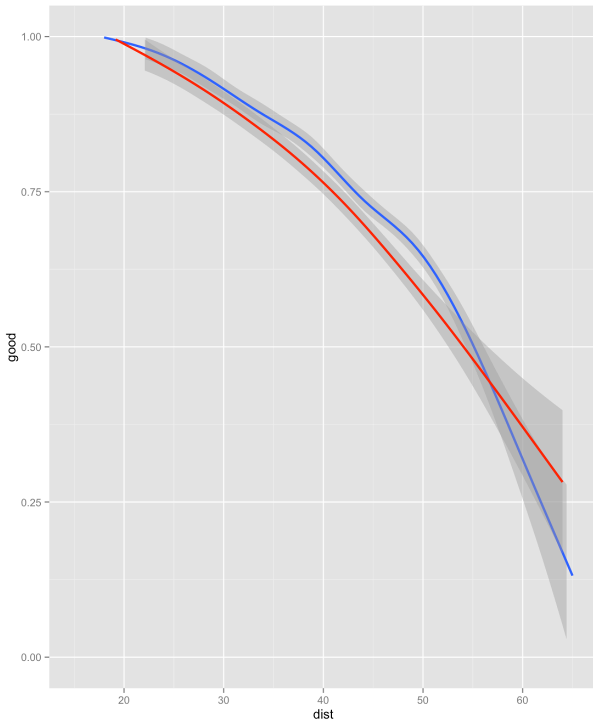
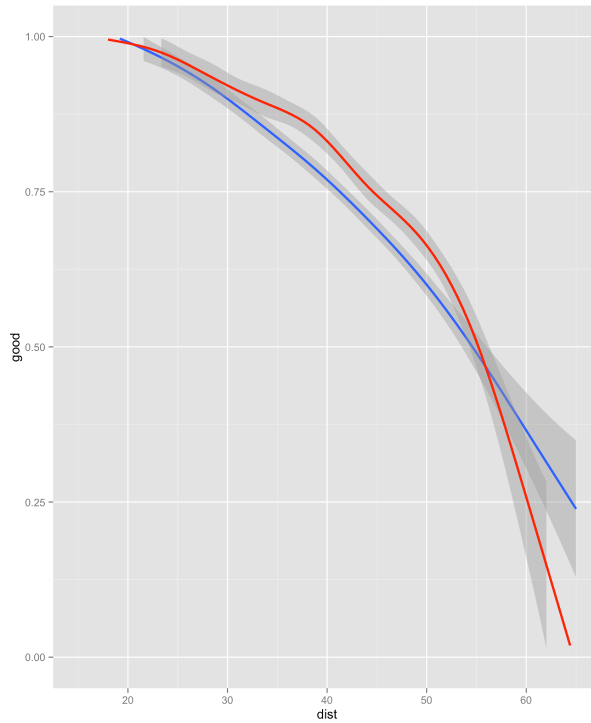


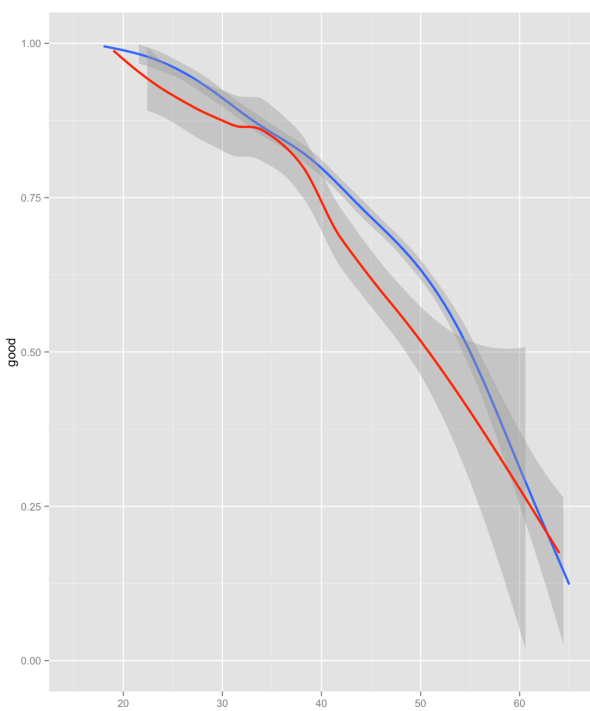
EXHIBIT B: Effects of Parameter Changes



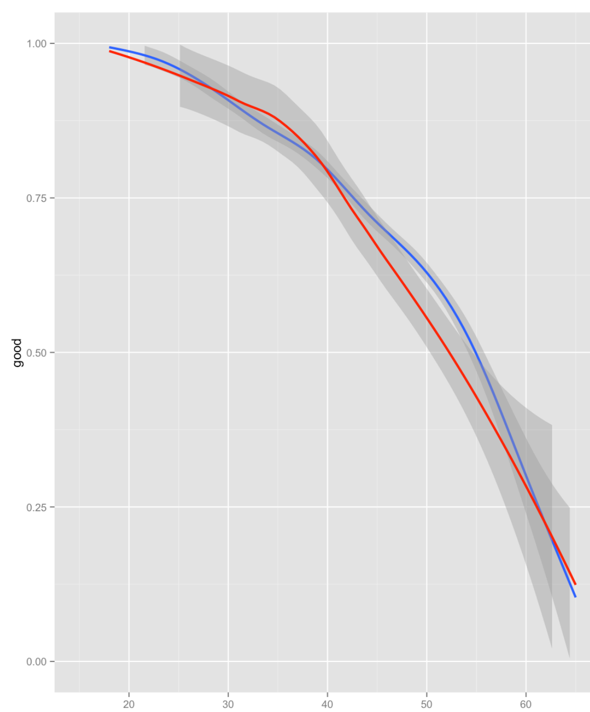
Wind Speed Effect: Red=Windy, Blue=Not Windy



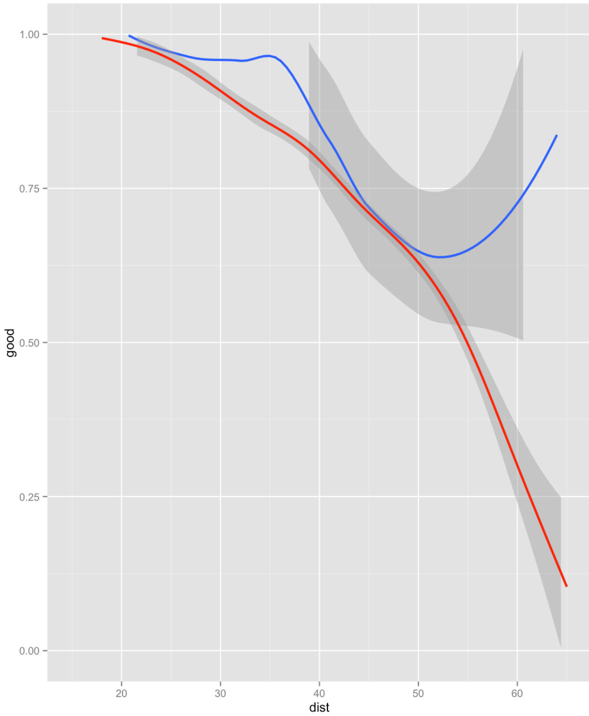
Turf Effect. Red=Turf, Blue=Not Turf



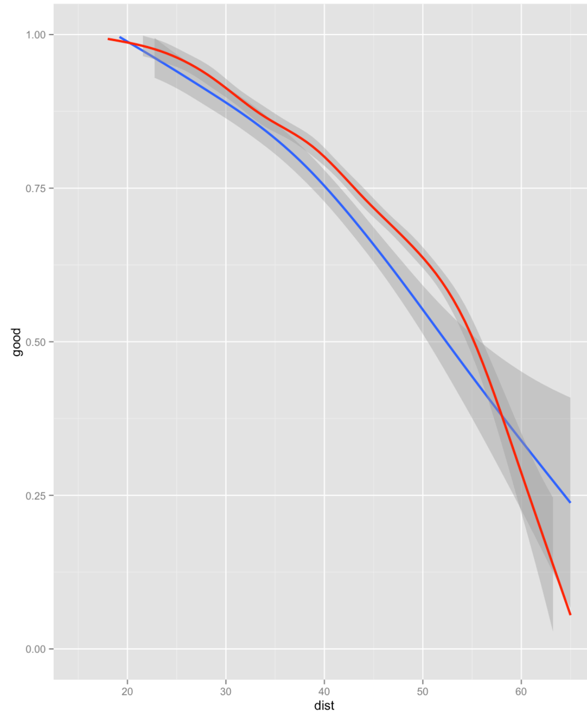
Precipitation Effect. Red=Precipitation, Blue=None



Icing the Kicker: Red=Iced, Blue=Not



Mile-High Effect. Blue=Game Played in Denver, Red= played elsewhere



Cold Weather Effect. Blue=Cold, Red=Not Cold

EXHIBIT C: Original Study Coefficients

Variable	Coefficient	Std. Error	P-Value
Constant	5.953	0.220	<0.0005
Distance (yards)	-0.106	0.003	<0.0005
Cold (<50°F)	-0.341	0.061	<0.0005
Precipitation	-0.280	0.099	0.005
Windy (≥10 mph)	-0.140	0.055	0.011
Turf	0.299	0.053	<0.0005
Altitude (≥4000 ft.)	0.694	0.157	<0.0005

EXHIBIT D: Probability Predictions from Logistic Model #2

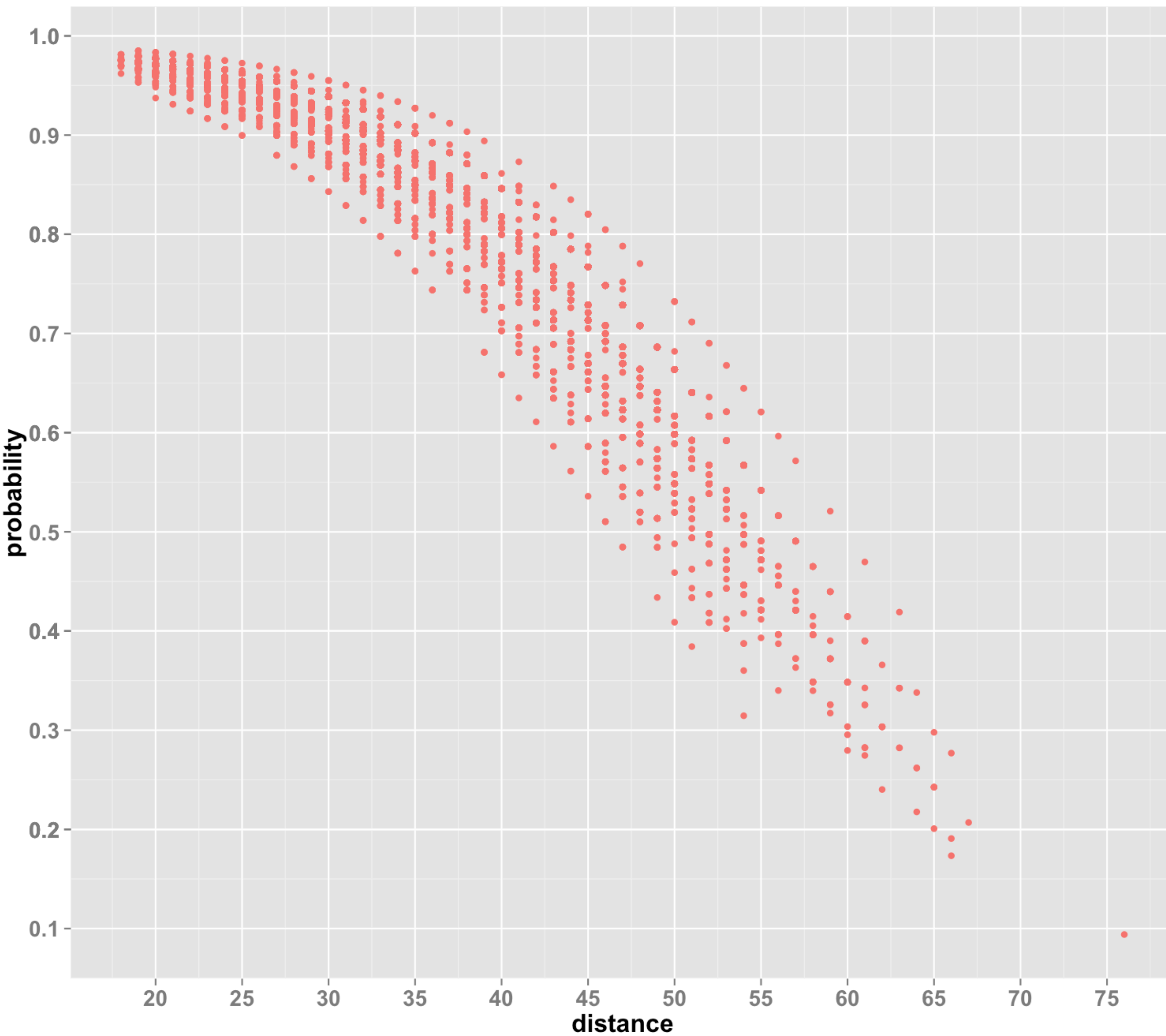


EXHIBIT E: Best Kickers by Career Points-Added per Attempt

Kicker	Points Added	Total Attempts	Make Percent	Points Added Rank	Percent Rank	Degree Underrated
Justin Tucker	0.372615842	78	0.923076923	1	1	0
Dan Bailey	0.304864407	98	0.908163265	2	2	0
Kai Forbath	0.30107944	40	0.875	3	4	1
Blair Walsh	0.282920419	69	0.898550725	4	3	-1
Greg Zuerlein	0.198175934	59	0.830508475	5	24	19
Rob Bironas	0.19545974	283	0.855123675	6	10	4
Connor Barth	0.187260008	120	0.841666667	7	19	12
Robbie Gould	0.166262695	278	0.863309353	8	7	-1
Steven Hauschka	0.15511815	125	0.872	9	5	-4
Dan Carpenter	0.145006194	192	0.838541667	10	21	11
John Kasay	0.122759211	322	0.838509317	11	22	11
Phil Dawson	0.111306907	395	0.855696203	12	9	-3
Joe Nedney	0.103201599	250	0.848	13	15	2
Sebastian Janikowski	0.094237196	448	0.801339286	14	44	30
Adam Vinatieri	0.082278075	456	0.842105263	15	17	2
Jason Hanson	0.078895436	360	0.836111111	16	23	7
Matt Stover	0.069996328	335	0.868656716	17	6	-11
Alex Henery	0.068124529	88	0.852272727	18	11	-7
Josh Scobee	0.064687163	271	0.808118081	19	40	21
Shayne Graham	0.056064831	306	0.849673203	20	13	-7
Nate Kaeding	0.049250632	225	0.84	21	20	-1
Jeff Wilkins	0.048260208	253	0.841897233	22	18	-4
Stephen Gostkowski	0.047780284	264	0.859848485	23	8	-15
Josh Brown	0.044610945	331	0.821752266	24	31	7
Ryan Succop	0.043994456	150	0.813333333	25	36	11
Ryan Longwell	0.035812387	354	0.824858757	26	30	4
Matt Bryant	0.034029428	310	0.848387097	27	14	-13
Mike Vanderjagt	0.032939242	208	0.850961538	28	12	-16
Jason Elam	0.032210377	311	0.826366559	29	28	-1
Matt Prater	0.030412498	186	0.817204301	30	35	5
Gary Anderson	0.013872664	122	0.844262295	31	16	-15
Caleb Sturgis	0.01259131	34	0.764705882	32	54	22
Shaun Suisham	0.011586012	233	0.819742489	33	33	0
Jeff Reed	0.00760398	277	0.826714801	34	27	-7
David Akers	0.003802662	516	0.817829457	35	34	-1
Mason Crosby	0.001348368	248	0.794354839	36	47	11
Jay Feely	-0.00178621	414	0.821256039	37	32	-5
Nick Novak	-0.00389161	125	0.808	38	41	3
Neil Rackers	-0.009086235	343	0.795918367	39	46	7
Randy Bullock	-0.021045149	35	0.742857143	40	63	23
Morten Andersen	-0.021497815	182	0.82967033	41	26	-15
John Carney	-0.031961265	288	0.829861111	42	25	-17
Nick Folk	-0.038181811	212	0.801886792	43	43	0
Graham Gano	-0.04462322	119	0.781512605	44	50	6
Rian Lindell	-0.044645628	377	0.811671088	45	38	-7
Garrett Hartley	-0.046479034	109	0.825688073	46	29	-17
Olindo Mare	-0.060422661	336	0.80952381	47	39	-8
Mike Nugent	-0.061526922	208	0.8125	48	37	-11
Brett Conway	-0.081891753	59	0.796610169	49	45	-4
Dan Bailey	-0.083950688	32	0.75	50	59	9
Bill Gramatica	-0.09013164	47	0.787234043	51	49	-2
Doug Brien	-0.092099728	114	0.789473684	52	48	-4
Lawrence Tynes	-0.092412104	252	0.805555556	53	42	-11
Mike Hollis	-0.096380888	87	0.770114943	54	52	-2
Paul Edinger	-0.108234418	181	0.751381215	55	58	3
Kris Brown	-0.116877351	307	0.762214984	56	55	-1
Martin Gramatica	-0.130215505	183	0.759562842	57	56	-1
Jason Hanson	-0.159399519	171	0.748538012	58	62	4
Billy Cundiff	-0.163136036	224	0.767857143	59	53	-6

Todd Peterson	-0.178490496	138	0.775362319	60	51	-9
Dave Rayner	-0.214793473	90	0.722222222	61	65	4
Steve Christie	-0.241079036	120	0.75	62	60	-2
Jeff Chandler	-0.305111432	30	0.733333333	63	64	1
Al Del Greco	-0.330507691	37	0.756756757	64	57	-7
Brad Daluiso	-0.338455124	32	0.75	65	61	-4
Tim Seder	-0.360883813	62	0.709677419	66	66	0
Jose Cortez	-0.424914359	75	0.706666667	67	67	0
Wade Richey	-0.502447418	56	0.660714286	68	68	0
Seth Marler	-0.642635312	33	0.606060606	69	69	0

EXHIBIT F: Best Kickers by Career Field Goal Percent

Kicker	Points Added	Total Attempts	Makes Percent	Score Rank	Percent Rank	Degree Underrated
Justin Tucker	0.372615842	78	0.923076923	1	1	0
Dan Bailey	0.304864407	98	0.908163265	2	2	0
Blair Walsh	0.282920419	69	0.898550725	4	3	-1
Kai Forbath	0.30107944	40	0.875	3	4	1
Steven Hauschka	0.15511815	125	0.872	9	5	-4
Matt Stover	0.069996328	335	0.868656716	17	6	-11
Robbie Gould	0.166262695	278	0.863309353	8	7	-1
Stephen Gostkowski	0.047780284	264	0.859848485	23	8	-15
Phil Dawson	0.111306907	395	0.855696203	12	9	-3
Rob Bironas	0.19545974	283	0.855123675	6	10	4
Alex Henery	0.068124529	88	0.852272727	18	11	-7
Mike Vanderjagt	0.032939242	208	0.850961538	28	12	-16
Shayne Graham	0.056064831	306	0.849673203	20	13	-7
Matt Bryant	0.034029428	310	0.848387097	27	14	-13
Joe Nedney	0.103201599	250	0.848	13	15	2
Gary Anderson	0.013872664	122	0.844262295	31	16	-15
Adam Vinatieri	0.082278075	456	0.842105263	15	17	2
Jeff Wilkins	0.048260208	253	0.841897233	22	18	-4
Connor Barth	0.187260008	120	0.841666667	7	19	12
Nate Kaeding	0.049250632	225	0.84	21	20	-1
Dan Carpenter	0.145006194	192	0.838541667	10	21	11
John Kasay	0.122759211	322	0.838509317	11	22	11
Jason Hanson	0.078895436	360	0.836111111	16	23	7
Greg Zuerlein	0.198175934	59	0.830508475	5	24	19
John Carney	-0.031961265	288	0.829861111	42	25	-17
Morten Andersen	-0.021497815	182	0.82967033	41	26	-15
Jeff Reed	0.00760398	277	0.826714801	34	27	-7
Jason Elam	0.032210377	311	0.826366559	29	28	-1
Garrett Hartley	-0.046479034	109	0.825688073	46	29	-17
Ryan Longwell	0.035812387	354	0.824858757	26	30	4
Josh Brown	0.044610945	331	0.821752266	24	31	7
Jay Feely	-0.00178621	414	0.821256039	37	32	-5
Shaun Suisham	0.011586012	233	0.819742489	33	33	0
David Akers	0.003802662	516	0.817829457	35	34	-1
Matt Prater	0.030412498	186	0.817204301	30	35	5
Ryan Succop	0.043994456	150	0.813333333	25	36	11
Mike Nugent	-0.061526922	208	0.8125	48	37	-11
Rian Lindell	-0.044645628	377	0.811671088	45	38	-7
Olindo Mare	-0.060422661	336	0.80952381	47	39	-8
Josh Scobee	0.064687163	271	0.808118081	19	40	21
Nick Novak	-0.00389161	125	0.808	38	41	3
Lawrence Tynes	-0.092412104	252	0.805555556	53	42	-11
Nick Folk	-0.038181811	212	0.801886792	43	43	0
Sebastian Janikowski	0.094237196	448	0.801339286	14	44	30
Brett Conway	-0.081891753	59	0.796610169	49	45	-4
Neil Rackers	-0.009086235	343	0.795918367	39	46	7
Mason Crosby	0.001348368	248	0.794354839	36	47	11

Doug Brien	-0.092099728	114	0.789473684	52	48	-4
Bill Gramatica	-0.09013164	47	0.787234043	51	49	-2
Graham Gano	-0.04462322	119	0.781512605	44	50	6
Todd Peterson	-0.178490496	138	0.775362319	60	51	-9
Mike Hollis	-0.096380888	87	0.770114943	54	52	-2
Billy Cundiff	-0.163136036	224	0.767857143	59	53	-6
Caleb Sturgis	0.01259131	34	0.764705882	32	54	22
Kris Brown	-0.116877351	307	0.762214984	56	55	-1
Martin Gramatica	-0.130215505	183	0.759562842	57	56	-1
Al Del Greco	-0.330507691	37	0.756756757	64	57	-7
Paul Edinger	-0.108234418	181	0.751381215	55	58	3
Dan Bailey	-0.083950688	32	0.75	50	59	9
Steve Christie	-0.241079036	120	0.75	62	60	-2
Brad Daluiso	-0.338455124	32	0.75	65	61	-4
Jason Hanson	-0.159399519	171	0.748538012	58	62	4
Randy Bullock	-0.021045149	35	0.742857143	40	63	23
Jeff Chandler	-0.305111432	30	0.733333333	63	64	1
Dave Rayner	-0.214793473	90	0.722222222	61	65	4
Tim Seder	-0.360883813	62	0.709677419	66	66	0
Jose Cortez	-0.424914359	75	0.706666667	67	67	0
Wade Richey	-0.502447418	56	0.660714286	68	68	0
Seth Marler	-0.642635312	33	0.606060606	69	69	0

EXHIBIT G: Most Underrated Kickers

Kicker	Points Added	Total Attempts	Percent Made	Score Rank	Percent Rank	Degree Underrated
Sebastian Janikowski	0.094237196	448	0.801339286	14	44	30
Randy Bullock	-0.021045149	35	0.742857143	40	63	23
Caleb Sturgis	0.01259131	34	0.764705882	32	54	22
Josh Scobee	0.064687163	271	0.808118081	19	40	21
Greg Zuerlein	0.198175934	59	0.830508475	5	24	19
Connor Barth	0.187260008	120	0.841666667	7	19	12
Dan Carpenter	0.145006194	192	0.838541667	10	21	11
John Kasay	0.122759211	322	0.838509317	11	22	11
Ryan Succop	0.043994456	150	0.813333333	25	36	11
Mason Crosby	0.001348368	248	0.794354839	36	47	11
Dan Bailey	-0.083950688	32	0.75	50	59	9
Jason Hanson	0.078895436	360	0.836111111	16	23	7
Josh Brown	0.044610945	331	0.821752266	24	31	7
Neil Rackers	-0.009086235	343	0.795918367	39	46	7
Graham Gano	-0.04462322	119	0.781512605	44	50	6
Matt Prater	0.030412498	186	0.817204301	30	35	5
Rob Bironas	0.19545974	283	0.855123675	6	10	4
Ryan Longwell	0.035812387	354	0.824858757	26	30	4
Jason Hanson	-0.159399519	171	0.748538012	58	62	4
Dave Rayner	-0.214793473	90	0.722222222	61	65	4
Nick Novak	-0.00389161	125	0.808	38	41	3
Paul Edinger	-0.108234418	181	0.751381215	55	58	3
Joe Nedney	0.103201599	250	0.848	13	15	2
Adam Vinatieri	0.082278075	456	0.842105263	15	17	2
Kai Forbath	0.30107944	40	0.875	3	4	1
Jeff Chandler	-0.305111432	30	0.733333333	63	64	1
Justin Tucker	0.372615842	78	0.923076923	1	1	0
Dan Bailey	0.304864407	98	0.908163265	2	2	0
Shaun Suisham	0.011586012	233	0.819742489	33	33	0
Nick Folk	-0.038181811	212	0.801886792	43	43	0
Tim Seder	-0.360883813	62	0.709677419	66	66	0
Jose Cortez	-0.424914359	75	0.706666667	67	67	0
Wade Richey	-0.502447418	56	0.660714286	68	68	0
Seth Marler	-0.642635312	33	0.606060606	69	69	0
Blair Walsh	0.282920419	69	0.898550725	4	3	-1
Robbie Gould	0.166262695	278	0.863309353	8	7	-1

Nate Kaeding	0.049250632	225	0.84	21	20	-1
Jason Elam	0.032210377	311	0.826366559	29	28	-1
David Akers	0.003802662	516	0.817829457	35	34	-1
Kris Brown	-0.116877351	307	0.762214984	56	55	-1
Martin Gramatica	-0.130215505	183	0.759562842	57	56	-1
Bill Gramatica	-0.09013164	47	0.787234043	51	49	-2
Mike Hollis	-0.096380888	87	0.770114943	54	52	-2
Steve Christie	-0.241079036	120	0.75	62	60	-2
Phil Dawson	0.111306907	395	0.855696203	12	9	-3
Steven Hauschka	0.15511815	125	0.872	9	5	-4
Jeff Wilkins	0.048260208	253	0.841897233	22	18	-4
Brett Conway	-0.081891753	59	0.796610169	49	45	-4
Doug Brien	-0.092099728	114	0.789473684	52	48	-4
Brad Daluiso	-0.338455124	32	0.75	65	61	-4
Jay Feely	-0.00178621	414	0.821256039	37	32	-5
Billy Cundiff	-0.163136036	224	0.767857143	59	53	-6
Alex Henery	0.068124529	88	0.852272727	18	11	-7
Shayne Graham	0.056064831	306	0.849673203	20	13	-7
Jeff Reed	0.00760398	277	0.826714801	34	27	-7
Rian Lindell	-0.044645628	377	0.811671088	45	38	-7
Al Del Greco	-0.330507691	37	0.756756757	64	57	-7
Olindo Mare	-0.060422661	336	0.80952381	47	39	-8
Todd Peterson	-0.178490496	138	0.775362319	60	51	-9
Matt Stover	0.069996328	335	0.868656716	17	6	-11
Mike Nugent	-0.061526922	208	0.8125	48	37	-11
Lawrence Tynes	-0.092412104	252	0.805555556	53	42	-11
Matt Bryant	0.034029428	310	0.848387097	27	14	-13
Stephen Gostkowski	0.047780284	264	0.859848485	23	8	-15
Gary Anderson	0.013872664	122	0.844262295	31	16	-15
Morten Andersen	-0.021497815	182	0.82967033	41	26	-15
Mike Vanderjagt	0.032939242	208	0.850961538	28	12	-16
John Carney	-0.031961265	288	0.829861111	42	25	-17
Garrett Hartley	-0.046479034	109	0.825688073	46	29	-17

CITATIONS

[i] Clark, Torin, Aaron Johnson, and Alexander Stimpson. "Going for Three: Predicting the Likelihood of Field Goal Success with Logistic Regression." *MIT Sloan Sports Analytics Conference* (2013). Web. 11 Jan. 2016.

[ii] "Affordable NFL Data Plans." *Armchair Analysis*. Web.

[iii] Burke, Brian. "Altitude and Field Goals." *Advanced Football Analytics*. 9 Jan. 2013. Web. 11 Jan. 2016.

[iv] Clark et al.

[v] "Stephen Gostkowski." *NFL.com*. Web.

[vi] Mazur, Matt. "A Step by Step Backpropagation Example." 17 Mar. 2015. Web. 11 Jan. 2016.