# A Model of Score Minimization and Rational Behavior in Golf

Andrew B. Mack[1]

Harvard University

Original Version: May 7, 2015

This Version: July 24, 2015

**Abstract:**

Previous research on golf focuses on data analysis but no theory exists to explain the findings. This paper represents the first step of a larger research project whose goal is to build formal models which can be used to study the strategic decisions made by golfers. As a starting point on which future papers can build, it seems important to have a model rational choice and perfect information which assumes that the representative golfers utility comes only from shooting as low a score as possible. Thus, for each shot that he hits, he will choose, from a number of possible strategies, the one which, ex ante, will produce the lowest expected score. First, I define golf club distances and golf shot error probability using a series of simple equations. Next, I use these in order to examine the choices that the score-minimizing agent will make in different scenarios and under different assumptions. Among other things, my model predicts that the rational golfer will always choose additional yardage over additional accuracy for his tee shots.

## I. Introduction

Golf has been the subject of serious research which it often involves using data to identify the skill elements which make a professional golfer successful. Work published by Mark Broadie and others has mostly sought to improve the statistical methods used to understand player performance but is less focused on tee-to-green[2] strategy decisions.[3] Previous analyses fail to consider the incentives which cause golfers to choose a certain strategy over others. Additionally, there are no theoretical models with which to understand these decisions or to explain the findings of the statistical analyses. The goal of this research project is to develop both a formal framework and the necessary models with which to examine a choice between competing strategies.

In the game of golf, a player is allowed to carry fourteen different golf clubs as established by the *Rules of Golf*. This constitutes the set from which he will choose in order to hit shots of varying distances. The strict objective of the game is to use a set of clubs to navigate a golf ball (1.68 inches in diameter) from a designated starting point called the tee into a hole or cup (4.25 inches in diameter) in as few stokes of the club as possible. This process is referred to as playing a hole. A standard golf course has 18 holes all of which are played in a round of golf.

Every hole has its own unique features which must be navigated. Golf is the only sport played on a field with no specific dimensions.[4] Because courses tend to be laid out over large pieces of land, the array of potential shots which a golfer could face during a round is literally infinite. This forces players to make quick decisions on how to most effectively advance the ball to the hole given his current position. This the strategic behavior which I am interested in exploring. In this paper, I establish expected score minimization as the foundational standard model of rational behavior. In this paper, I define strategic behavior to be irrational if it is inconsistent with a strategy that minimizes expected score.

For the purpose of modeling hyper-rational golf course behavior, the concept of *homo economicus* will prove very useful. Imagine, every Sat-

---

[2]For a definitions of golf terms, refer to the first page of the appendix.

[3]I focus on tee-to-green strategy in this research and assume, for all intents and purposes, that putting does not exist.

[4] *Golfs Grand Design*, PBS.

urday morning homo economicus plays a round of golf. This paper attempts to characterize the way he makes decisions. Henceforth, whenever I refer to the player or the golfer, assume that the individual playing golf is homo economicus.

In Section II, I discuss the incomplete nature of the previous statistical research done by Mark Broadie, arguing that, for all its strengths, he lacks of a theoretical basis for understanding how golf is played.

One might argue that golf behavior could be rational if the player were maximizing his golf utility function. In Section III, I explain why utility is important and argue that this particular baseline rationality model ought to be developed before exploring other drivers of a representative golfers utility.

In Section IV, I create a function which defines the distance the player is able to hit each club in the set. I also discuss the assumptions which I make in the process. Section V creates a model of how the golfer will make decisions if he has perfect control over his shots. This captures the essence of the rational decision making process. In Section VI, I dispense with the perfects shot-control model and develop a probabilistic framework for uncertain shot outcomes, discuss the assumptions I make for simplification, and show that the choice of strategy is different for the two very similar scenarios captured in **Figure 3** I explain the motivation for this shift the context of **Figure 4**. In Section VII, I augment the model in the previous section by adding creating a function for the uncertainty of the outcomes for shots played from bad positions. I incorporate this into the expected score equation to identify how the players new strategy choice for the scenario in **Figure 3**. Section VIII discusses my findings.

## II. All evidence but no theory

In a series of papers[5], Mark Broadie, professor of business and finance at Columbia, used the PGA Tours ShotLink data[6] and quantitative methods to develop the strokes gained approach in order to compare player proficiencies in many individual aspects of the game of golf. This new

---

[5]See references for a list.

[6]ShotLink Intelligence, PGATour.com. (Link)

method of measuring performance has become broadly popular in golf analysis and has changed the way golfers skills are assessed. Broadies excellent new book, Every Shot Counts: Using the Revolutionary Strokes Gained Approach to Improve Your Golf Performance and Strategy, details his insights on lower scoring from the ShotLink data and explains his strokes gained methods. The book also gives recommendations to amateur golfers on how to make better strategic decisions on the golf course. Broadies formulation of the strokes gained statistical method constitutes an important advance for golf analysis as well as for the relative comparison of the different golf skills between PGA Tour professionals and I find his results and methods persuasive.

One aspect of Broadies book troubles me, however. In Chapter Eight of his book, Tee-to-Green Strategy: How Data and Optimization Can Lower Your Score he uses his insights about professional golf to attempt to coach the amateur golfer through several strategic situations so you can learn how to develop a game plan to shoot lower scores.[7] Broadie can identify tee-to-green strategies which have the highest likelihood for success and discusses how amateur golfers might employ them. Previously, Broadie discusses his use of the ShotLink dataset in his simulation models. These suggestions are well supported and I do not necessarily disagree with his suggestions but I find this approach problematic because all of his suggestions rely on insights from data on the actions and outcomes of golf shots hit by professionals who play at the highest level of the game. Broadie *does* acknowledge the tremendous heterogeneity of ability and playing style that exists among golfers and asserts that the choice of an optimal strategy depends on a players individual shot pattern[8].

Neither in his research papers nor in his book does Broadie offer a formal model of how golf is played. A more traditional and robust method of analysis begins with a model of agent behavior and then uses data to improve that model. Instead he uses data on holes have been played in the past to gauge which strategies have been most efficient. The lack of a model to explain the behavior in the data along with a focus a non-representative sample of observations matters a lot when it comes to understanding optimal strategic behavior and making strategy recommendations. To illustrate this consider a counterfactual case where Broadie collected and analyzed

---

[7]Broadie, M. 2014. Every Shot Counts, 162.

[8]In Every Shot Counts, Broadie defines a shot pattern as the range of outcomes that can happen on any golf shot. (pg. 163)

ShotLink-like data from amateur golfers and then tried to characterize the strategies that were most effective. It seems unlikely that he would reach the same conclusions and if he were to make recommendations on how professional golfers might improve their scoring based on this hypothetical data, the recommendations might be misleading. In neither case is he able characterize a generally optimal strategic approach to golf. What works for Tiger Woods, might not work for an average Sunday golfer. Thus the need for a model.

It must be said, however that this is not his primary research focus. What follows is an attempt to provide the beginnings of a general model which will make it possible to have a broader understanding of optimal strategic decision making and allow more insightful conclusions to be drawn from the data.

## III. Why choose a minimization of expected score approach?

Rational behavior in golf could be modeled using two different methods. I chose the method by which a player minimizes expected score given an array of strategies. It could be argued that a player maximizing his golf utility function would also be behaving rationally. This utility function would take into account golf preferences other than scoring. This assumes that a golfers behavior is motivated by more than just his raw score. While both could be equally valid, expected score minimization reflects a golfers desire to achieve the strict goal of the game playing the course in as few strokes as possible.

In no way do I assume that all golfers behave in this way. It is possible that some players behave in a way such that it is *as if* they are minimizing their expected score but for a vast majorities of golfers at all levels (including PGA Tour players) strategic golf course decisions do not reflect an expected-score minimization process of decision making. Future papers will substantially explore the drivers of golfers utility and the complexities of a golfers utility function. That said, the fundamental objective of golf is scoring and developing a standard model of rationality must start with the assumption that utility is driven simply by a desire to shoot the lowest score possible. To develop a utility maximization model which considers a golfers secondary objectives would be putting the cart before the horse.

I do not think that any type of rational behavior model will not be able to account for many of the inexplicable choices which can be systematically observed in the game but having both a model of expected-score minimization and a model of utility-maximization will necessary for examining players strategic decisions so as to properly identify and assess the reasoning-failures which cause these surprising strategic decisions to be made.

## IV. Defining golf club distances [9]

Let $club_j$ represent a unique club where $j$ can take discrete values between 1 and 13. **Table 1** in the appendix lists the common names for each $j$. The **natural** length (where $d_{std} = 0$) of a shot hit with $club_j$ is given by $Dist_j$. This is not fixed, however. At will, a player can adjust his swing in order to increase or decrease this displacement. If the club is not a wedge ($j = 12, 13$), he can add or subtract a distance ($d_{std}$) of up to 5 yards[10] to a shot without choosing a different club. The maximum and minimum displacements of shots played with $club_j$ are as follows:

$$Dist_j^{max} = Dist_j + 5$$

$$Dist_j^{min} = Dist_j - 5.$$

Thus, a player is able to hit $club_j$ any distance within the 10 yard interval, $[Dist_j^{min}, Dist_j^{max}]$.

The driver ($j = 1$) is always the longest club in the bag. That is:

$$Dist_{j=1} > Dist_{j=2...13}.$$

I use $Dist_{j=1}$ as reference point from which I can calculate the length of shots hit using every other club. For simplicity, I assume a distance

---

[9]Of the fourteen clubs which make up a set, the putter is irrelevant for the purposes of this paper.

[10]In keeping with standard practice in U.S. golf, all distances discussed in this paper are measured in yards.

function which is linear in $j$.[11] An increase from $club_j \to club_{j+1}$ corresponds to a 10 yard decrease in the shot-distance it can naturally achieve. For all $j \leq 11$, the yardage associated with $club_j$ can be given by the function, $D(j)$:

$$D(j) = 10(1 - j) + Dist_{j=1}, \tag{1}$$

where
$$D'(j) < 0; D''(j) = 0.$$

I am going to assume that $club_{j=12}$ and $club_{j=13}$ both wedge clubs but not a pitching wedge. Wedges cannot be hit a long way but they can be hit any distance less than their maximum distances. For wedges, these are the maximum and minimum distances:

$$Dist_{j=12}^{max} = Dist_{j=12} + 5 \qquad Dist_{j=12}^{min} > 0$$

$$Dist_{j=13}^{max} = Dist_{j=13} + 5 \qquad Dist_{j=13}^{min} > 0$$

The **Figure 1** illustrates how this model of golf club distances allows the golfer to hit any yardage in the range of $(0, Dist_{j=1}^{max}]$.

---

[11]This is a very useful assumption and I one which I think produces a fairly good approximation of reality.
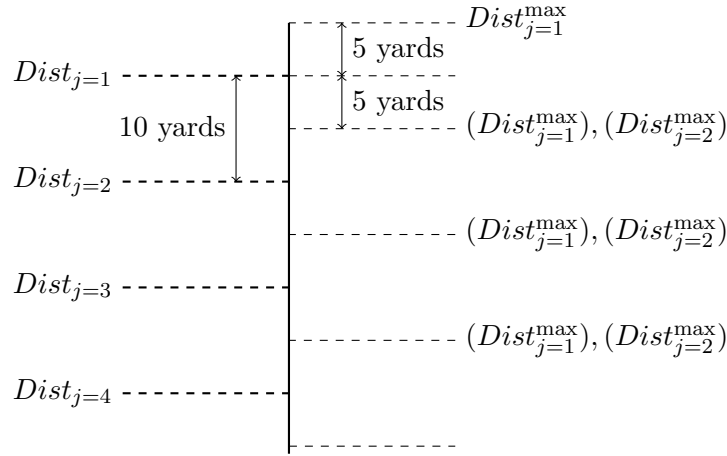
Figure 1: A diagram of natural, maximum and minimum distances for $j = 1, ..., 4$.

## V. Decisions with perfect ball-control

In this first model of strategic decision making, a representative golfer is able to hit every shot exactly where he wants it to go. **Figure 2** shows a standard golf hole.[12] In order to minimize his score, the player chooses which club to hit given his distance to the target. Commonly-used technology such as GPS and laser rangefinders have made it easy for the player to know his precise yardage to the target. Therefore, I assume that the will know the true yardage ($Y^{true}$) for his shots.

$Y^{true}$ must be adjusted, however, because the distance that a golf ball travels in the air can be affected by certain exogenous factors. Before each shot, the player must assess the relative elevation of the target, wind velocity, and wind direction.[13] Failing to do this increases the likelihood of a bad outcome and is inconsistent with score minimization.

---

[12] I chose to use a one-shot hole because my assumption of perfect ball-control, creates a situation where only approach shots are relevant. There is no decision to for a player to make if position does not affect the outcome of the approach shot.

[13] There are more factors which might affect the golf balls flight but these are the ones which must be reassessed before every shot.

The player estimates an adjusted yardage which reflects how far he actually needs to hit the ball for the best possible outcome.

$$Y_{(F_2)}^{adj} = Y_{(F_2)}^{true} + effect_{elevation} + effect_{wind}$$

This extremely simple model is important because it shows the general decision-making process. Rational behavior requires that the player consider the relevant factors before executing the golf shot. Though, here, I incorporate two factors which affect a shots distance. Going forward, I will assume that their effects do not exist and therefore, $Y_{(F_2)}^{adj} = Y_{(F_2)}^{true}$.
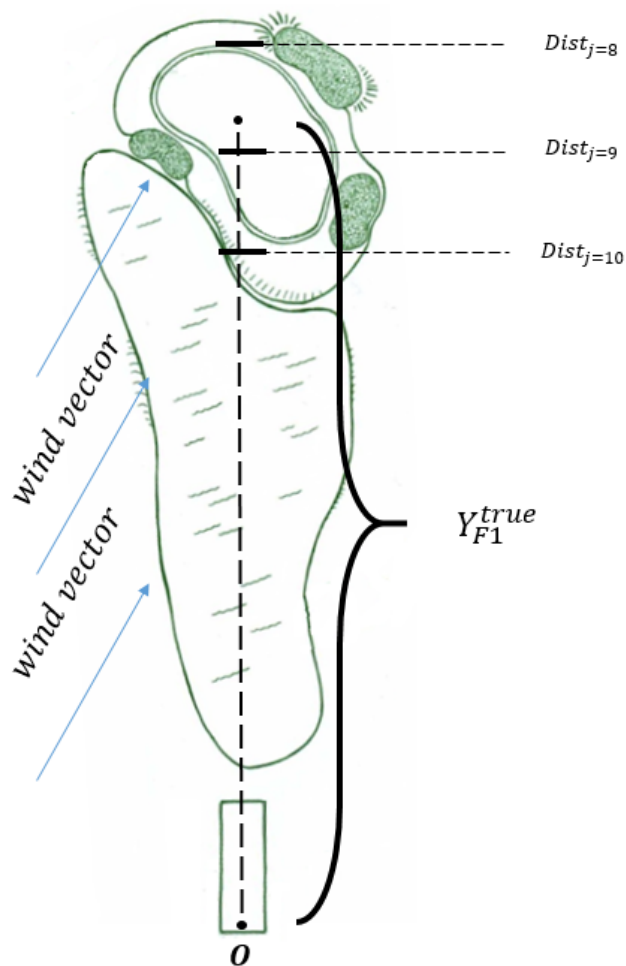
Figure 2: This hole, $F_2$ is a fairly generic 1-shot hole. It is within reach for the player with either $club_{j=9}$ or $club_{j=8}$. (Source: Byrdy 2005)]

**Note for this and all other hole diagrams**: All of the golf holes in the diagrams in this paper were taken from Stan Byrdys book, *Alister Mackenzies Masterpiece: The Augusta National Golf Club* as drawings of golf holes at Augusta National Golf Club. *They have been altered and adapted significantly for this paper and should be taken as generic golf holes rather than the hole which Byrdy originally intended to represent.*

## VI. Minimizing expected score with uncertain shot outcomes

VI.1    *Probabilities associated with good and bad positions*

In practice, no matter how much talent one has for the game of golf, it is impossible to control exactly where the ball is going to go. Mark Broadie correctly emphasizes shot pattern. He says when a golfer hits a shot, he observes only one outcome, but he needs to plan for the range of outcomes that could happen.

Simplistically, lets say that shots either end up in good positions or bad positions. I define a good position to be the fairway. Any position which is not the fairway is a bad position. I will abbreviate a position in the fairway as $\boldsymbol{p}^{good}$ and a position not in the fairway as $\boldsymbol{p}^{bad}$. The probability of a good position will vary inversely with the length of the shot. Expressed mathematically, I assert:

$$P(good|club_j) < P(good|club_{j+1}) < ... < P(good|club_{j=13}) \qquad (2)$$

Alternatively, I could show this by taking the probability of hitting a good shot as a function of $j$:

$$\text{P(good)} : j \rightarrow f(j).$$

I define the function $f$ as:

$$f(j) = (j - 1)\eta + P(good|club_{j=1}). \qquad (3)$$

Thus $\frac{d\text{P(good)}}{dj} = f'(j) > 0$ and $f''(j) = 0$.

In **Table A.2**, I illustrate how the golfers probability of hitting a good shot increases linearly in $j$. I use this functional form in my models

because it will help produce a more concrete result. **Figure A.1** graphs P(good) : $j$ where $\eta$ is a constant fraction ($0 \leq \eta < 1$). This shows how as the clubs natural distance become shorter, shot accuracy increases. **Table A.3** also models the probability of a good outcome given club selection. It has the same purpose as **Table A.2**, to capture how this probability increases as shot length decreases. This alternative probability function captures a more realistic effect: decreasing marginal accuracy gains over shot distance. This can be seen in **Figure A.2**. Exploring the impact which this will have on strategic decisions is an extension for future research. In this paper, for the sake of simplicity, I assume the first case: accuracy gains over shot distance are linear.

This matters because the players position after hitting a first shot influences his likelihood of hitting the second shot well hence being in a good position after he hits the second shot. If the ball is in the fairway, it is much easier to hit accurately than if it is in the rough. Holding $club_j$ constant,

$$P(good|\boldsymbol{p}^{good}) \gg P(good|\boldsymbol{p}^{bad}). \tag{4}$$

Not all bad positions are created equally, however. Some bad positions only require that the player hit out of long grass from an area which is not the fairway. This is called the rough. Henceforth, I will use $\boldsymbol{p}^{bad}_{rough}$ to symbolize these positions. Other bad positions might require the player to hit through a forest, from a sand trap, or add a 1-shot penalty to his score.

This has important implications for an agent attempting to minimize expected score $E[S]$. The probabilities of being in a good position after playing shots from different categories of bad position also vary.

$$P[good|\boldsymbol{p}^{bad}_{rough}] > P[good|\boldsymbol{p}^{bad}_{trees}]$$

$$P[good|\boldsymbol{p}^{bad}_{rough}] > P[good|\boldsymbol{p}^{bad}_{sand}] \tag{5}$$

12

## VI.2    *Minimization of expected score in* **Figure 3**, *Panel A*

     In **Figure 3**, $O$ is the origin or tee box from which the first shot is always played. A rational golfer will choose the strategy $(s_i)$ for this hole which minimizes $E[S|\boldsymbol{p}^O]$. The tee shot will be played to a position in the fairway according to the golfers choice of strategy, either $s_i = s_A$ or $s_i = s_B$. In *Panel A*, I amended the diagram by removing the bunkers on the left side of the hole. This has implications for the players strategy choice. A shot played from $\boldsymbol{p}^O$ using $club_{j=1}$ will come to rest at either in the fairway at Position A, $\boldsymbol{p}_{(}(s_A)^{good}$, with $P(good_O|club_{j=1} = P(good_{(s_A)}))$ or in the rough parallel with Position A, $(\boldsymbol{p}^{bad}_{rough}, s_A)$, with $P(bad_O|club_{j=1}) = P(bad_{(s_A)})$.
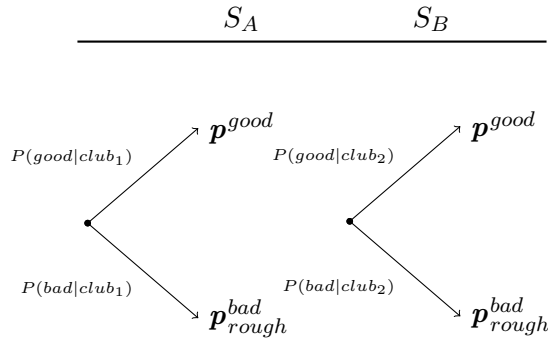
Note that:

$$P(good_{(s_A)}) + P(bad_{(s_A)}) = 1.$$

     A tee shot using $club_{j=2}$ will finish in the fairway at Position $B$, $\boldsymbol{p}^{good}_{(s_B)}$, with $P(good_O|club_{j=2}) = P(good_{(s_B)})$ or in the rough parallel with Position B, $[\boldsymbol{p}^{bad}_{rough}, s_B]$, with $P(bad_O|club_{j=2}) = P(bad_{(s_B)})$. Similarly,

$$P(good_{(s_B)}) + P(bad_{(s_B)}) = 1.$$

These are the outcomes which could be possible after the tee shot.
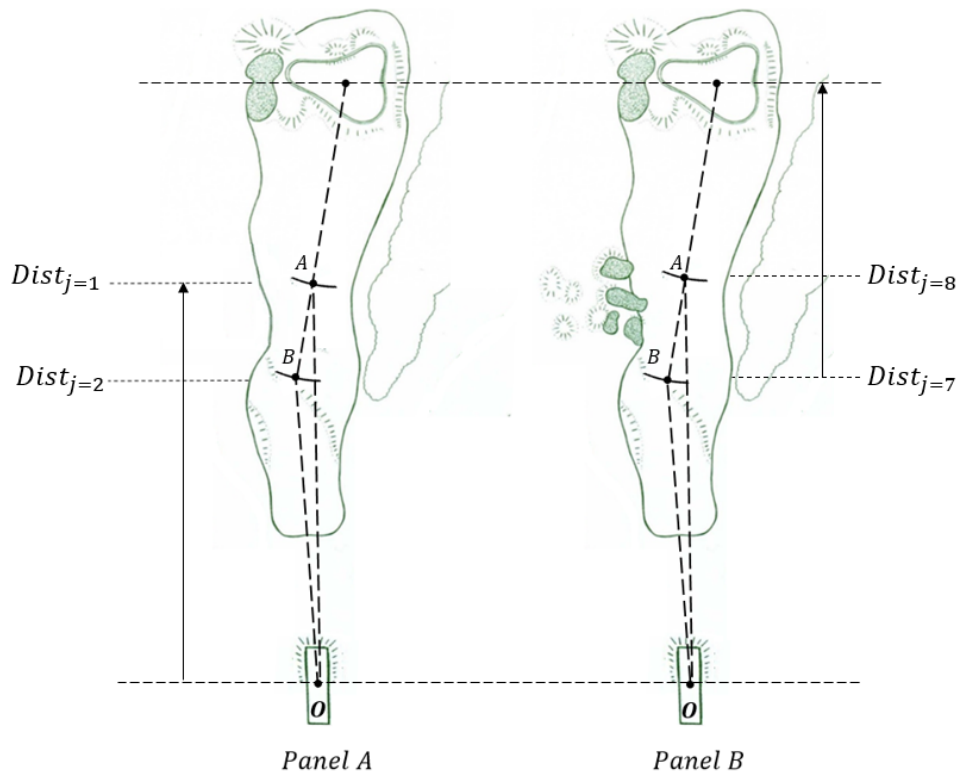
Figure 3: This hole, $F_3$ is a generic 2-shot hole. It requires a tee shot and then an approach shot. (Source: Byrdy (2005))

As I have already shown, $P(good_O|club_{j=1}) < P(good_O|club_{j=2})$. This implies that he will be in the fairway more often if he uses $club_{j=2}$ than if he uses $club_{j=1}$. It is important not to over-interpret this statement. Although the player will hit more fairways with $club_{j=2}$, it follows that he will need to hit a longer shot into the green. This is captured in both panels of **Figure 3**. If he chooses strategy, $s_i = s_A$, the golfer can expect to hit an approach shot with $club_{j=8}$. If he chooses $s_B$, he can expect to hit $club_{j=7}$ into the green. I am describing two effects which run in the opposite directions.

Effect 1  better chance of being in the fairway with $s_B$

$$P(\boldsymbol{p}^{good}|s_B) > P(\boldsymbol{p}^{good}|s_A) \tag{6}$$

Effect 2  better chance of hitting the green from the fairway at Position $A$

$$P(green|\boldsymbol{p}^{good}_{(s_B)}, club_{j=7}) < P(green|p^{good}_{(s_A)}, club_{(j=8)}) \tag{7}$$

At this point, it is impossible to know which effect will have a greater magnitude.

So far, I have looked at the potential scenarios which might arise if the players shot finishes in a good position. If it finishes in a bad spot, $p^{bad}$, the obstacles which the golfer might face given both strategies in Panel $A$ are the same:

$$\boldsymbol{p}^{bad}, s_A \to \boldsymbol{p}^{bad}_{rough} \quad and \quad \boldsymbol{p}^{bad}, s_B \to \boldsymbol{p}^{bad}_{rough}$$

This means that there are no obstacles which would be in play with one strategy but which would not be in play with the other.

All of this can be incorporated into an expected score equation. Note that in minimizing $E[S]$, he also includes his expected likelihood of

hitting the green with his second shot given his choice of strategy. Having strategy in this equation incorporates the players club choices and thus it also incorporates the expected club uncertainties. Assuming that he is able to reach the green in two shots regardless of strategy, the player chooses:

$$minE[S] = 1 + E\{P[hitting\ green\ with\ 2^{nd}\ shot|s_i]\} \quad where \quad s_i \in [A, B, ..., i]$$

For the scenario illustrated by **Figure 3**, *Panel A*, the equations for expected score are:

$$E[S|\boldsymbol{p}^O, s_A] = 1 + P(good_{(s_A)}) \cdot P(green|\boldsymbol{p}^{good}, s_A) + P(bad_{(s_A)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_A)$$

and

$$E[S|\boldsymbol{p}^O, s_B] = 1 + P(good_{(s_B)}) \cdot P(green|\boldsymbol{p}^{good}, s_B) + P(bad_{(s_B)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_B).$$

It follows that any advantage that the golfer gains by being in the fairway more often after choosing $s_i = s_B$ will be offset by his need to hit a longer club into the green. Therefore, this leads me to conclude that the expected scores given both strategies are equal and suggests that the player will be indifferent between the two strategies in *Panel A*.

$$P(good_{(s_A)}) \cdot P(green|\boldsymbol{p}^{good}, s_A) + P(bad_{(s_A)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_A)$$

$$=$$

$$P(good_{(s_B)}) \cdot P(green|\boldsymbol{p}^{good}, s_B) + P(bad_{(s_B)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_B)$$

$$\therefore \quad \underline{E[S|\boldsymbol{p}^O, s_A] = E[S|\boldsymbol{p}^O, s_B]}$$

This is a very interesting result. It shows that in the absence of fairway obstacles, the golfer gains no advantage by being in the fairway off the tee over hitting the ball longer. Under the assumptions of this model, the advantage of $\boldsymbol{p}^{good}_{(s_B)}$ is cancelled by the increased probability $P(green|\boldsymbol{p}^{good}, s_A)$.

### VI.3 *Minimization of expected score in* **Figure 3**, *Panel B*

It is common in the game of golf for obstacles to be placed in the vicinity of the landing area to punish inaccurate tee shots of a certain length. In these cases, different positions in the fairway bring into play unique combinations of challenges, penalties, or both. This is illustrated in *Panel B* of **Figure 3**. *Panel A* and *Panel B* are diagrams of the same hole but for the sand traps to the left of the fairway in *Panel B*. These sand traps or bunkers are in the range for $club_{j=1}$ but not for $club_{j=2}$.

This has important implications for the golfers choice of either $s_A$ or $s_B$. Given these two options, the possible outcomes after hitting a tee shot are different than they were in the previous example.
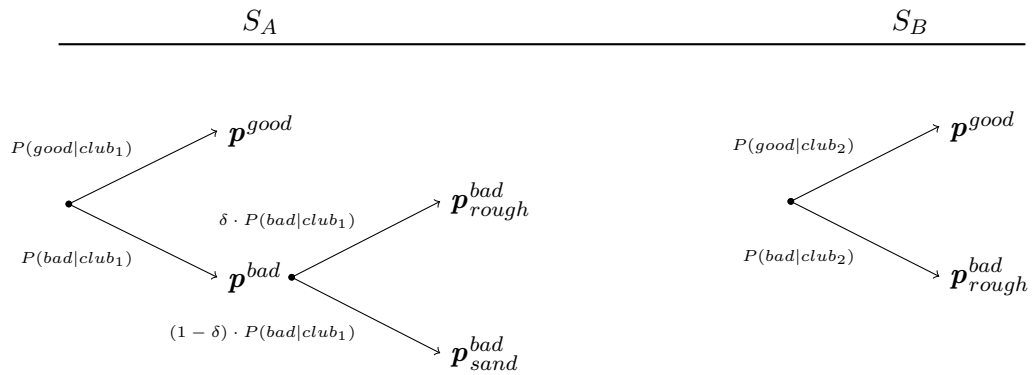


Here, the probability of hitting a bad shot with the driver is the same as before except, now, there is an additional bad potential outcome for an

inaccurate drive: the sand. The probability of hitting the green from the sand trap is lower than it would be from the rough. This is shown in **Eq. (5)**.

Since the probability that a player will hit a bad shot from $\boldsymbol{p}^O$ if he uses $club_{j=1}$ is $P(bad_O|club_{j=1}) = P(bad_{(s_A)})$, the ball will find the rough with the probability $\delta \cdot P(bad_{(s_A)})]$ and will find the sand with the probability $(1-\delta) \cdot P(bad_{(s_A)})$. This can be incorporated into the expected score equation as follows:

$$E[S|\boldsymbol{p}^O, s_A] =$$

$$= 1 + P(good_{(s_A)}) \cdot P(green|\boldsymbol{p}^{good}, s_A) +$$

$$+ \delta \cdot P(bad_{(s_A)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_A) + (1-\delta) \cdot P(bad_{(s_A)}) \cdot P(green|\boldsymbol{p}^{bad}_{sand}, s_A)$$

and

$$E[S|\boldsymbol{p}^O, s_B] =$$

$$= 1 + P(good_{(s_B)}) \cdot P(green|\boldsymbol{p}^{good}, s_B) + P(bad_{(s_B)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_B).$$

The players new minimization problem has to take this into account. When he chooses:

$$minE[S] = 1 + E[P(hitting\ green\ with\ 2^{nd}\ shot|s_i)] \quad where \quad s_i \in [A, B, ..., i]$$

In the **Eq. (6)** and **Eq. (7)**, I assert that when giving up distance for accuracy off the tee by playing $club_{j=2}$ rather than $club_{j=1}$, the player then faces increased inaccuracy on the approach shot because he needs to use $club_{j=7}$ rather than $club_{j=8}$. This still holds. Unlike in the previous example, however, the presence of the bunkers will change the magnitudes of both effects, thus they will no longer cancel.

$$\delta \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_A) + (1-\delta) \cdot P(green\boldsymbol{p}^{bad}_{sand}, s_A) < P(green|\boldsymbol{p}^{bad}_{rough}, s_A)$$

By this logic, the model predicts the following choice:

$$P(good_{(s_A)}) \cdot P(green|\boldsymbol{p}^{good}, s_A) + \delta \cdot P(bad_{(s_A)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_A) +$$

$$+(1 - \delta) \cdot P(bad_{(s_A)}) \cdot P(green|\boldsymbol{p}^{bad}_{sand}, s_A)$$

$$>$$

$$P(good_{(s_B)}) \cdot P(green|\boldsymbol{p}^{good}, s_B) + P(bad_{(s_B)}) \cdot P(green|\boldsymbol{p}^{bad}_{rough}, s_B)$$

and therefore, for the situation faced in *Panel B*, the player will choose strategy, $s_B$:

$$E[S|\boldsymbol{p}^O, s_A] > E[S|\boldsymbol{p}^O, s_B].$$

In words, the player will choose the strategy where he gives up yardage off the tee. Importantly, it seems that this is not in order to gain the increased accuracy of $club_{j=2}$ but rather because it eliminates the possibility of hitting the ball into the bunkers.

### VI.4 *Expected score minimization in* **Figure 4**

The scenario presented in **Figure 4** is nearly identical to the one presented in *Panel B* of **Figure 3**. The only difference is that, by hitting $club_{j=1}$, the player can avoid the bunker on the right while $club_{j=2}$ brings the sand into play. This can be seen clearly in *Panel B* of **Figure 4**. Without doing the calculations, it is safe to say that:

$$E[S|\boldsymbol{p}^O, s_A] < E[S|\boldsymbol{p}^O, s_B]$$

$$E[S|\boldsymbol{p}^O, s_A] < E[S|\boldsymbol{p}^O, s_B]$$

Though it may seem superfluous, I have included this example in order to underscore the implication from the previous example. The players choice of accuracy over distance in *Panel B* of **Figure 3** is seemingly driven entirely by his desire to eliminate $p_{sand}^{bad}$. This same motivation is responsible, here, for his choice of $s_A$.
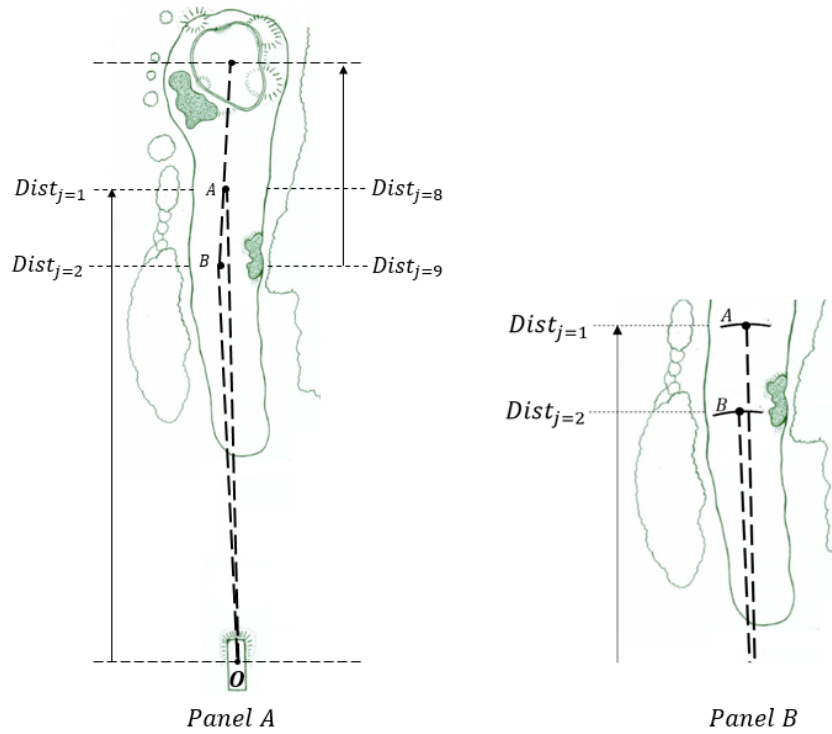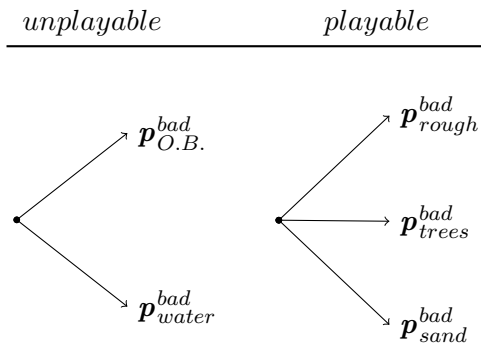


Figure 4: This hole, $F_4$ is a generic 2-shot hole. It requires a tee shot and then an approach shot. (Source: (Byrdy (2005)))

## VII. Alternate assumptions for shots played from bad positions

Throughout this body of research, I emphasize the fact that not all positions off the fairway are equally bad. First, I will separate the $\boldsymbol{p}^b ads$ into two categories, playable and unplayable. These two categories can be divided into subcategories as follows:

$$\underline{unplayable} \qquad\qquad \underline{playable}$$

$$\boldsymbol{p}^{bad}_{O.B.} \qquad \boldsymbol{p}^{bad}_{rough}$$

$$\boldsymbol{p}^{bad}_{trees}$$

$$\boldsymbol{p}^{bad}_{water} \qquad \boldsymbol{p}^{bad}_{sand}$$

Though it is true that under very unusual circumstances there are positions in the water which can be playable, and positions in the rough or trees which can unplayable, I will assume that these exceptions do not exist for the purposes of this research. Hitting a ball into the water or out-of-bounds adds a penalty to the players score but allows him to play his next shot from positions which are generally favorable. Shots which must be played from the rough, trees, or sand are subject to a different penalty.

Understanding this penalty requires me to briefly explain a little bit of golf reality. It is much easier for a golfer to hit a short club out of a bad position than it is for him to hit a long club. In good positions, small imperfections in the players golf swing can still produce a good result. This is not true if the position is bad. Here, the margin-for-error in a players swing becomes smaller as the attempted shot becomes longer. This is an effect which I think ought to be present in my models.

Keeping **Eq. (4)** and **Eq. (2)** in mind given that the player is in the rough, I add the fraction, $\mu$ to further diminish the probability of hitting

a good shot as $j$ decreases. Again, I am going to make this effect linear for the sake of simplicity. The details of this fraction, $\mu$, can be seen in **Table A.4** and **Figure A.3** and are defined by the following equation:

$$\text{P(good|club}_j, \ \mathbf{p}_{rough}^{bad}) : j \to \mu(j-2) + P(good|club_{j=2}, \ \mathbf{p}_{rough}^{bad}), \qquad (8)$$

$$where \quad j \in [2,11]$$

In **Eq. (5)**, I said that $P(good|\boldsymbol{p}_{rough}^{bad}) > P(good|\boldsymbol{p}_{sand}^{bad})$. I capture this effect by augmenting $\mu$ with $\varepsilon$, a fraction which is also described in Table A.4 is defined by Eq. (9).

$$\text{P(good|club}_j, \ \mathbf{p}_{sand}^{bad}) : j \to [\mu(j-2) - \varepsilon(j-1) + P(good|club_{j=2}, \boldsymbol{p}_{sand}^{bad}), \ (9)$$

$$where \quad j \in [2,11]$$

VII.1
*Incorporating $\mu$ into the expected-score minimization for* **Fig. 3**, *Panel A*

In a previous example, I show that the golfer is indifferent between $s_A$ and $s_B$. This is true under the assumptions which I made. If I incorporating the concept that the probability of hitting a good shot from a bad position decreases as the length of the attempted shot increases, I expect this result to change. I previously concluded that for the scenario in *Panel A*,

$$\boldsymbol{p}^{bad}, s_A \to \boldsymbol{p}_{rough}^{bad} \quad and \quad \boldsymbol{p}^{bad}, s_B \to \boldsymbol{p}_{rough}^{bad}.$$

While this continues to signify that there are no obstacles which would be in play with one strategy but which would not be in play with the other.

The statement does not imply, however, that the potential downsides are equal. I can now be more specific using **Eq. (8)**. ($\boldsymbol{p}^{bad}$, $s_A$) implies the use of $club_{j=8}$ while ($\boldsymbol{p}^{bad}$, $s_B$) will require the use of $club_{j=7}$

$$P(good|club_{j=8}, \boldsymbol{p}^{bad}_{rough}) = \mu + P(goodclub_{j=7}, \boldsymbol{p}^{bad}_{rough})$$

If I take this into consideration in the players expected-score minimization equation,

$$E[\boldsymbol{p}^O, s_A] = 1 + P(good_{(s_A)}) \cdot P(green|\boldsymbol{p}^{good}, s_A) + P(bad_{(s_A)} \cdot \overbrace{P(green|\boldsymbol{p}^{bad}_{rough}, s_A)}^{P(good|club_{j=8}, \boldsymbol{p}^{bad}_{rough})}$$

and

$$E[\boldsymbol{p}^O, s_B] = 1 + P(good_{(s_B)}) \cdot P(green|\boldsymbol{p}^{good}, s_B) + P(bad_{(s_B)} \cdot \overbrace{P(green|\boldsymbol{p}^{bad}_{rough}, s_B)}^{P(good|club_{j=7}, \boldsymbol{p}^{bad}_{rough})}$$

Therefore:

$$P(good_{(s_A)}) \cdot P(green|\boldsymbol{p}^{good}, s_A) + P(bad_{(s_A)}) \cdot P(good|club_{j=8}, \boldsymbol{p}^{bad}_{rough})$$

$$\geq$$

$$P(good_{(s_B)}) \cdot P(green|p^{good}, s_B) + P(bad_{(s_B)}) \cdot P(good|club_{j=7}, \boldsymbol{p}^{bad}_{rough})$$

I have increased the models connection with reality by augmenting it with this new effect $\mu$. It now predicts:

$$\therefore E[S|\boldsymbol{p}^O, s_A] > E[S|\boldsymbol{p}^O, s_B]$$

The scenario in *Panel A* of **Figure 3**, an expected-score minimizer chooses a strategy which delivers greater distance over additional accuracy.

## VIII. Discussion

This paper is able to make a few helpful conclusions. In Section III, I show that correct club choice is an essential element of rational behavior. I also show that the expected-score minimizer adjusts for exogenous factors such as before every shot. These predictions may be obvious, but there is a more subtle point, here.

(**a**) I show that the golfer must adjust his yardage so as to account for how certain factors will affect his specific shot;

(**b**) The rational golfer will be able to choose the correct club for the shot at hand;

(**c**) Every shot is independent because (a) and (b) are different for every shot a player might face;

(**d**) Rational decisions require (a) and (b) for every shot.

(**e**) Therefore, I can characterize the golfers general strategic process in the following statement:

*When behaving rationally, the golfer will treat every shot as independent of the previous one. He, therefore, will select a new strategy to minimize expected-score before each shot.*

In Section IV, I show that on a generic two-shot hole, where bad positions are not differentiated, the golfer will be indifferent between a conservative strategy from the tee which increases his probability of hitting the fairway and an aggressive strategy which improves the prospects for his second shot, given that he hits a good tee shot. I then differentiate bad positions so that a bad shot might find either rough or sand. Assuming that it is more difficult to hit from the sand than from the rough, the golfer will minimize expected score by choosing a strategy such that the worst outcome (the sand) is not within the range. He does this by selecting a club which will either travel beyond the sand or stay short of it.

The golfers indifference between the two strategies in *Panel A* of **Figure 3**, as predicted by the model in Section IV, is not entirely in line with what I anecdotally observe in watching people play golf. Rarely does the golfer choose 3-wood over driver for his tee shots. Once I factor in the $\mu$ effect, the model now predicts a robust preference for distance over accuracy. This is the most important prediction of this paper because it provides a rational justification for the driver-off-the-tee strategy which is so easily to observe in reality.

**REFERENCES:**

**Broadie**, **Mark**. 2008. Assessing Golfer Performance Using Golfmetrics, in *Science and Golf V: Proceedings of the 2008 World Scientific Congress of Golf*, eds. Debbie Crews and Rafer Lutz, 253-226. Mesa, AZ: Energy in Motion Inc. (Link)

**Broadie**, **Mark**. 2011. Putts Gained Measuring Putting on the PGA Tour," Working Paper. Unpublished. (Link)

**Broadie**, **Mark**. Assessing Golfer Performance on the PGA Tour," *Interfaces*, 42(2): 146-165. (Link)

**Broadie**, **Mark**. 2014. *Every Shot Counts: Using the Revolutionary Strokes Gained Approach to Improve Your Golf Performance and Strategy*, New York, NY: Gotham. (Website)

**Broadie**, **Mark** and **Soonmin Ko**. 2009. "A Simulation Model to Analyze the Impact of Distance and Direction on Golf Scores," in *Proceedings of the 2009 Winter Simulation Conference*, eds. Manuel Rossetti, Raymond R. Hill, Bjrn Johansson, Ann Dunkin, and Ricki G. Ingalls, 3109-3120. (Link) [see also Ko, S. 2012. Doctoral Dissertation (Link)]

**Byrdy**, **Stan**. 2005. *Alister Mackenzies Masterpiece: The Augusta National Golf Club*. Ann Arbor, MI: Sports Media Group.

*Golfs Grand Design*. PBS Distribution, 2012. Film. (Website)

## APPENDIX

<u>Definition of Golf Terminology</u> [14]

**Golf Club**  This is the implement used in playing the game. A club is assembled by gluing a small metal head to the end of a long, thin shaft made of graphite or steel. A player carries fourteen clubs which vary in loft and in length. A specific combination of these variables allows the player to hit the golf ball a given distance.

***Rules of Golf***  A book of rules which govern the playing of the game. Golf played in the United States and Mexico is governed by a version of the rules written by the United States Golf. Association. Golf played anywhere else in the world follows the rules as specified by The Royal and Ancient Golf Club of St. Andrews.

**Shot**  same as stroke; one swing of the club made with the intention of hitting the ball

**Tee-to-green strategy**  A unique set of steps for moving the ball from the tee to the green.

**The tee**  A small patch of ground where the golfer begins playing a hole by hitting his drive.

**Drive**  The first shot hit on any hole. Generally, a shot which is played from the tee to the fairway.

**Landing area**  the region of the fairway where most players will likely to hit their tee shots

**Fairway**  An intermediate area of very short grass which guides the player to the green. The golfer intends for his drive to settle in the fairway thus providing him the best opportunity to reach the green.

---

[14]Although I will give more formal definitions, Robin Williams's 2-minute sketch mocking the game of golf in which he explains the objective and most of these terms far more effectively than I. (Link)

**Approach shot**  A shot played either from the fairway or the rough which has the green as its immediate target. The green  A circular area of extremely short grass which surrounds the cup.

**A Putt**  a golf shot made with a putter; the type of stroke used to advance the ball towards the hole once the ball is already on the green

**Natural distance of a club**  The distance traveled by a ball hit solidly with a specific club by a player using an ordinary golf swing (no effort to hit it extra hard or soft), given perfect weather conditions.

**Out-of-bounds**  an area where a player might hit the ball but from where he is not allowed to play it, usually because it lies outside of a golf courses property line. The player must replay a shot which settles OB from the original position using his next shot. An additional one-stroke penalty is incurred.

**Sand trap, Bunker**  a defined area of sand on a golf course where a ball may come to rest; an obstacle out of which it is difficult for a golfer to hit a good shot

**Table A.1**

| $Club_j$ | Common Name | Distance (yrds) | Max, Min Distance |
|---|---|---|---|
| $j=1$ | Driver | $Dist_{j=1}$ | $Dist_{j=1}^{max} = Dist_{j=1} + 5$ <br> $Dist_{j=1}^{min} = Dist_{j=1} - 5$ |
| $j=2$ | wood club 1 | $Dist_{j=1} - 10$ | $Dist_{j=2} \pm 5$ |
| $j=3$ | wood club 2 | $Dist_{j=1} - 20$ | $Dist_{j=3} \pm 5$ |
| $j=3$ | hybrid club 1 | $Dist_{j=1} - 20$ | $Dist_{j=3} \pm 5$ |
| $j=3$ | 2-iron | $Dist_{j=1} - 20$ | $Dist_{j=3} \pm 5$ |
| $j=4$ | wood club 3 | $Dist_{j=1} - 30$ | $Dist_{j=4} \pm 5$ |
| $j=4$ | hybrid club 2 | $Dist_{j=1} - 30$ | $Dist_{j=4} \pm 5$ |

**Table A.1 cont.**

| $Club_j$ | Common Name | Distance (yrds) | Max, Min Distance |
|---|---|---|---|
| $j = 4$ | hybrid club 3 | $Dist_{j=1} - 30$ | $Dist_{j=4}^{max} = Dist_{j=4} + 5$ <br> $Dist_{j=4}^{min} = Dist_{j=4} - 5$ |
| $j = 5$ | 3-iron | $Dist_{j=1} - 40$ | $Dist_{j=5} \pm 5$ |
| $j = 5$ | hybrid club 4 | $Dist_{j=1} - 40$ | $Dist_{j=5} \pm 5$ |
| $j = 5$ | 4-iron | $Dist_{j=1} - 40$ | $Dist_{j=6} \; pm5$ |
| $j = 6$ | 5-iron | $Dist_{j=1} - 50$ | $Dist_{j=6} \pm 5$ |
| $j = 7$ | 6-iron | $Dist_{j=1} - 60$ | $Dist_{j=7} \pm 5$ |
| $j = 8$ | 7-iron | $Dist_{j=1} - 70$ | $Dist_{j=8} \pm 5$ |

**Table A.1 cont.**

| $Club_j$ | Common Name | Distance (yrds) | Max, Min Distance |
|---|---|---|---|
| $j = 9$ | 8-iron | $Dist_{j=1} - 80$ | $Dist_{j=9}^{max} = Dist_{j=9} + 5$ <br> $Dist_{j=8}^{min} = Dist_{j=9} - 5$ |
| $j = 10$ | 9-iron | $Dist_{j=1} - 90$ | $Dist_{j=10} \pm 5$ |
| $j = 11$ | pitching wedge | $Dist_{j=1} - 100$ | $Dist_{j=11} \pm 5$ |
| $j = 12$ | sand wedge | $Dist_{j=1} - 110$ | $Dist_{j=12}^{max} = Dist_{j=12} + 5$ <br> $Dist_{j=12}^{min} > 0$ |
| $j = 13$ | lob wedge | $Dist_{j=1} - 120$ | $Dist_{j=13}^{max} = Dist_{j=13} + 5$ <br> $Dist_{j=13}^{min} > 0$ |

Table A.1: **Golf Club Distances**: The commonly recognized golf club names which are associated with each value of $j$. (Note: The composition of a club set can vary but significantly. This composition does not affect the $Dist_j$s. In this model and also in practice, players choose between woods, hybrids, and irons based on which they find easiest to hit.)

**Table A.2**

| $Club_j$ | **Category** | $P(good|club_j)$ | $P(good|Club_j) : j \rightarrow$ $\eta(j-1) + P(good|club_{j=1})$ |
|---|---|---|---|
| $j = 1$ | Driver | $P(good|club_{j=1})$ | $P(good|club_{j=1})$ |
| $j = 2$ | 3-wood | $P(good|club_{j=2})$ | $\eta + P(good|club_{j=1})$ |
| $j = 3$ | 5-wood | $P(good|club_{j=3})$ | $2\eta + P(good|club_{j=1})$ |
| $j = 4$ | 3-iron | $P(good|club_{j=4})$ | $3\eta + P(good|club_{j=1})$ |
| $j = 5$ | 4-iron | $P(good|club_{j=5})$ | $4\eta + P(good|club_{j=1})$ |
| $j = 6$ | 5-iron | $P(good|club_{j=6})$ | $5\eta + P(good|club_{j=1})$ |
| $j = 7$ | 6-iron | $P(good|club_{j=7})$ | $6\eta + P(good|club_{j=1})$ |
| $j = 8$ | 7-iron | $P(good|club_{j=8})$ | $7\eta + P(good|club_{j=1})$ |

**Table A.2 cont.**

| $Club_j$ | **Category** | $P(good|club_j)$ | $P(good|Club_j) : j \to$ $\eta(j-1) + P(good|club_{j=1})$ |
|---|---|---|---|
| $j = 9$ | 8-iron | $P(good|club_{j=9})$ | $8\eta + P(good|club_{j=1})$ |
| $j = 10$ | 9-iron | $P(good|club_{j=10})$ | $9\eta + P(good|club_{j=1})$ |
| $j = 11$ | wedge | $P(good|club_{j=11})$ | $10\eta + P(good|club_{j=1})$ |
| $j = 12$ | wedge | $P(good|club_{j=12})$ | $10\eta + P(good|club_{j=1})$ |
| $j = 13$ | wedge | $P(good|club_{j=13})$ | $10\eta + P(good|club_{j=1})$ |

Table A.2: The probability of hitting a good shot by club, given as a linear function of $j$, illustrates that as the clubs natural distance become shorter, shot accuracy increases. $\eta$ is a constant fraction.     $(0 \le \eta < 1)$
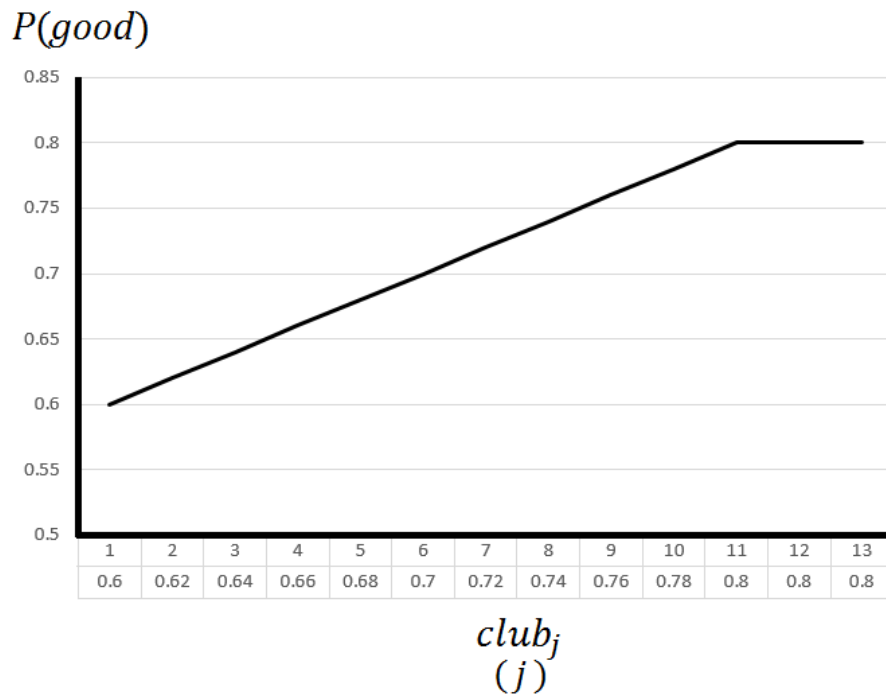
$P(good)$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0.6 | 0.62 | 0.64 | 0.66 | 0.68 | 0.7 | 0.72 | 0.74 | 0.76 | 0.78 | 0.8 | 0.8 | 0.8 |

$club_j$

$(j)$

Figure A.1: $P(good|club_j)$ **Note**: To construct this graph, I choose $P(good|club_{j=1}) = 0.60$ and take $\eta = 0.02$.

34

**Table A.3**

| $Club_j$ | Category | $P(good|club_j)$ | $P(good|Club_j) : j \rightarrow$ $\tau(j-1) + P(good|club_{j=1})$ |
|---|---|---|---|
| j=1 | Driver | $P(good|club_{j=1})$ | - |
| j=2 | fairway wood | $P(good|club_{j=2})$ | $\tau + P(good|club_{j=1})$ |
| j=3 | fairway wood | $P(good|club_{j=3})$ | $(\tau + 20\tau^2) + P(good|club_{j=1})$ |
| j=3 | hybrid club | $P(good|club_{j=3})$ | $(2\tau - 10\tau^2) + P(good|club_{j=1})$ |
| j=4 | hybrid club | $P(good|club_{j=4})$ | $2\tau + P(good|club_{j=1})$ |
| j=4 | long-iron | $P(good|club_{j=4})$ | $(3\tau - 20\tau^2) + P(good|club_{j=1})$ |
| j=5 | long-iron | $P(good|club_{j=5})$ | $3\tau + P(good|club_{j=1})$ |
| j=6 | long-iron | $P(good|club_{j=6})$ | $(3\tau + 10\tau^2) + P(good|club_{j=1})$ |
| j=7 | mid-iron | $P(good|club_{j=7})$ | $(4\tau - 10\tau^2) + P(good|club_{j=1})$ |

**Table A.3 cont.**

| j=8 | mid-iron | $P(good\|club_{j=8})$ | $4\tau + P(good\|club_{j=1})$ |
|---|---|---|---|
| j=9 | mid-iron | $P(good\|club_{j=9})$ | $(4\tau + 5\tau^2) + P(good\|club_{j=1})$ |
| j=10 | 9-iron | $P(good\|club_{j=10})$ | $(4\tau + 10\tau^2) + P(good\|club_{j=1})$ |
| j=11 | wedge | $P(good\|club_{j=11})$ | $(4\tau + 15\tau^2) + P(good\|club_{j=1})$ |
| j=12 | wedge | $P(good\|club_{j=12})$ | $(4\tau + 15\tau^2) + P(good\|club_{j=1})$ |
| j=13 | wedge | $P(good\|club_{j=13})$ | $(4\tau + 15\tau^2) + P(good\|club_{j=1})$ |

Table A.3: The probabilities of hitting a good shot with each club in a form which exhibit **decreasing returns** to distance. $\tau$ is a constant fraction. $(0 \leq \tau < 1)$
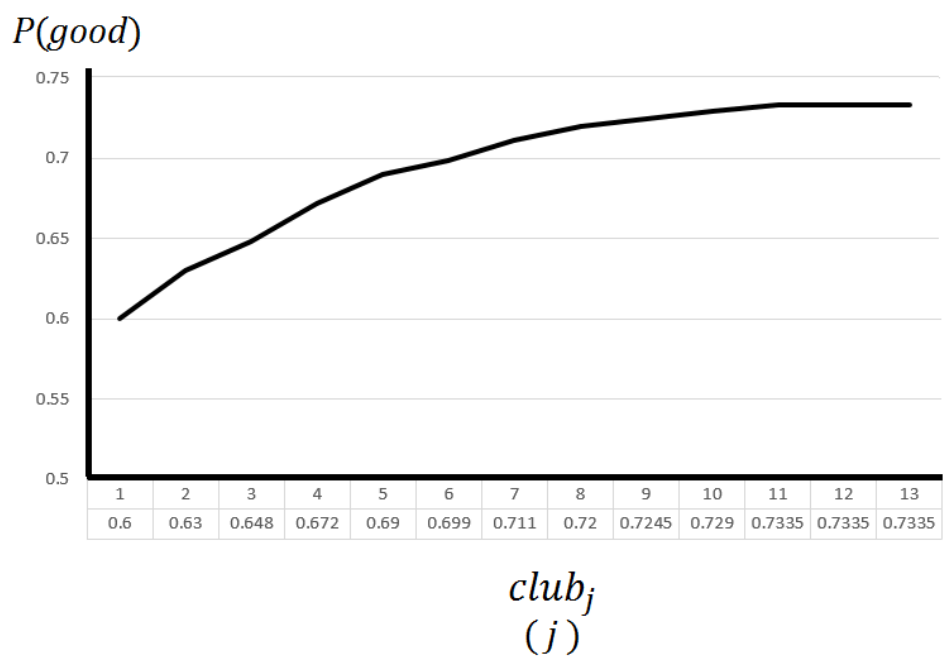
Figure A.2: $P(good club_j)$ **Note:** To construct this graph, I choose $P(good|club_{j=1}) = 0.60$ and take $\tau = 0.03$.

**Table A.4**

| $Club_j$ | $P(good)$ | $P(good\|p^{bad}_{rough})$ | $P(good\|p^{bad}_{sand})$ |
|---|---|---|---|
| $j = 2$ | $P(good\|club_{j=2})$ | $P(good\|club_{j=2})$ | $-\varepsilon + P(good\|club_{j=2})$ |
| $j = 3$ | $P(good\|club_{j=3})$ | $\mu + P(good\|club_{j=2})$ | $-2\varepsilon + \mu + P(good\|club_{j=2})$ |
| $j = 4$ | $P(good\|club_{j=4})$ | $2\mu + P(good\|club_{j=2})$ | $-3\varepsilon + 2\mu + P(good\|club_{j=2})$ |
| $j = 5$ | $P(good\|club_{j=5})$ | $3\mu + P(good\|club_{j=2})$ | $-4\varepsilon + 3\mu + P(good\|club_{j=2})$ |
| $j = 6$ | $P(good\|club_{j=6})$ | $4\mu + P(good\|club_{j=2})$ | $-5\varepsilon + 4\mu + P(good\|club_{j=2})$ |
| $j = 7$ | $P(good\|club_{j=7})$ | $5\mu + P(good\|club_{j=2})$ | $-6\varepsilon + 5\mu + P(good\|club_{j=2})$ |
| $j = 8$ | $P(good\|club_{j=8})$ | $6\mu + P(good\|club_{j=2})$ | $-7\varepsilon + 6\mu + P(good\|club_{j=2})$ |
| $j = 9$ | $P(good\|club_{j=9})$ | $7\mu + P(good\|club_{j=2})$ | $-8\varepsilon + 7\mu + P(good\|club_{j=2})$ |
| $j = 10$ | $P(good\|club_{j=10})$ | $8\mu + P(good\|club_{j=2})$ | $-9\varepsilon + 8\mu + P(good\|club_{j=2})$ |
| $j = 11$ | $P(good\|club_{j=11})$ | $9\mu + P(good\|club_{j=2})$ | $-10\varepsilon + 9\mu + P(good\|club_{j=1})$ |

Table A.4: The probabilities of hitting a good shot from a bad position with each club as a linear functions of $j$. $\mu$ and $\varepsilon$ are constant fractions. Note: (a) $club_{j=1}$ is never used in bad positions. $(0 \leq \mu,\ \varepsilon < 1)$
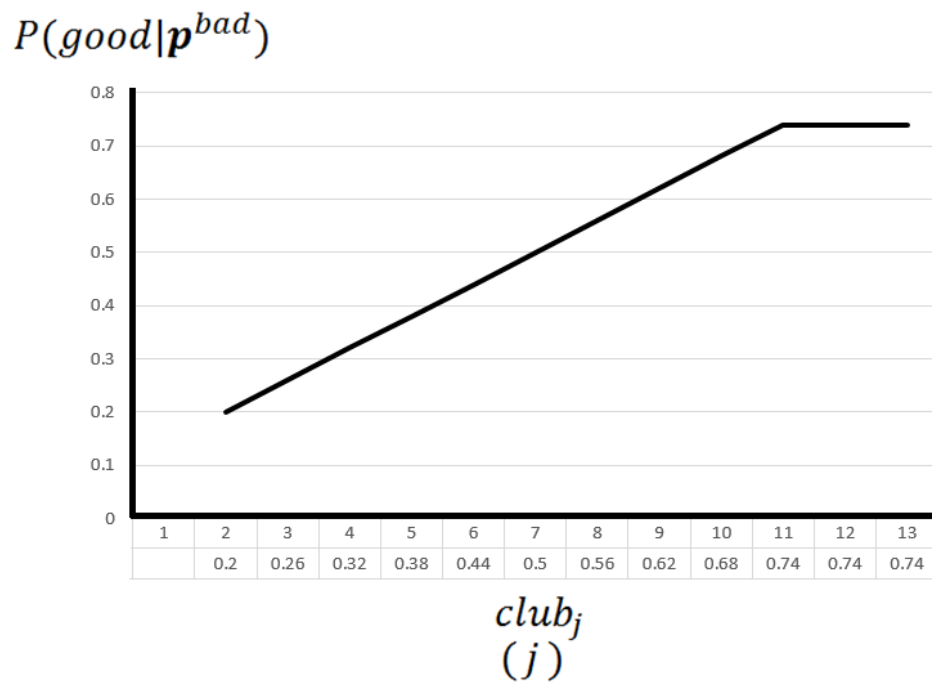
$P(good|\boldsymbol{p}^{bad})$



| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.2 | 0.26 | 0.32 | 0.38 | 0.44 | 0.5 | 0.56 | 0.62 | 0.68 | 0.74 | 0.74 | 0.74 |

$club_j$
$(j)$

Figure A.3: $P(good|\boldsymbol{p}^{bad},\ club_j)$   **Note**: To construct this graph, I choose $P(good|\boldsymbol{p}^{bad},\ club_{j=2}) = 0.20$ and take $\mu = 0.06$.