THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF RISK MANAGEMENT

Using Strokes Gained on the PGA Tour to Analyze U.S. Ryder Cup Performance

WILLIAM GERITY
SPRING 2024

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Actuarial Science
with honors in Actuarial Science

Reviewed and approved* by the following:

Andrew Wiesner
Teaching Professor of Statistics
Thesis Supervisor

Amanda Hammell
Professor in Practice of Risk Management
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

This thesis looks at one of the only team tournaments in men's professional golf, the Ryder Cup, and the strategy behind selecting and organizing the players on the team. With the tournament having the opportunity for a team captain to choose golfers, there is undoubtedly many ways to come up with the best lineup. In the past, some methods have been more qualitative such as how a player's game looks currently, or if they will have chemistry with the other players on the team. However, since the introduction of strokes gained statistics, there is more opportunity to quantitatively choose the optimal lineup of players.

Specifically, I explored the strokes gained metric for measuring players' performance on the PGA Tour, and the application of new methods to analyze performance in the Ryder Cup. I looked at the strokes gained statistics by category, to quantitatively find the best Ryder Cup lineup based the players' performance. If this study proves to be successful, then it could promote further methodologies for Ryder Cup player selection and potentially be used in future Ryder Cups.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my honors advisor Amanda Hammell for her continued support throughout the long thesis writing process. I would also like to express my appreciation to my thesis supervisor Andrew Wiesner for his valuable insight and guidance.

Chapter 1

Introduction

In the modern age of professional sports, there are increasingly more statistics available for analysis. In the NFL, the introduction of *Quarterback Rating* in 2011 has completely changed the way quarterbacks are analyzed, replacing the previous *passer rating* metric (Burke and Katz, 2016). In other sports such as basketball and baseball, more advanced statistical performance measures have been in place even longer. However, professional golf has not been as quick to create and adopt new advanced statistics. The traditional way of tracking players' performance in golf has always been a combination of driving distance, fairway hit percentage, green in regulation (GIR) percentage, and number of putts per round/hole. These metrics may seem fairly comprehensive when tracking all parts of a player's golf game, but ultimately fail to capture the whole picture. For example, one way in which putts per round is flawed is when a player chips their ball into the hole from off the green. When just using this statistic, it should show that the player had zero putts for that hole and subsequently lower their total putts per round, making it seem as though that player's putting was exceptional. However, in reality that player's putting was not exceptional on that hole they chipped in, because they did not even putt. This type of statistic was one of many motivating factors that lead to the creation of strokes gained, a revolutionary new way to measure a golfer's performance.

The concept of gaining strokes was first being researched in the 1990's and early 2000's by Peter Sanders, but it was widely popularized by Columbia professor Mark Broadie (Sanders). Strokes gained (SG) is essentially a way to compare a player's performance to a benchmark.

Broadie (as cited by Golf Channel, 2023) explained in an interview that; in finance, just saying that a stock went up 10% sounds great, but what if the market average went up 25% in that same time period. This example shows the importance of having a benchmark, because in golf it works the same way where it is much more useful to compare your performance to the average performance of someone with similar skills as you. This means that the benchmark for professional golfers on the PGA Tour, and the one that Broadie uses in his calculations, is the tour average. Around the same time that strokes gained were being introduced, the PGA Tour was implementing ShotLink, which is “laser technology to track where every shot starts and where every shot finishes” (Chupaska, 2021). In utilizing this new way of tracking every single shot that is hit in every PGA Tour event, strokes gained started to gain traction and eventually be used by the PGA Tour for daily use. The way strokes gained is actually calculated is by using the benchmark or average strokes per hole, then taking the difference of the player’s shot and the average shot. For example, Mark Broadie further explains this in an interview with golf commentator Brandel Chamblee (as cited by Golf Channel, 2023). Consider that Justin Thomas is playing a par 5 where the PGA Tour average is 4.6 strokes. First, imagine that the hole is 4.6 strokes long not 525 yards, then he hits his tee shot very well and only has 3.4 strokes left to get in the hole on average. Justin Thomas’ corresponding strokes gained Off the Tee would be the original 4.6 – 3.4 because that’s where he ended up, which equals 1.2. Then you subtract 1.0 since it took him 1 stroke to get his ball there, and you end up with 0.2 strokes gained Off the Tee.

After the use of strokes gained was implemented onto the PGA Tour, it quickly gained traction and widespread use amongst all professional golfers. Naturally, strokes gained became a new way to quantify the game of golf and can also be used to determine the optimal players for each of the U.S. and European Ryder Cup teams.

Introduction to the Ryder Cup

The sport of golf is unique in that it is one of the most individual sports where players play for themselves and are only playing in competition with the help of one caddy. However, there is a time where golf becomes a team sport, and that happens once every two years at the Ryder Cup. The Ryder Cup is a team golf tournament between the United States and Europe, where each team consists of 12 golfers, and they compete against each other as a team.

A unique aspect of the Ryder Cup is that it is the only professional golf tournament that is not played for any prize money, yet most of the players claim that it is one of the best tournaments in golf. This is because they believe that playing for something bigger than yourself such as your country or continent, means more than simply playing for money like every other tournament. Three-time major champion Jordan Spieth once said in an interview that, “I’d rather play better in the Ryder Cup than in the [PGA] Tour Championship” (as cited in McAfee, 2021). The Ryder Cup is also played in a match play format, which is different than the stroke play format typically played every other week on the PGA Tour. In match play, you are truly competing against an opponent, where each hole is worth 1 point and whichever team wins more holes wins that match. For instance, if a U.S. pairing scores a 3 on a hole and Europe gets a 4, then the US would win that hole and be 1 up in the match. If the two pairings tie on a hole, then neither team wins the hole and you play the next hole with the same score as before, then if Europe were to win the next hole, the U.S.’s 1 up lead would decrease to all-square. Once a specific match is finished, if the U.S. is leading, they earn 1 point for Team USA and vice versa with Europe. However, if they play all 18 holes and are all square at the end, both USA and Europe obtain ½ point each.

The overall format of the Ryder Cup is split into 3 different categories: Four-Ball, Foursomes, and Singles. The Four-Ball matches consist of 2 U.S. players against 2 European players, where everyone plays their own ball and then each team counts the lower score between the two players as their team score on that hole. The Foursomes format is similar to Four-Ball in that it is 2v2, but instead of everyone playing their own ball, it is an alternate shot format where you and your partner alternate hitting consecutive shots. Lastly, the singles format is as it sounds, where 1 player from the U.S. team plays against 1 player from the European team in a match play format. It is also worth noting that the Ryder Cup takes place over a 3-day span, with two sessions of play on the first and second days. On each of the first two days, there are 4 Foursomes matches and 4 Four-Ball matches played, for a total of 8 2v2 matches on each of the first two days of competition. On these first two days, there is no equal participation requirement from the captain to the players, meaning that one player could play in all 4 sessions while another player could play in only 1 session. This decision is the captain's responsibility and is strategically selected based on factors such as current play, course fit, and pairing/opponent fit. The final day of competition is the singles day, where all 12 players on the United States team get matched up to play each of the 12 European players in a 1v1 match play format. That means, given there is a maximum of 1 point that can be earned for every match, there are a total of 28 points that can be earned at the Ryder Cup (8 from Foursomes, 8 from Four-Ball, 12 from Singles). The Ryder Cup is won by the team that accumulates 14½ points first, and in the unlikely event of a 14-14 tie after 3 days of competition, the Ryder Cup is retained by the defending champions.

In terms of the makeup of the 12 team members, they are chosen partially based on a rankings system and partially by the Team Captain of each side. For each team U.S. and Europe, there exists a player ranking system and the 6 players with the greatest number of points are automatically chosen to play in the Ryder Cup for a given year. Then the last 6 spots on each team are chosen solely by each captain’s discretion. Therefore, much strategy is put into the selection of the captain’s picks and there are many different factors to consider. In this thesis, I analyzed the Ryder Cup pairings and how they can be made in the most optimal way using strokes gained data.

Chapter 2

Literature Review

When looking at the past literature relating to golf statistics, most of them deal with more simple statistical performance measures such as driving distance, driving accuracy, green in regulation percentage, and putts per round. After the PGA Tour adopted the use of strokes gained to measure player performance, coupled with the widespread use of shotlink, it became the hot topic in the world of golf statistics.

There was a paper published on this exact topic, investigating how synergy works withing the context of the Ryder Cup. The study by Cynthia and Beach Clark tries to look at how other sports and professions utilized synergy to overcome skill and then applied their model towards professional golf. More specifically, breaking down team sports and the Ryder Cup team competition into a ranking of all the individual players that made up that given team. Then quantifying the skills of all the members and comparing that to the outcome of how well they performed. For example, if a Ryder Cup team or pairing does not have many highly ranked players on it, yet they perform very well compared to expectations, then there must be something else at play which they call "synergy". They tested two different hypotheses, the first being "When synergy is identified in dyads, performance improves" and the other looking at the team as a whole instead of the dyad or pairings (Clark and Clark, 2015). To quantify the skill level of the players, they used scoring average for that season to rank the players, and then created a t-test to look at the averages between the American and European teams. In the end, they found that synergy was present for testing hypothesis 1, but it was not significant enough to reject the null hypothesis. However, when testing their hypothesis 2, they found that the t-test showed the

Americans had “significantly more individual skill and consistency” than the European team. Out of the 72 observations considered in the study from 6 Ryder Cup events, the mean scoring average for the European team was 72.36 with a variance of 1.56, and the American’s was 70.81 with a variance of 0.41. Given that the United States team has a fairly large skill advantage over the European team, it would be expected that they win the majority of the Ryder Cup tournaments. However, this is not the case, as Europe has won 56% of the total points compared the USA’s 44%, as well as 5 of the 6 Ryder Cup trophies from the years 2004-2014 which comprises the study’s data. The Clarks then saw this as evidence that something beyond just skill was at play and demonstrated their synergy concept that accounts for the overwhelming success of the European team over the Americans. Therefore, they concluded that there was enough evidence in the data and from the t-test to reject the null hypothesis, accepting that synergy present in the whole team improves performance (Clark and Clark, 2015). Overall, it is quite useful to understand synergy and that there is more to the Ryder Cup than just the skill level of the players. The ideas brought up are very interesting, and I will be looking at the Ryder Cup similarly, by asking why Europe has been so dominant given their team is not more skilled. This is a multi-faceted question that could be investigated deeply in any sort of lens, however for my paper I will primarily be looking at the Foursomes and Four-Ball pairings. Assuming evenly matched teams, one would expect the difference in total points won to be fairly equal between Europe and the United States. However, Europe has performed significantly better in Foursomes while Four-Ball has gone as expected.

Another interesting paper by Rose Baker and Ian McHale published in 2016, uses empirical Bayes’ analysis to rank the players in the Ryder Cup. Their goal was to create a model to quantitatively estimate a power ranking of both the American and European players in the

Ryder Cup from 1979-2004. This quickly becomes complicated for two reasons: not every golfer plays in an equal number of matches in the Ryder Cup, and a player's performance is subjectively related to their opponent's performance. This means that some players don't have very much useable data which makes measuring with any degree of certainty difficult. Nonetheless, they overcome this challenge by assuming that the strengths or skills of the players come from some common distribution, which then becomes very useful for modelling and making predictions. Another issue that is relevant to determining the best Ryder Cup players is that generally speaking, there is a small number of competitors which can cause some uncertainty in the predictions. Baker and McHale fit the model using Monte Carlo Integration which allows them to consider a greater number of observations and create a more accurate prediction. The authors used a strength parameter, α given $\alpha > 0$, which is given to each player and allows them to calculate the probability as a function of the ratio between the strength parameters of the competitors. They then used a paired comparison model to rank competitors competing in head-to-head matches out of a larger pool of players, such as the Ryder Cup. One of their findings is that, based on their rankings, the U.S. teams of 2004 and 2006 underperformed the most relative to their expected outcomes. That is, in terms of skill, they had a sizeable advantage over the European teams for those two years, yet Europe won both of those Ryder Cup's. This implies that something beyond skill caused these two upsets. Ultimately, Baker and McHale found that there tends to be very little difference in the strengths between the very best players, with the α only ranging from 1.009 to 1.003 within the top 14 players. This is backed by their claim that there is no statistical significance supporting the differences in skill, however there is relatively not that much data, and they acknowledged that they would need more data to test it further. Lastly, one of Baker and McHale's most intriguing results is that the overall strength of a foursomes match pairing is closely aligned with the strength of the strongest player (Baker and McHale, 2015). This is quite counterintuitive, as the traditional logic is that in foursomes matches, you can only be as good as your weakest player.

As mentioned in the introduction, Columbia professor Mark Broadie laid the groundwork for strokes gained in his 2008 paper, *Assessing Golfer Performance using Golfmetrics*. Later on, he continued this work on strokes gained by expanding onto the PGA Tour with his 2012 paper titled, *Assessing Golfer Performance on the PGA Tour*. This paper specifically takes a deep dive into the four major parts of the game. Each part is represented by strokes gained statistics and looks at which are the most impactful for scoring on the PGA Tour. Broadie first compares the strokes gained metrics to more traditional types of golfer performance statistics then argues that strokes gained can help get a wider and clearer view of relative performance. Overall, his strategy is to use strokes gained as a way to incorporate a benchmark into the PGA Tour's statistics, where everything under strokes gained is comparative to the Tour average for that specific part of the game (driving, irons etc.). He calculates a benchmark function that varies based on "distance to the hole and depends on course condition at the location of the ball (i.e., tee, fairway, rough, green, sand, or recovery)" (Broadie, 2012). This benchmark utilized the ShotLink data provided by the PGA Tour to analyze over 8 million golf shots over the years 2003-2010. Broadie broke this data down into three broad categories: Long Game, Short Game, and Putting. After using his formulas, Broadie did a comparative analysis and found overwhelming evidence that supports long game as being the most important factor in determining golfer's skill. For example, Broadie found that in the 10 players with the highest total strokes gained per round on Tour, long game accounts for 65% of their strokes gained advantage. Conversely, in the 10 players with the lowest total strokes gained in that same time span, they lose 71% of their stokes due to long game (Broadie, 2012). These findings tell us that when compared to the PGA Tour average, long game is by far the most important and influential factor in explaining the variability in scores. In fact, Broadie concluded that "the contributions to total strokes gained are 72 percent, 11 percent, and 17 percent for the long game, short game, and putting, respectively" (Broadie, 2012) putting further emphasis on the long game aspect of golf for Tour professionals. Assuming that Mark Broadie's definition of *long game* consists of both tee shots and iron shots (SG OTT and SG APP), then in my paper I will be looking at these two specific categories when comparing Ryder Cup players and pairings. Additionally, I will combine these two categories into long game to see if there is any difference related to Ryder Cup performance.

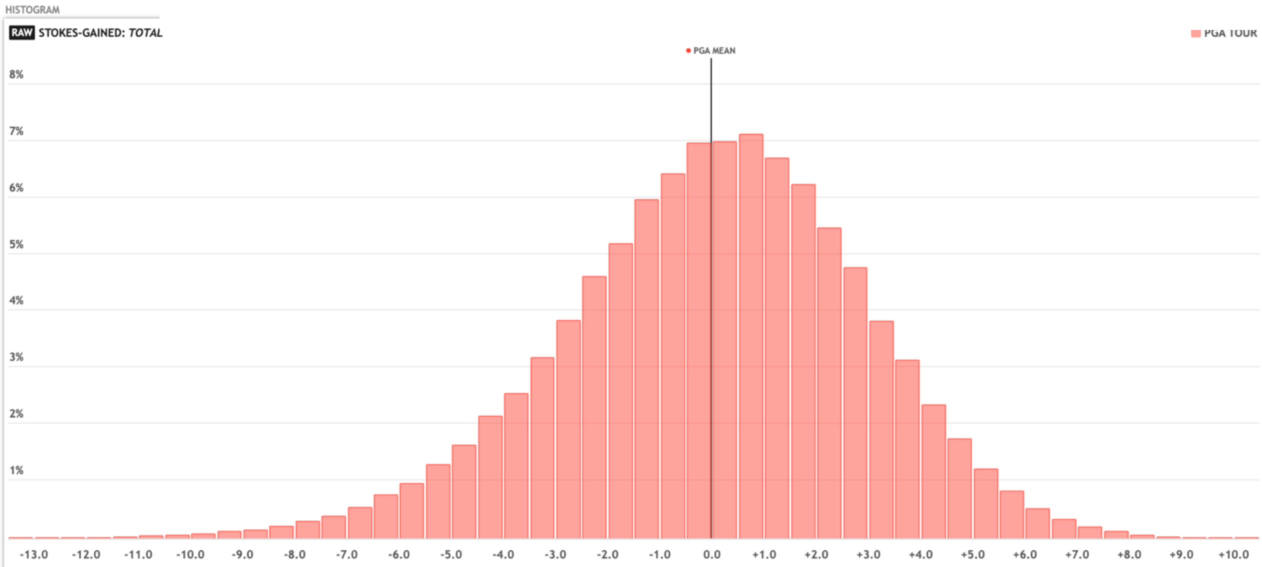
Chapter 3

Data Selection and Assumptions

The data that I obtained came from a couple of online sources through the use of data scraping into Microsoft Excel. The first source was the PGA Tour official website (Golf Stat and Records), where all strokes gained data for every player on the PGA Tour is publicly available. The players' strokes gained data only goes back to the 2004 season because that is the first year when the PGA Tour had both the strokes gained and ShotLink technology track every shot hit in every tournament. Additionally, I only considered the data for the United States Ryder Cup team since their data is readily available on the PGA Tour website since they play on the PGA Tour. However, many of the players for the European Ryder Cup teams play on the DP World Tour, which is the premier professional golf tour in Europe. Therefore, the strokes gained data for the European players is not as readily available, as this tour doesn't use the ShotLink system.

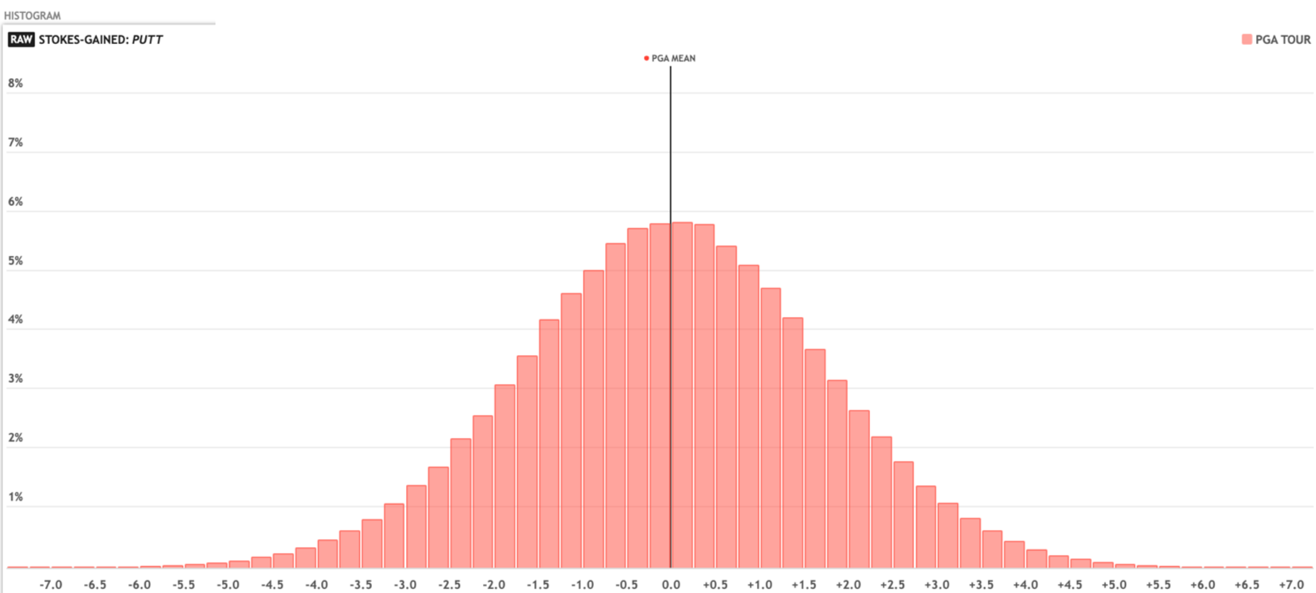
An assumption that I made was that for any given strokes gained category on the PGA Tour, the distribution was normally distributed with an expected value of 0. This assumption is for both SG OTT, SG APP, SG ATG, and SG PUTT categories as well as SG TOTAL which is where you simply add up all of the individual categories. Figures 1 and 2 below from datagolf.com does a great job at showing this distribution visually, for SG Total and SG Putt.

Figure 1. SG: Total Discrete Distribution on PGA Tour



(SG Distributions)

Figure 2. SG: Putt Discrete Distribution on PGA Tour



(SG Distributions)

This also makes sense intuitively because when you refer back to how strokes gained were first started, the benchmark was always PGA Tour average. Naturally the mean of all the player's strokes gained for all rounds on Tour should equal zero. This assumption allowed me to compare a player's strokes gained average in a given year to the PGA Tour's average strokes gained in a specific category in any given year. For instance, if a player that's playing in the Ryder Cup gained 0.05 in the SG APP category, then it's known he is doing better than the PGA Tour average, no matter which year he played.

Data Collection

First, I created a data frame consisting of every American golfer who played in a Ryder Cup between the years of 2004-2023, totaling 10 Ryder Cups. Next, I obtained the season averages for each of the four strokes gained categories; Off the Tee (OTT), Approach the Green (APP), Around the Green (ATG), and Putting (PUTT) for every Ryder Cup player in every Ryder Cup season. Since the same players might be in consecutive Ryder Cups, but their strokes gained averages are different in each of those years, that resulted in a total of 120 unique players' strokes gained data elements. The reason for using Ryder Cup players' strokes gained averages is to provide a way to quantify each player's performance in each general category during that PGA Tour season. This allows me to be able to break down each player's performance into the four SG categories so, for example, I can analyze how being a good or bad putter during the PGA Tour season is possibly related to Ryder Cup performance.

The websites datagolf.com (Ryder Cup) and golfcompedium.com (Ryder Cup Scores) were used to scrape the pairings and players' win/loss records in each of the last 10 Ryder Cups. As mentioned previously, in every Ryder Cup pairing or singles match, there are only three possible outcomes: win (1 point), tie (½ point), or lose (0 points). This was combined with the strokes gained season averages data for all the American Ryder Cup player's, to match them into one data frame. Figure 3 below is an example of how the data was organized. This data reflects the 24 Foursomes matches for the last three Ryder Cups (2018, 2021, and 2023).

Figure 3. Foursomes Pairings Data Set

Pairing Last Names	Year	SG_OTT	SG_APP	SG_ATG	SG_PUTT	SG_OTT2	SG_APP3	SG_ATG4	SG_PUTT5	Points Earned
Scheffler/Burns	2023	1.021	1.194	0.399	-0.301	0.266	-0.1	0.046	0.584	0
Homa/Harman	2023	0.317	0.471	0.24	0.621	0.241	-0.03	0.031	0.446	0
Fowler/Morikawa	2023	0.126	0.74	0.269	0.29	0.487	1.012	0.047	-0.109	0
Schauffele/Cantlay	2023	0.265	0.88	0.056	0.667	0.852	0.564	0.147	0.305	0
Thomas/Spieth	2023	0.207	0.306	0.44	-0.222	0.2	0.173	0.25	0.086	0
Scheffler/Koepka	2023	1.021	1.194	0.399	-0.301	-	-	-	-	0
Homa/Harman	2023	0.317	0.471	0.24	0.621	0.241	-0.03	0.031	0.446	1
Schauffele/Cantlay	2023	0.265	0.88	0.056	0.667	0.852	0.564	0.147	0.305	0
Thomas/Spieth	2021	0.235	0.887	0.296	0.026	-0.091	0.365	0.407	0.389	0
Johnson/Morikawa	2021	0.361	0.315	0.147	0.429	0.281	1.17	0.077	-0.457	1
Koepka/Berger	2021	0.631	0.482	0.058	0.187	0.264	0.834	-0.017	0.214	1
Schauffele/Cantlay	2021	0.266	0.65	0.087	0.48	0.569	0.486	0.363	0.402	1
Koepka/Berger	2021	0.631	0.482	0.058	0.187	0.264	0.834	-0.017	0.214	0
Johnson/Morikawa	2021	0.361	0.315	0.147	0.429	0.281	1.17	0.077	-0.457	1
Thomas/Spieth	2021	0.235	0.887	0.296	0.026	-0.091	0.365	0.407	0.389	1
Schauffele/Cantlay	2021	0.266	0.65	0.087	0.48	0.569	0.486	0.363	0.402	1
Johnson/Fowler	2018	0.919	0.829	0.238	0.385	0.244	0.494	0.242	0.296	0
Watson/Simpson	2018	0.78	0.005	-0.231	0.001	-0.116	0.486	0.37	0.692	0
Mickelson/DeChambeau	2018	-0.225	0.555	0.091	0.51	0.586	0.556	0.07	0.346	0
Thomas/Spieth	2018	0.408	0.844	0.311	0.272	0.271	0.409	0.184	-0.034	0
Johnson/Koepka	2018	0.919	0.829	0.238	0.385	0.658	0.232	0.162	0.168	0
Watson/Simpson	2018	0.78	0.005	-0.231	0.001	-0.116	0.486	0.37	0.692	1
Woods/DeChambeau	2018	0.061	0.883	0.385	0.266	0.586	0.556	0.07	0.346	0
Thomas/Spieth	2018	0.408	0.844	0.311	0.272	0.271	0.409	0.184	-0.034	1

The Ryder Cup was not held in 2020 due to COVID-19, thus switching the tournament years from even to odd. A similar data set was created for the Four-Ball matches played, where there was a total of 80 rows for each data set. This is because there are 8 Foursomes and 8 Four-Ball matches played in each Ryder Cup, and there is data for the past 10 Ryder Cups. Additionally, there was one missing value as you can see in the Figure 3, where Brooks Koepka's strokes gained season averages are not filled in. This is because he did not play on the PGA Tour in the 2023 season. To account for this lack of data, I did not consider the singles match that he played in and only considered his partner's averages for his Foursomes/Four-Ball matches. One noteworthy aspect about the Points Earned column is that the format of every Ryder Cup match is team vs. team and match play as opposed to stroke play. This means that just because a player plays well in their match, doesn't mean that they will win their match due to factors such as partner performance or opponent performance. Likewise, even if the U.S. team plays well during the entire Ryder Cup, it doesn't mean they will win the Ryder Cup as they could be outplayed by Europe.

Furthermore, between all three types of Ryder Cup matches: Singles, Foursomes, and Four-Ball, I only considered the pairings format of matches (Foursomes and Four-Ball) for modelling purposes. When thinking about the team captain's picks for players and pairings, there is an element of randomness as well as many other factors that need to be considered. As outlined in the introduction, Ryder Cup captain's picks are a long and thoughtful process, full of much deliberation of factors such as statistics, course fit, team fit and more. However, with the singles matches, I assumed that they are mostly based on skill level represented by each player's seasonal strokes gained averages. This means one can assume that in general the player with the better skill level and course fit should win, which is not relevant to the captain's picks and pairings selections. Therefore, I will model both the Foursomes and Four-Ball 2v2 matches from the United States team's perspective, comparing them to each other and analyzing the strategies that go into them.

Lastly, another assumption I made about the outcomes of the Ryder Cup is that I assume the United States and European Ryder Cup teams are evenly matched. Similar to a coin flip, each outcome in this case (U.S. win or U.S. loss) has a probability of 0.5 and no matter how many times that event is repeated the probability stays the same. That is, it is expected for each team to win the same amount in all 160 pairings matches. For instance, from the American team's perspective, it can be assumed that they expect to have won 40 of the 80 Foursomes matches and 40 of the 80 Four-Ball matches in the last 10 Ryder Cups. In summary, I have made these assumptions with the intent of simplification and to make the modeling easier to perform and understand.

Chapter 4

Modeling

To quickly recap how the modelling will work, using the data set created for all of the Foursomes and Four-Ball matches. I analyzed all of the players' seasonal strokes gained averages, and then compare them to their Ryder Cup points earned to determine which factor or factors are most impactful.

First, given the assumptions made in the previous chapter that the U.S. team is equally likely to win as the European team, we can expect the U.S. team to earn 40 out of 80 points in both Foursomes and Four-Ball. To officially test this assumption against the actual results of the 160 Ryder Cup matches, I used a discrete binomial distribution. The binomial distribution is a discrete probability distribution that is useful for modelling an event or series of events with only two outcomes. The formula is defined by the probability density function below.

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

where n = # of trials
x = # of successes
p = probability of success

Based on the assumption that both teams are equally likely to win, that means the p in the formula is equal to 0.5 and 1-p would also be equal to 0.5. For the purposes of this modeling, I will only consider the Foursomes and Four-Ball matches that ended in a win (1 point) or loss (0 points). This is to make it easier and simpler to model, as now there are only two outcomes being win or lose, as opposed to before when there was a third outcome which was tie. For the 80

Foursomes matches, there were 11 matches that ended in a tie. After omitting them there were 69 Foursomes matches. Based on the assumption we could expect the United States team to win half of the matches or have an expected value of $E[x] = np = 34.5$ wins. Similarly, for the Four-Ball matches there were 16 that ended in a tie, resulting in 64 matches to consider under this model. The expected value of wins for the United States team would be $E[y] = np = 32$ wins. However, when looking at the actual results the United States team won 28 of the 69 Foursomes matches and 33 of the 64 Four-Ball matches. At first glance, it seems as though in Four-Ball, the U.S. has a slight advantage over the European team. But also, that the total amount of wins is very close to the expected value calculated above. However, for the Foursomes matches, it seems clear that the United States team performs below expectations. In order to quantify these results, I ran a binomial t-test in R Studio using a 95% confidence interval and alpha value of 0.05. Essentially, this t-test uses the binomial distribution as its model and compares the realized value (28 and 33 successes) to all outcomes and determines how likely it is to have happened. It does this using the implied probability that was input based on the assumptions of 0.5 for probability of success. Ultimately, the results showed that for both the Foursomes and Four-Ball outcomes, the hypothesis of $p = 0.5$ hold true as the p-values were both > 0.05. This suggests that the results were not significantly different enough from expected to reject $p = 0.5$. Now that the assumption was reinforced by the t-test, I continued modelling the number of outcomes in each type of match using the binomial distribution with $p = 0.5$. This is the graph of both the Foursomes and Four-Ball probability density functions (pdf), highlighting the number of successes in each with a white bar.

Figure 4. Binomial Distribution PDF – Foursomes Matches

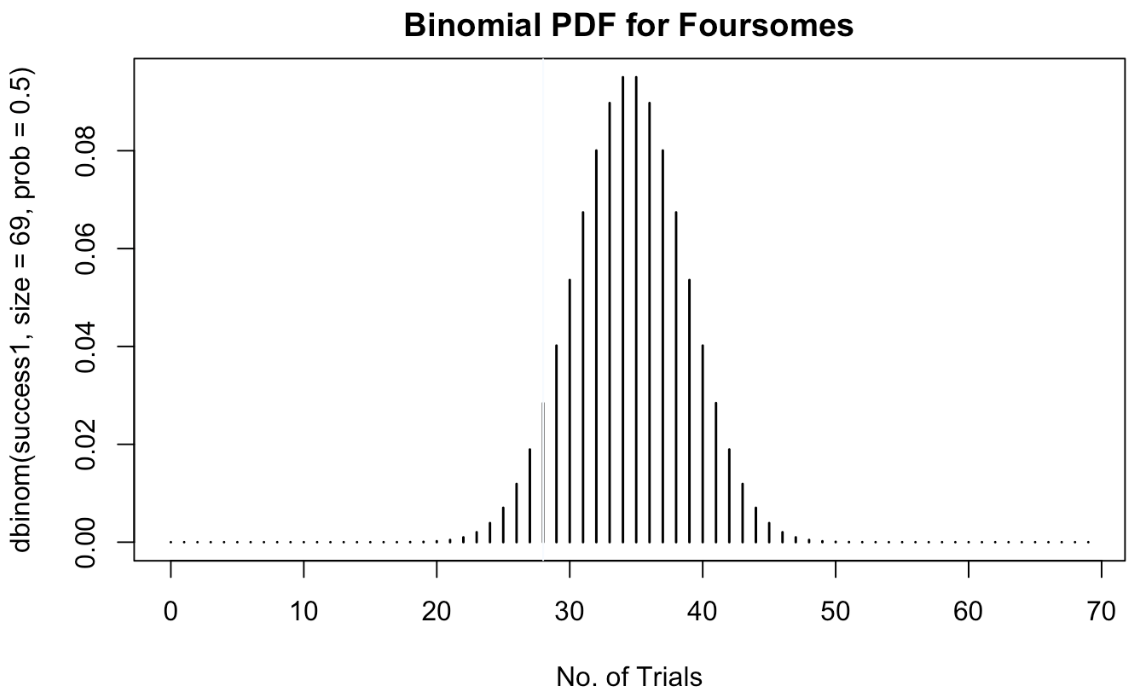
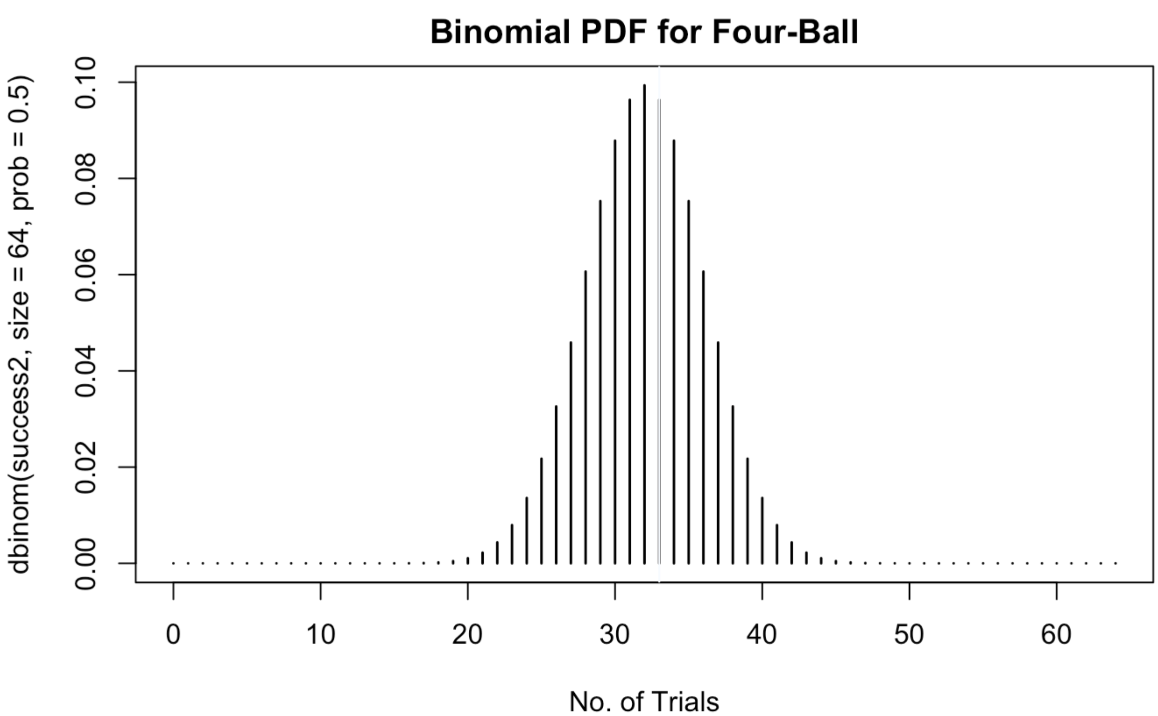


Figure 5. Binomial Distribution PDF – Four-Ball Matches



From the graphs, it's evident how due to the binomial distribution assumption with probability p = 0.5, there are high probabilities in the middle and much less towards each of the ends. To illustrate, this can be essentially compared to a flip of a fair coin since the probability of success is 0.5. Imagine a fair coin is flipped over 60 times, naturally it makes sense to expect that there is a similar number of heads and tails, or in the case of this thesis it being Ryder Cup wins and losses. However, when there is an extreme value that is different from the expected, it suggests that the underlying probability might not be 0.5 or there could be some other confounding factor. Additionally, I used the cumulative distribution function (cdf) to determine the probability of seeing the x number of successes or a more extreme value (farther away from expected value). These are the figures of the cdfs for the Foursomes and Four-Ball matches.

Figure 6. Binomial CDF – Foursomes Matches

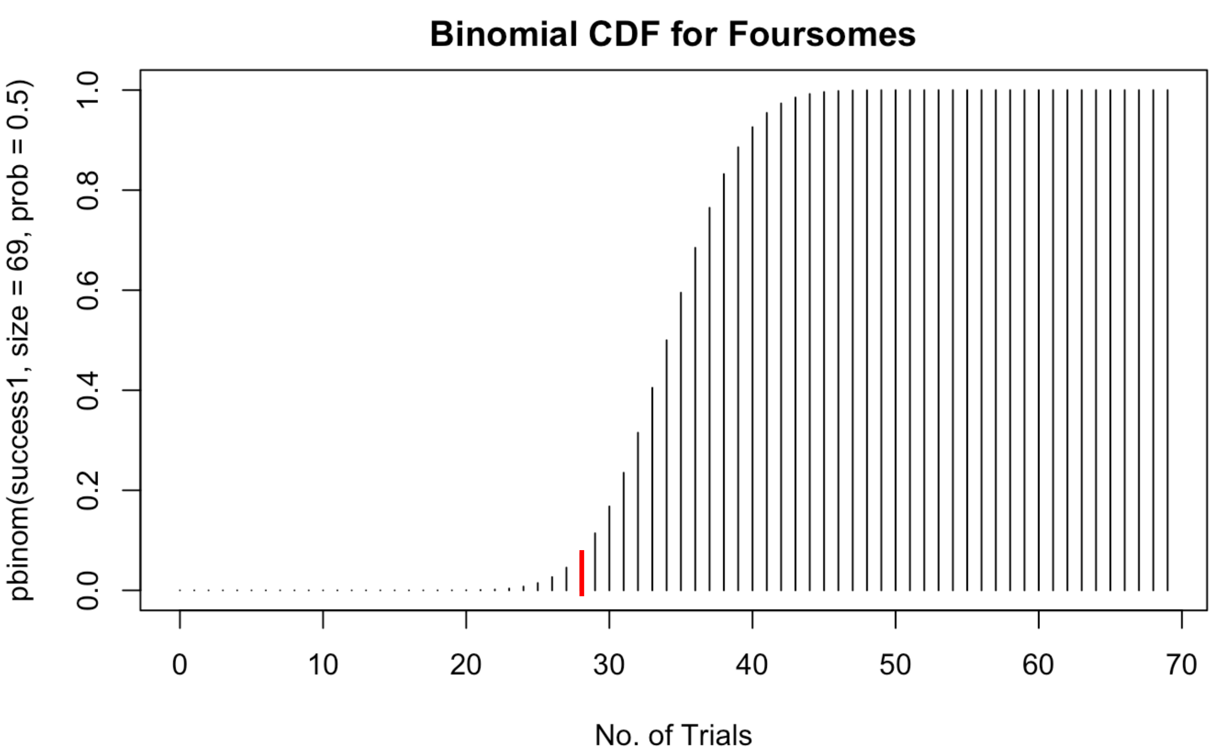
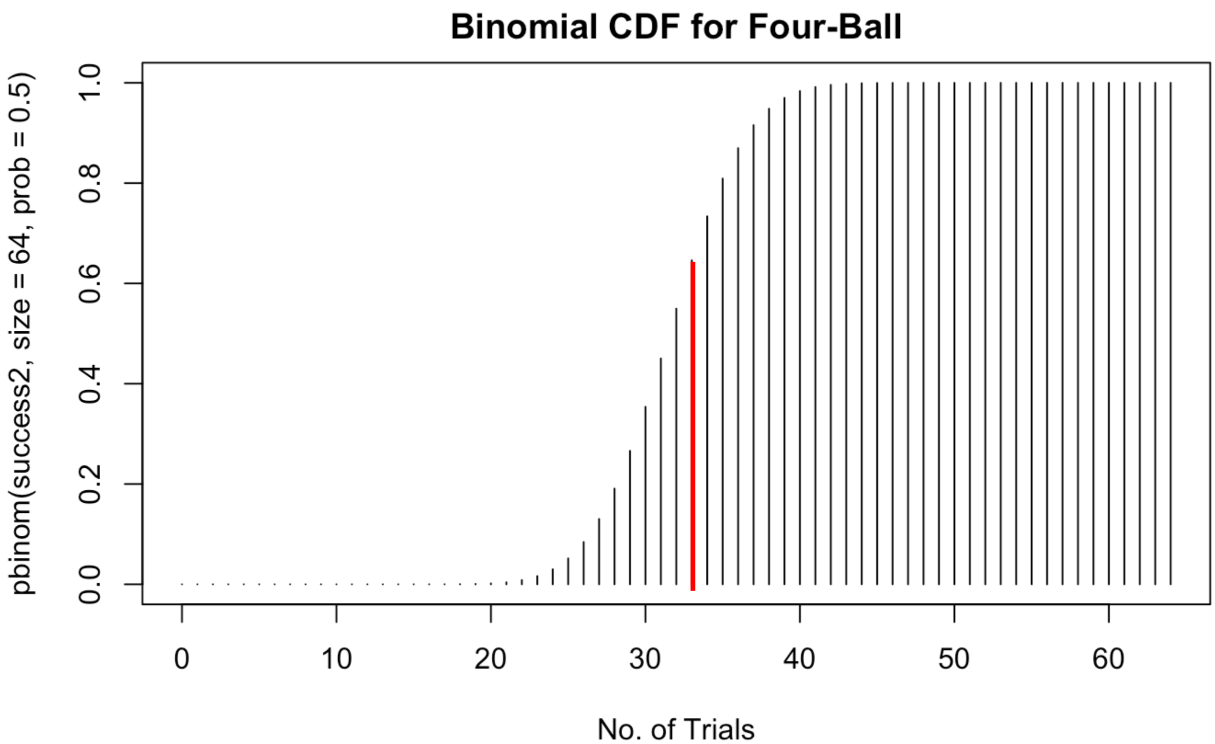


Figure 7. Binomial CDF – Four-Ball Matches



These Figures 6 and 7 are a more visual representation that the probability of outcomes increases as the number of successes goes up. In the beginning, between 0 and 20 successes, the probability is very low as these outcomes are highly unlikely, and then the cumulative probability increases rapidly just after 30 trials. The number of successes in each of the types of matches is represented by the red lines in the graphs and shows the cumulative probability of that event (x successes in n trials each with a 0.5 probability) happening.

As the results were quantified it showed that for the Foursomes matches, the probability of getting 28 or less wins in 69 matches, is equal to 0.07401 or 7.4%. Similarly for the Four-Ball matches, the probability of getting 33 or more wins in 64 total matches is equal to 0.646 or

64.6%. This stark difference shows exactly how unlikely it is that United States team won so little of the Foursomes matches. It also shows that the amount of wins the United States team has earned over the past Four-Ball matches is very close to the expected amount of 32, and therefore is represented by a high probability of that outcome or better happening. In general, these results suggest that while the American team has the slight advantage in the Four-Ball format, this is entirely predictable and within our assumptions. However, for the Foursomes it is clear that the European team has a significant advantage on the United States team. The binomial distribution alludes that there might be other factors that give the Europeans the edge in this format.

Foursomes Pairings Modeling

Based on the previous subchapter, the United States team has significantly underperformed compared to expectations in the Foursomes match format. There could be a number of reasons for this outcome, such as many highlighted by previous authors in the literature review chapter. Many others in the golfing world have theorized that it could be because the European team has more chemistry than the American team. According to former U.S. Ryder Cup player Justin Leonard (as cited in Beall, 2021), chemistry definitely plays a role and the winning teams he’s been on have all had good chemistry. This concept could be plausible as the Foursomes’ format requires much more teamwork than Four-Ball. In thinking about the game of golf in terms of the Foursomes format, it makes sense that pairing chemistry leads to higher performance. Since one of the players hits the drive, then the other player in the pairing hits the approach shot and so on, it puts a greater emphasis on communication and teamwork. For instance, consider if Player A has a strong long game (driving and irons) but poor putting/short game and Player B has a weak long game but strong short game. It may seem constructive to pair them together so that Player A can hit approach shots and Player B can putt. However, this is not ideal in Foursomes, as there will also be instances where Player B must hit the approach shot and Player A must hit the putt/chip. Additionally, a player simply having a poor short game can be problematic in Foursomes regardless of the strokes gained skills of their partner. This creates a sense of added stress on their partner in both trying to hit a closer approach shot, as well as on second putts if the poor putter hits first on the green.

First, in order to model this difference between the long and short game between players, I used the data set defined in the previous chapter. The first model I used was one that calculated the absolute difference between the players' SG APP and SG PUTT seasonal average values. This model simply takes the absolute difference of one player's SG APP minus their partner's SG PUTT and vice versa, and then returns whichever absolute value is greater. Then I compared these numbers to the number of points earned in each Ryder Cup match, which is only win or lose in this case. To determine whether or not the maximum difference between SG APP and SG PUTT was relatively large or not, I used conditional formatting to determine which values were "High" or above the mean of all of them and "Low" or below the mean.

Table 1. Maximum Difference between SG_APP and SG_PUTT in Foursomes

In losses:	**Count**	**%**
Total:	40	
# of High:	22	55%
# of Low:	18	45%

In Wins:		
Total:	28	
# of High:	12	43%
# of Low:	16	57%

These results show that for the Foursomes matches, there were 40 matches the U.S. team lost and 28 that they won. In 55% of the Foursomes matches that the U.S. team lost, there was a relatively high maximum difference between one of the player's SG APP and the other's SG PUTT, as opposed to only 43% in matches the U.S. won. The results for this model using Four-Ball data are almost identical to these, with 55% of losses having a high maximum difference player, and 45% in matches the U.S. won.

Similarly, I chose to next model the minimum sum of the pairing partner's SG APP and SG PUTT statistics. Specifically, this model adds together each players SG APP and SG PUTT season averages, then returns the lower of the two values. This model is useful for finding out how it affects match points earned when one player is strong in one area and their pattern is weak or they are both weak. The purpose of choosing just the lower of the two sums is to model how Ryder Cup pairings perform when at least one of their combinations of SG APP plus SG PUTT is poor relative to average. As before, these values are then compared to themselves by finding which are higher than the mean and lower than the mean, and then analyzed with their points earned results.

Table 2. Minimum Sum of SG_APP and SG_PUTT in Foursomes

In losses:	Count	%
Total:	40	
# of High:	24	60%
# of Low:	16	40%
In Wins:		
Total:	28	
# of High:	13	46%
# of Low:	15	54%

These results tell almost the opposite story as the first model, where in the majority of matches the U.S. team lost, there were pairings with at least one high sum of SG APP and SG PUTT. When comparing these percentages to that of the Four-Ball matches under the same model, they are different as well. Where 61% of the 31 matches the US team lost in Four-Ball had relatively low Min Sum values and only 45% of the 33 wins had the same. Based on these results I took the SG APP out of the modeling as it does not have as much of an effect on winning in the Ryder Cup as initially thought by Broadie.

I then tried to focus the model simply on putting, while continuing the earlier proposition about one player in a Foursomes pairing having relatively poor skill levels. This time, I used the model MIN PUTT, which simply returns the lower of the SG PUTT value between pairing partners.

Table 3. Minimum SG_PUTT for Foursomes

In losses:	Count	%
Total:	40	
# of High:	19	47%
# of Low:	21	53%
In Wins:		
Total:	28	
# of High:	16	57%
# of Low:	12	43%

Table 4. Minimum SG_PUTT for Four-Ball

In losses:	Count	%
Total:	31	
# of High:	14	45%
# of Low:	17	55%
In Wins:		
Total:	33	
# of High:	20	61%
# of Low:	13	39%

The results based on this model differ compared to the previous models with APP included, as now there is a larger percentage of player's having low SG PUTT averages in matches lost than in matches won for Foursomes. This also applied to the Four-Ball format where we see similar things. In 55% of losses there was at least one player with a low SG PUTT as opposed to 39% in matches the U.S. team won. In comparing the model's results between Foursomes and Four-Ball, it's evident that having a relatively bad putter in a pairing is detrimental to the pairings results. It is even more so in Four-Ball than Foursomes format, showing a few percentages increase. This makes sense because Four-Ball is more of an individual style match, where having one bad putter can be made up for by great partner play. To further explore these results, I chose to change the benchmark of comparing the MIN PUTT model for SG PUTT to the PGA Tour average rather than the average of past U.S. Ryder Cup players. This new model still returned the lower of the two SG PUTT values between each pairing member, but only counted if the number was below the overall PGA Tour average, which as defined in the previous chapter is zero. In total there were 16 Foursomes matches where at least one of the players' SG PUTT was negative.

Table 5. MIN PUTT Below PGA Tour Average - Foursomes Matches

Outcome of Match	# of times outcome occurred	%
Win	5	31%
Loss	11	69%

I also chose to compare the SG PUTT for the Four-Ball matches, and there was a total of 22 matches where at least one partner has negative values.

Table 6. MIN PUTT Below PGA Tour Average - Four-Ball Matches

Outcome of Match	# of times outcome occurred	%
Win	12	54.5%
Loss	10	45.5%

Based on these results, we can see that in Foursomes matches where at least one U.S. player's SG PUTT seasonal average is below PGA Tour average, the U.S. team loses 69% of the time. Conversely, in Four-Ball matches where at least one U.S. player's SG PUTT average is below PGA Tour average, the U.S. team loses 54.5% of the time. This gives us insight towards answering the question of why the U.S. Ryder Cup teams have been so consistently poor in the Foursomes match format. These results suggest an impact having a bad putter in a paring has towards earning a point in a given match. This also makes sense intuitively, because in a Foursomes style match, if a pairing contains a bad putter that puts added stress on both that player's putts and their partners approach shots to be closer. This type of added stress can definitely make a difference for a Ryder Cup team, especially when there are multiple bad putters on an overall team.

Chapter 5

Limitations and Conclusion

As determined in the previous chapter, the U.S. Ryder Cup team has historically performed poorly in the Foursomes format and one major cause appears to be poor putting. Specifically, having at least one player in a pairing that has a seasonal SG PUTT average below the PGA Tour average of zero. This kind of poor putting appears to be more detrimental in the Foursomes matches than in Four-Ball, as partners need to rely on their teammates much more in Foursomes opposed to Four-Ball. This also explains why when assuming that the U.S. and European teams are evenly matched, the European team dominates in Foursomes. Similarly, why the U.S. team has the historical advantage in the Four-Ball format, as it is more dependent on individual skill and performance which is what the American golfers excel at. Since the stokes gained data is only available on the PGA Tour, it is difficult to compare the U.S. players' statistics with the Europeans. There are also no strokes gained data available from any Ryder Cup tournament, which makes measuring performance accurately more difficult as well. These results can then be very useful for the Ryder Cup captain whose job it is to both choose the optimal six golfers to be on their team and organize the lineup of match pairings. Both of these tasks are difficult and subjective but could be made easier by selecting golfers with relatively high strokes gained putting averages during the season leading up to the Ryder Cup, as well as pairing players together who have similar play styles. For instance, the model also suggests that historically when you pair players together that have a large difference between their SG Approach and SG Putting statistics, they lose more often than not in Foursomes. This is also based on the logic that paring a good iron player with a good putter doesn't necessarily result in winning Foursomes matches if the good iron player is a poor putter and the good putter is a poor iron player.

However, there are some restrictions that are not able to be considered based on the current amount of available information. One aspect that the strokes gained statistics does not take into account is player-team comradery. As described by many past Ryder Cup captains and players, team chemistry is very important which also puts added pressure on the team captain to select players who will fit the team dynamic and not cause and unnecessary turmoil. For instance, 2023 U.S. Ryder Cup captain Zach Johnson explained “the last thing I want is to put any sort of potential for drama” when talking about making his picks (as cited in Mumm, 2024). Another aspect that isn’t taken into consideration is the player’s current form when selecting the six captain’s picks. By using the seasonal strokes gained averages for each player, it gives a great sense of how the player has performed throughout the season. However, it could make a difference for the Ryder Cup if a golfer played very well in the beginning of the season and poorly at the end of the season and vice versa. Since the Ryder Cup is always after the season is finished, then one potential future study could investigate how a player’s current form impacts Ryder Cup performance. The last important aspect that does not get captured by the data is pressure and course fit. Strokes gained statistics are great at capturing golfer’s shots and comparing them to a common benchmark, but it also assumes that all shots are created equal. For instance, strokes gained does account for a shot that is played from the rough vs fairway vs bunker, however it cannot account for whether that shot was hit in the first round of a tournament or the final round where the stakes are higher. This can also be very important for captains to consider, as the Ryder Cup provides some of the most high-pressured situations in all of professional golf. Presumably, players who perform better under pressure should perform better in the Ryder Cup. Similarly, is course fit, as strokes gained treats all shots hit from and into the rough as rough and so on. Stokes gained and ShotLink does have categories for type of lie such as, "fairway, intermediate rough, primary rough, greenside bunker" (Minton, 2012). However, an important distinction not considered is the variability of rough from one course to the next. Players who play well on a certain type of course can be reasonably expected to perform well on similar courses. So, if a player plays well on a course similar to that of the Ryder Cup, then it could be beneficial to select that player.

Finally, due to the limitations of this study there is much opportunity for further research. While it is beneficial to know how poor strokes gained putting averages affect Ryder Cup team performance, it may be more beneficial to model how a player's current form impacts performance. Strokes gained does not consider course type or difficulty in its calculations, but one could also try to quantify the difficultly by scoring average or winning score and model Ryder Cup performance by player that way. Overall, Ryder Cup captain's pick and pairing selection is an imperfect process without definitive results. If studies like this one can even slightly help with this process and give the U.S. team a better chance of winning the next Ryder Cup, then such research would be beneficial.

Appendix A

R Studio Code

````
```{r}
#H0: p = 0.5 vs H1: P =! 0.5
binom.test(28, 69, p = 0.5, alternative = c("two.sided", "less", "greater"), conf.level = 0.95)
binom.test(33, 64, p = 0.5, alternative = c("two.sided", "less", "greater"), conf.level = 0.95)
# Shows that p = 0.5
```
```{r}
# Probabilities of getting exactly x successes in n trials
# Foursomes
dbinom(28, 69, prob = 0.5)
dbinom(34, 69, prob = 0.5)

# Four-Ball
dbinom(33, 64, prob = 0.5)
dbinom(32, 64, prob = 0.5)
```
```{r}
# CDF
# Foursomes
# Pr(X <= 28)
pbinom(28, 69, prob = 0.5)
pbinom(34, 69, prob = 0.5)
# Four-Ball
four_ball_cdf <- pbinom(33, 64, prob = 0.5)
four_ball_cdf
# Pr(X >= 33)
#fb <- (1 - four_ball_cdf)
#fb
pbinom(33, 64, prob = 0.5)
```
```{r}
# Foursomes
success1 <- 0:69
plot(success1, dbinom(success1, size=69, prob = 0.5),type='h', lwd = 1.5, xlab = 'No. of Trials',
main = 'Binomial PDF for Foursomes' )
abline(v =28, col = blues9)
# Four-Ball
success2 <- 0:64
plot(success2, dbinom(success2, size=64, prob = 0.5),type='h', lwd = 1.5, xlab = 'No. of Trials',
main = 'Binomial PDF for Four-Ball')
abline(v = 33, col = blues9)
````
````

````
```
```{r}
# Foursomes
success1 <- 0:69
plot(success1, pbinom(success1, size=69, prob = 0.5),type='h', xlab = 'No. of Trials', main =
'Binomial CDF for Foursomes')
# Four-Ball
success2 <- 0:64
plot(success2, pbinom(success2, size=64, prob = 0.5),type='h', xlab = 'No. of Trials', main =
'Binomial CDF for Four-Ball')
````

Appendix B

Microsoft Excel Formulas

*1 refers to player A in the pairing and 2 refers to player B

Absolute Difference:

= MAX(ABS(SG_APP1-SG_PUTT2), ABS(SG_APP2-SG_PUTT1))

Minimum Summation:

= MIN(SG_APP1+SG_PUTT2, SG_APP2+SG_PUTT1)

Summation of Putting:

= SG_PUTT1 + SG_PUTT2

Minimum Putting:

= MIN(SG_PUTT1, SG_PUTT2)

BIBLIOGRAPHY

Baker, R. D., & McHale, I. G. (2015). An empirical Bayes' procedure for ranking players in Ryder Cup golf. *Journal of Applied Statistics*, *43*(3), 387–395. https://doi.org/10.1080/02664763.2015.1043869

Beall, J. (2021, September 18). *With eyes on Bryson-Brooks feud, how much does team chemistry matter at the Ryder Cup?* Golf Digest. https://www.golfdigest.com/story/ryder-cup-team-chemistry

Broadie, M. (2008). Assessing Golfer Performance using Golfmetrics. In *Energy in Motion Inc* (pp. 253–262). https://www.columbia.edu/~mnb2/broadie/Assets/broadie_wscg_v_200804.pdf

Broadie, M. (2012). Assessing Golfer Performance on the PGA TOUR. *Interfaces*, *42*(2), 146–165. https://doi.org/10.1287/inte.1120.0626

Burke, B., & Katz, S. (2016, September 8). *How is Total QBR calculated? We explain our quarterback rating*. ESPN.com. https://www.espn.com/blog/statsinfo/post/_/id/123701/how-is-total-qbr-calculated-we-explain-our-quarterback-rating

Chupaska, S. (2021, October 1). *The Golf Guru | Columbia Business School*. Business.columbia.edu. https://business.columbia.edu/magazine-winter21/research/the-golf-guru

Clark, C., & Clark, B. (2015). *IS THE SUM OF THE PARTS GREATER THAN THE WHOLE? SKILL VS. SYNERGY*. Academy of Strategic Management Journal. https://www.proquest.com/docview/1760499063/fulltextPDF/BED1D125CA7E4529PQ/1?accountid=13158&sourcetype=Scholarly%20Journals

Golf Channel. (2023). *Strokes gained with "Godfather of Golf Analytics".* [YouTube]. https://www.youtube.com/watch?v=MeNHbGhPFzU

Golf Stat and Records | PGA TOUR. (n.d.). PGATour. https://www.pgatour.com/stats

McAfee, P. (2021, August 18). Jordan Spieth Tells Pat McAfee He Would Rather Win The Ryder Cup Than Tour Championship [Interview]. In *www.youtube.com*. https://www.youtube.com/watch?v=s6KZer19BHw&t=147s

Minton, R. (2012). *ShotLink and Consistency in Golf*. https://apps.roanoke.edu/minton/consistency.pdf

Mumm , C. [Producer]. (2024, March 6). Pick Six (Season 2, Episode 6) [Documentary Series Episode]. In *Full Swing Netflix*. Box to Box Films.

Ryder Cup | Live Model. (n.d.). Datagolf.com. Retrieved April 2, 2024, from https://datagolf.com/ryder-cup/live-model

Ryder Cup Scores. (n.d.). Golf Compendium. Retrieved April 2, 2024, from https://www.golfcompendium.com/2019/05/2004-ryder-cup.html

Sanders, P. (n.d.). *History of Strokes Gained*. Www.shotbyshot.com. https://www.shotbyshot.com/History-of-Strokes-Gained

SG Distributions. (n.d.). Datagolf.com. Retrieved April 2, 2024, from https://datagolf.com/sg-distributions

ACADEMIC VITA

William Gerity

EDUCATION

The Pennsylvania State University, Schreyer Honors College **University Park, PA**
Smeal College of Business *May 2024*
Bachelor of Science in Actuarial Science, Minor in Statistics

RELEVANT ACTUARIAL

- SOA Exam P - Probability (Sitting) *July 2024*
- SOA Exam FM - Financial Mathematics Passed *August 2023*
- VEE - Mathematical Statistics, Accounting & Finance, and Economics Completed *December 2022*

EXPERIENCE

Gallagher Re **Philadelphia, PA**
Actuarial Analyst Starting July 2024

Munich Re **Princeton, NJ**
Corporate Reserving Intern *May 2023-August 2023*

- Analyzed the impact of combining umbrella and non-umbrella segments, while gaining valuable skills and knowledge in LDF selection, tail curve fitting, and ultimate method selection
- Compared actuarial segments on a UY and OY basis using chain-ladder and transition point calculations in Excel to determine if a meaningful difference exists

Munich Re **Princeton, NJ**
Corporate Reserving Intern *May 2022-August 2022*

- Expanded my core competencies of actuarial science by regularly performing segment analysis and technical checks of the reserve review process
- Leveraged my technical skills to help streamline various data processes related to outlier loss expectations
- Updated current year reserve review data in Power BI dashboard to be presented to upper management

Bridge Street Golf at Snipes **Morrisville, PA**
Customer Service Associate (Shift Manager) *June 2017-August 2021*

- Promoted to shift manager in 2019 and performed closing duties that met operational standards
- Exemplified superior customer service for an average of over 45 walk-in customers per shift and quickly identified solutions to resolve customer issues as they arose

INVOLVMENT/CAMPUS ENGAGEMENT

SOA Case Competition **University Park, PA**
Semi-finalist *January 2023-April 2023*

- Created a social insurance program for the fictional country of "Storslysia" to help minimize the risk of displacement from increased climate-related threats
- Utilized quantitative modeling techniques to predict climate risks and money saved after implementing the program

Penn State Actuarial Science Club **University Park, PA**
Finance Team Member/Mentor *August 2020-Present*

- Committed to pursuing a career as an actuary and proactively seeking development and networking opportunities
- Responsible for making a dashboard in Excel to keep track of all the club's assets and liabilities

Penn State Club Golf Team **University Park, PA**
Athlete *August 2020-Present*

- Actively contributing to the club golf team by participating in tournaments, team practices, and team run volunteer activities

Penn State Department of Mathematics **University Park, PA**
Learning Assistant (Course: Calculus w/ Analytic Geometry 1) *August 2021-May 2021*

- Hosted a one-hour weekly review session to help students with calculus problems and to gain extra practice

TECHNICAL SKILLS

- Excel, R, Power BI, Python