Using Spreadsheet-Based Simulation To Evaluate The Fairness of The USGA Golf Handicap Index

Andrew Tiger  
*Southeastern Oklahoma State University*

Kashif Ur-Rehman  
*Southeastern Oklahoma State University*

Chandra Hurst  
*Southeastern Oklahoma State University*

Follow this and additional works at: [http://scholars.fhsu.edu/jbl](http://scholars.fhsu.edu/jbl)

Part of the [Business Commons](http://scholars.fhsu.edu/jbl) and the [Education Commons](http://scholars.fhsu.edu/jbl)

**Recommended Citation**


Available at: [http://scholars.fhsu.edu/jbl/vol2/iss1/23](http://scholars.fhsu.edu/jbl/vol2/iss1/23)
USING SPREADSHEET-BASED SIMULATION TO EVALUATE THE FAIRNESS OF THE USGA GOLF HANDICAP INDEX

Andrew Tiger, Southeastern Oklahoma State University
Kashif Ur-Rehman, Southeastern Oklahoma State University
Chandra Hurst, Southeastern Oklahoma State University

Using spreadsheet-based simulation, the USGA Handicap Index was shown to be an unfair statistic in one-on-one and team competition in two common types of scoring: stroke and match play. Experiments were developed in which players of different abilities (based on central tendency and variation) competed against each other over many trials. The results showed that in some situations, based on identifiable differences in abilities, some players won/lost a disproportionate (unfair) number of times. The causes of unfairness are different in one-on-one and team play. Alternative procedures were proposed that proved to reduce the unfairness of the index.

INTRODUCTION

The United State Golf Association (USGA) handicap system is a mathematical procedure that culminates in a statistic defined as a player’s handicap index. The purpose of the USGA handicap system is to make the game of golf more enjoyable by enabling individuals or teams of differing abilities to compete fairly. The two most common types of competition where handicapping is used are medal and match play. Medal play is competition based on total strokes. When player 1’s score, adjusted by the handicap index, is lower than player 2’s score, player 1 wins. For example, assume player 1 has a handicap index of 10 and player 2 has a handicap of 6. If player 1 scores an 83, the adjusted score is 73. If player 2 scores an 80, the adjusted score is 74. Player 1 wins.

Match play is competing hole by hole. The player who wins the most holes wins the match. For the two players mentioned above, player 1 receives 4 strokes. Since player 1’s handicap is 4 points more than that of player 2, player 1 receives 4 strokes on the 4 most difficult holes. On the remaining holes, the players compete with no strokes given. In team golf, scoring is slightly modified. In both medal and match play, each team’s hole-specific score is the best (minimum) handicap adjusted score. For medal play, the scores are totaled. In match play, the team with the best score wins the hole.

If these two players or teams compete using their handicap index, a fair competition would give each player or team an equal chance of winning. We propose that the Handicap Index is an unfair statistic and does not provide each player or team an equal chance of winning. In this paper, the USGA handicap index was shown to be an unfair statistic because of two specific reasons. The first is discarding the higher (poorer) scores prior to calculating the statistic. Discarding the high scores unfairly gives an advantage to the player with less variable scores. The second is specific to team golf. For team competition, the index does not account for score variability, giving the advantage to the team composed of players with high scoring variability. Initially, these two types of unfairness seem contradictory. The advantage switches from the golfer with less variability to the team with more variability. We believe this is the source of confusion within the literature, and our research eliminates the confusion.

Spreadsheet-based simulation was used to model 1-on-1 and team competition using the USGA handicap system. Experiments were developed where players of different abilities (based on central tendency and variation) competed against each other over many trials. The results showed that in some situations, based on identifiable differences in abilities, some players or teams won/lost a disproportionate (unfair) number of times.

The paper is outlined as follows. Relevant literature is summarized. Then, the spreadsheet simulation model for individual and team scoring is explained. Next, the experiment, analysis and proposal for 1-on-1 competition is covered, followed by the experiment, analysis and proposal for team competition. The paper concludes by incorporating the proposals for both 1-on-1 and team competition into a unified system for reducing unfairness and suggesting research extensions.

LITERATURE REVIEW

Several papers from 1975 to 2000 have addressed the unfairness of the USGA handicap procedure. Particularly,
Scheid (1975, 1977, and 1990) was instrumental in the modification of the handicap differential and index formulae. Pollack (2002) studied both medal and match play. Mosteller and Youtz (1992) studied the scoring of professional golfers on final tournament days and found them to be Poisson in nature. That is, scores were found to be independent, and earlier scores did not impact the final days' scores. With respect to golf handicaps, the literature can be classified into two schools of thought: advantage to the low handicap player (Scheid, 1975; Pollack, 2002) and advantage to the high handicap player (Bingham and Swartz, 2000).

Although this paper briefly addresses the influence of handicap on unfairness, stronger drivers of unfairness were discovered. Among other factors, the handicap index is a direct result of a player's score variability, and we demonstrate that a player's variability contributes more to unfairness than a player's handicap. In some of the literature, high handicap players were assumed to have high variability. The reasoning for this assumption was not provided. We disagree with this assumption. Good players, those with low handicaps, can have high variability if they are aggressive golfers and take risks. Similarly, golfers with high handicaps can be very consistent.

Tallis (1994) studied handicapping team golf primarily by creating a nonlinear function of individual handicaps. However, Tallis only studied game formats that allowed players to choose which shots, that is, scramble format, or to alternate shots within a hole. Our research does not address within hole decisions, rather, we focus on hole-to-hole decisions. Our research only focuses on stroke or match play. That is, each player must play his/her own ball for the entire round of golf. In the literature, two types of solution techniques have been applied to investigate handicap unfairness: closed-form probability theory (Pollock, 2002) and empirical studies of actual scores (Scheid, 1975). Bingham and Swartz (2000) used both methods.

A relatively new but increasingly popular method for studying stochastic systems is spreadsheet-based simulation. MS-Excel has a variety of functions, specifically a random number generating function that allows MS-Excel to be used as a powerful simulation modeling tool. This popularity is evidenced by the explosion of simulation modeling textbooks such as practical management science (Winston and Albright, 2000); simulation modeling using @Risk (Winston, 2001); advanced modeling in finance using Excel and VBA (Jackson and Staunton, 2002); spreadsheet modeling and decision analysis (Ragsdale, 2004) and operations management: a process approach with spreadsheets (Shafer and Meredith, 1998). We believe that spreadsheet simulation is ideally suited for investigating handicap unfairness, and the next section presents the spreadsheet modeling techniques to accomplish this.

Simulating Golf Scoring Using a Spreadsheet Model

Rather than using actual golf scores, scores were generated using a spreadsheet-based model. Instead of relying on actual scores, generating scores guarantees knowledge of a player's scoring ability. Relying on actual scoring data creates nonsampling error due to scoring mistakes that may be unintentional or intentional. The sport of golf relies heavily on an honor system when golfers record their scores. When this honor system is broken, nonsampling error occurs. Simulating scoring eliminates the nonsampling error. Generating golf data requires two steps: establishing a player’s handicap and running a competition between two players or teams to generate wins losses and ties.

Developing a player's handicap

A player's hole-specific scores were randomly generated by using MS-Excel functions (RAND(), NORMINV(), ROUND() and SQRT()); the player's 18-hole aggregate scoring average and standard deviation; and the average percentage of strokes used on each hole. Equation [1] displays the MS-Excel formula for a hole-specific score assuming 5.9% of strokes are used on this hole and a player's 18-hole a player’s average score and standard deviation of 84 and 3, respectively.

\[ \text{ROUND}(\text{NORMINV}(\text{RAND}(), 84*5.9\%, 3*\text{SQRT}(5.9\%)), 0) \]  \[ [1] \]

The RAND() function generates a random number uniformly between zero and one. Recalculation of the spreadsheet generates a new number. For example, if RAND() returns 0.79635965, the formula returns a player’s score of six. Similarly, if RAND() returns 0.032904798, the formula returns a score of four. A score for a complete round is the sum of the individual 18 holes.

Armed with the ability to randomly generate a player's score, the USGA handicap formula procedure was reproduced to generate a player’s handicap index. The USGA handicap index is calculated by taking 96% of the average of the best ten out of the last 20 handicap differentials. A handicap differential is computed from four elements: adjusted gross score, USGA course rating, USGA slope rating and 113 (the slope rating of a course of standard difficulty). To determine the handicap
differential, subtract the USGA course rating from the adjusted gross score; multiply the difference by 113 then divide the resulting number by the USGA slope rating. The final number is rounded to the nearest tenth. Equation [2] summarizes the formula. Reproducing this procedure in MS-Excel required using the SMALL(), TRUNC() and AVERAGE() functions. See figure 1 for an example.

\[ \text{Handicap Differential} = \frac{(\text{Gross Score} - \text{Course Rating}) \times 113}{\text{Slope Rating}} \]

**Figure 1: Spreadsheet Used to Simulate Scores and Determine a Handicap Index**

---

**Modeling competition**

Handicap development logic was duplicated for a second player, which provided two players with established handicap indices. Next, 100 head-to-head games were simulated between the two players in both medal (stroke) and match play. For each of the 100 games, the winning player received one point. In case of a tie, each player was awarded one-half point; thus 100 points were available. Each player’s winning proportion was calculated by dividing his points by 100. For example, if Player one scored 58 points, the proportion was 0.58, which meant player one wins 58% of the time. After determining each player’s winning proportion, a hypothesis test on two proportions was used to test if either player significantly wins more than the other. The test assumed a normal approximation of the binomial distribution. For all tests, \( \alpha = 0.05 \).

In team golf, scoring is slightly modified. In both medal and match play, each team’s hole-specific score is the best (minimum) handicap adjusted score. To elaborate, for each team, both players generate a (gross) score for each hole. The score for each player is adjusted based on the player’s handicap to produce a net score for each hole. In an attempt to reduce unfairness, the USGA offers an additional allowance: “Men receive 90% of course handicap; women receive 95% of course handicap” [USGA handicap system manual 9.4.b.ii, 11]. The smallest score is used as the team’s score for that hole. For medal play, the scores are totaled. In match play, the team with the best score wins the hole. No additional allowance is given by the USGA. Beyond this, the 100 head-to-head games and scoring are identical with the one-on-one competition.

The spreadsheet model is based on several assumptions. The first was that all scores followed a rounded normal probability density function as suggested by Scheid [9]. Although Scheid only focused on the
aggregate score, we assume normality on individual holes. A strength of the simulation model is that if this assumption is questioned, other distributions could be investigated. Another assumption is that a golfer’s ability, defined by an eighteen-hole scoring average and standard deviation, remains unchanged throughout the competition. No upward or downward trends in abilities occur. Again, if necessary, the model could be modified. A third assumption is that the fictitious golf courses where the two golfers compete have a USGA course and slope rating of 72 and 113, respectively. Using 72 and 113 for the course and slope rating simplifies the handicap differential and eliminates its impact on the analysis. As with the other assumptions, this assumption could also be eliminated by minor modification. For team golf, to address the additional USGA allowance, the players are assumed to be men; a final assumption is that scoring for each hole is independent of scoring on the other holes. If desired, this assumption could be eliminated if hole-to-hole correlation was of interest. Of the assumptions, we believe the last has the most potential for new research.

One-On-One Competition Experiment and Analysis

There were four independent variables in the experiment: both players’ average scores and standard deviations. The average score for both players were integers that ranged from 75 to 84 (10 levels), and the score standard deviation for both players were integers that ranged from one to three (three levels). Each of the 900 trials was collected as a record in a separate worksheet using a visual basic macro. Spreadsheet simulation software such as @risk or crystal ball would have simplified the data collection and would be recommended for those unfamiliar with visual basic. If the USGA handicap index is fair, each of the trials should level the playing field and produce statistically insignificant differences between the winning proportions of player 1 and player 2. In stroke play, 460 of the 900 trials (51%), the winning proportion was statistically different (unfair). Similarly, match play produced unfairness 52% of the time.

As summarized in the literature review, two schools of thought exist. One group believes that higher handicapped players have the advantage; whereas, the other group believes the opposite. To address this issue in this research, Figure 2 shows wins as a function of handicap. Of the statistically significant wins, lower handicap golfers won 64% of the time; thus providing some evidence that the advantage goes to the low handicap golfer. However, if unfairness was only related to lower handicaps, high handicap golfers would never generate statistically significant wins. However, high handicap golfers did statistically win 27% of the time. Further analysis revealed other causes of unfairness.

Figure 2: Breakdown of Wins Based on Handicap

![Wins as a Function of High/Low Handicap](image)

Rather than focusing on handicap, consistency, as measured with the standard deviation, created unfairness. For example, if player 1 had a standard deviation of 3 and player 2 had a standard deviation of 1, then the difference between the standard deviations is 2. The trials where the difference is produced statistically significant wins 88% and 87% of the time for stroke and match play, respectively. To elaborate, of the 900 trials, 200 instances had a standard deviation of two. In stroke play, of the 200, 175 (87.5%) resulted in a significantly higher
winning proportion, and only 25 showed no significance. Additionally, in all 175 significant wins, the player with the lower standard deviation won.

The reason for the statistically significant wins is the procedure of throwing out the high (poor) scores prior to calculating the index. Recall that the USGA handicap index is calculated by taking 96% of the average of the best 10 out of the last 20 handicap differentials. This procedure unfairly gives the advantage to the player with less variable scores. To illustrate, consider 2 players competing on a course with a USGA course and slope rating of 72 and 113, respectively. Assume player 1 is a very consistent golfer, and has scored an 80 in the last 20 rounds. Player 2 also has averaged an 80 in the last 20 rounds; however, this player is not as consistent, having scored 75 in 10 rounds and 85 in 10. If the USGA handicap system is followed, player 1 has a handicap index of 8.0, and player 2 has a handicap index of 3.0. Despite the fact that both players score 80 on average, player 2 seems to be the better golfer. If these two players competed, player 2 would have to give player 1 five strokes. Without the extra strokes, player 1 will win half the time. With the extra strokes, player 1 would clearly win half the time and tie the other half. Clearly, player 2 is at a disadvantage because the player’s more variable scoring produces a relatively lower handicap index.

When the difference in standard deviations was zero, significant wins are attributed to sampling error. For example, in one of the trials, both players have a scoring average and standard deviation of 75 and two, respectively. However, the 100-game stroke play trial resulted in player 1 winning 37 times, player 2 winning 51 times, and 12 ties. This difference was statistically significant. How could this happen? Figure 3 shows the difference in the calculated handicap index for the 100-game trials. When the line is positive, player 2 has the higher handicap. Since both players have the same abilities, player 2 has the advantage. Similarly, when the line’s value is negative, player 1 has the higher handicap and advantage. Notice that player 2 has a handicap of at least one 30% of the time. However, player 1 has a handicap advantage of at least one only 4% of the time. Therefore, due to nothing but random error, player 2 wins a significant amount of the time. This sampling error could be eliminated by running a trial of more than 100 games. However, another reason for the unfairness is that only ten of the 20 scores are used when calculating the handicap. The smaller sample size also increases sample error and can be reduced by using a larger sample size (all 20 scores).

Figure 3: One-on-One Play Unfairness

To combat the two issues addressed previously, two changes are suggested. The first is not to throw out any data, but to use a simple average of the twenty most recent handicap differentials. The second proposal is not to multiply the average by the 0.96 scaling factor. The scale factor’s purpose was to provide fairness (Scheid,
1977), and it has been modified based on empirical evidence. However, we propose abandoning the scale factor completely. With the proposed changes, only 1.6% (stroke) and 14.1% (match) of the trials generated significant wins. Figure 4 shows the significant win percentage as a function of players’ difference in standard deviation for both the USGA and proposed method.

Figure 4: The Difference in Handicap (Player 2 - Player 1) for the Trial When Both Have an Average and Scoring Standard Deviation of 75 and 2, Respectively

Team Competition Experiment and Analysis

Team competition was between two 2-player teams competing in both stroke (4-ball stroke play) and match play (4-ball match play). To reduce experiment time, player-specific attributes (scoring average and standard deviation) were not set to specific values. Rather, each player’s attributes were randomly generated from a uniform distribution. For the 18-hole average score, the limits for the uniform distribution were 80 and 90. The 18-hole standard deviation limits were 1 and 3. As with the 1-on-1 competition, a trial was composed of 100 head-to-head games between the 2 teams; thus 100 points were available. After determining each team’s winning proportion, a hypothesis test on 2 proportions was used to test if either team wins significantly more than the other. For all tests, \( \alpha = 0.05 \). The experiment consisted of 400 of these 100 head-to-head games.

Based on the 1-on-1 competition results, the USGA handicap index was expected to produce unfairness as player’s scoring variability increased. The more variable team was defined as the team with the highest summed scoring variance. However, the proposed method was expected to remove unfairness. Unfortunately, this did not occur. For team competition, the proposed method was even more unfair than the USGA method. This result is exactly opposite the result of the 1-on-1 competition in which the player with less variability had the advantage. In the team competition, variability is advantageous. To reduce this advantage, that is, unfairness, additional measures beyond modification of the statistic were needed.

A second experiment (400 trials) was run, consisting of using a trial-and-error search routine, written in visual basic for MS-Excel that identified the number of additional strokes that the more variable scoring team should give to the less variable scoring team. Additional strokes are those beyond those determined based on the 1-on-1 competition. The proposed model, given in equation 3, states that the number of additional strokes is based on the golfer-specific scoring standard deviation for both players from both teams.

\[
\text{Additional Strokes} = c_1 \sigma_{\text{HH}} + c_2 \sigma_{\text{HL}} + c_3 \sigma_{\text{LH}} + c_4 \sigma_{\text{LL}} \quad [3]
\]

The code HH indicates the high variability team, high variability player. The code HL indicates the high variability team, low variability player. The code LH represents the low variability team, high variability player. Finally, the LL represents the low variability...
team, low variability player. The coefficients for each term are represented by $c_1$, $c_2$, $c_3$, and $c_4$.

From these experiment results, a multiple regression analysis provided the math relationship between strokes and player scoring variability (see Table 2 and equation 4). Table 1 shows the MS-Excel regression output and an $r^2$ of 0.41. Equation 4 reflects that the more variable team, which consists of the HH and HL players, has an unfair advantage and should give strokes to provide fairness. To implement, the additional strokes in equation 3 must be rounded to the nearest integer. Table 2 provides an easy to use chart to identify additional strokes as a function of the players' scoring standard deviation.

Table 1: Team Competition Additional Strokes Multiple Regression Output from MS-Excel

<table>
<thead>
<tr>
<th>Regression Statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.64</td>
</tr>
<tr>
<td>R Square</td>
<td>0.42</td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.41</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.71</td>
</tr>
<tr>
<td>Observations</td>
<td>400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
<td>4</td>
</tr>
<tr>
<td>SS</td>
<td>144.6</td>
</tr>
<tr>
<td>MS</td>
<td>36.2</td>
</tr>
<tr>
<td>F</td>
<td>70.9</td>
</tr>
<tr>
<td>Significance F</td>
<td>3.02E-45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>#N/A</td>
<td>#N/A</td>
<td>#N/A</td>
<td></td>
</tr>
<tr>
<td>$c_{HH}$</td>
<td>1.220</td>
<td>0.099</td>
<td>12.4</td>
<td>6.8E-30</td>
</tr>
<tr>
<td>$c_{HL}$</td>
<td>0.85</td>
<td>0.090</td>
<td>9.4</td>
<td>4.8E-19</td>
</tr>
<tr>
<td>$c_{Lie}$</td>
<td>-1.05</td>
<td>0.099</td>
<td>-10.6</td>
<td>2.15E-23</td>
</tr>
<tr>
<td>$c_{LiC}$</td>
<td>-1.04</td>
<td>0.124</td>
<td>-8.3</td>
<td>1.24E-15</td>
</tr>
</tbody>
</table>

Additional Strokes = $1.22\sigma_{HH} + 0.85\sigma_{HL} - 1.05\sigma_{Lie} - 1.03\sigma_{LiC}$ [4]

Table 2: Additional Strokes for Team Competition as a Function of Players' Scoring Standard Deviation

A final 400-trial experiment was performed and produced the results illustrated in figures 5. In both stroke and match play, the original proposed method is solely based on the 1-on-1 results, and the modified proposed method, on the other hand, incorporates the results from the multiple regression analysis. Clearly, for team competition, the additional strokes reduced unfairness.
CONCLUSION

The USGA handicap system is a mathematical procedure that culminates in a statistic defined as a player’s handicap index. The purpose of the USGA handicap system is to make the game of golf more enjoyable by enabling golfers of differing abilities to compete fairly. In this paper, the USGA handicap index was shown to be an unfair statistic because of two procedures: (1) eliminating the highest (poorest) scores prior to calculating the index and (2) failing to account for the additional strokes needed by a team with less variable scoring.

The USGA’s reasoning for disregarding a player’s higher scores is that they “bear little relation to the player’s potential ability” (USGA, 2005). We have no argument against the index’s ability to measure ability; however, it is a poor statistic for providing fair competition between two players. Specifically, the very act of eliminating the higher scores was shown to create an unfair advantage for a player with less variable scoring.

We believe implementation is straightforward. In addition to providing the current USGA handicap index, two additional indices are required: the average and standard deviation of the last twenty handicap differentials. The former provides a measure of potential, and the latter two would be used for competition, whether in individual, team, stroke or match play.

Finally, this research is relevant because of the use of spreadsheet based simulation modeling. MS-Excel was used to simulate the stochastic system of players competing in medal and match play using both the USGA handicap system and the proposed method. Since MS-Excel is inexpensive and commonly used, future studies to evaluate or extend this research are easily implemented available.

For future research, both the 1-on-1 and team competition need to be extended beyond the experiments in this paper to address both more advanced players as well as less skilled players. An additional extension is to generalize to teams with more than two players. Another interesting extension is modifying the model for team building. Based on the current USGA method, the advantage can be quantified, and teams could use this method to select players based on their abilities. Finally, many golf tournaments require a team composed of players with varying skill levels, with the best player labeled the ‘A’ player, the second labeled the ‘B’ player, et cetera. The question becomes ‘what constitutes an ‘A’ player? Should he have a handicap of less than 5? Or, should he have a handicap of less than 8? Similar questions exist for all members. A modified version of the model could evaluate various rules to achieve the fairest tournament.

The most important extension is implementation. Once the model has been generalized beyond 2 person teams and for all relevant skill levels, the new handicapping method should be implemented on a trial basis at a golf course, hopefully with the assistance of the USGA. The implementation process would require at
least 1 year allowing golfers to learn the new system and submit scores knowing that the local course will use the new system when hosting tournaments.

REFERENCES


Scheid, F. 1975. You are not getting enough strokes. Golf Digest, August: 52.


Andrew Tiger is department chair and associate professor of management at Southeastern Oklahoma State University. He received his Ph.D. in industrial engineering from University of Houston. His current research interests include simulation, spreadsheet modeling, and golf pace-of-play modeling. He has published in INFORMS Transaction on Education, Academy of Information and Management Science Journal, Issues in Information Systems, and Global Journal of Flexible Systems Management.

Kashif Ur-Rehman is financial consultant. He received his MBA from Southeastern Oklahoma State University. His current interests include asset management, derivatives, and financial engineering.

Chandra Hurst is a current MBA student at Southeastern Oklahoma State University. She has published in Issues in Information Systems.