INCREASED COMPETITION ON THE LPGA TOUR: PROOF OF THE HYPOTHESIS?

C. Barry Pfitzner, Department of Economics and Business, Randolph-Macon College, Ashland, VA 23005, <u>bpfitzne@rmc.edu</u>

ABSTRACT

This paper documents a serendipity finding from an analysis of the determinants of success on the Ladies Professional Golf Association (LPGA) Tour. The percentage of the variance in scoring average across players on the LPGA tour that is explained by the variance in Greens-In-Regulation (*GIR*) has fallen from about 80% to near 50% in the span of the last five years. Investigation into the cause(s) of this phenomenon entertained several competing hypotheses, ultimately concluding that an increased level of competition on the LPGA tour is the determining factor in the fall in the explanatory power of a regression of scoring average on *GIR*. Regression results and F-tests serve as documentation and explanation for the fall in the power of *GIR* in predicting scoring average.

INTRODUCTION

Professional golf tours keep a variety of performance statistics presumed to measure important skills related to success. One dominant statistic is greens in regulation (*GIR*)—the percentage of golf holes for which the player reaches the surface of the green in at least two fewer strokes than the par score for that hole. Other major statistics include driving distance (*DD*), driving accuracy (*DA*) which measures the percentage of drives in the fairway of the hole being played, sand saves (*SS*) which measures the percentage with which a player takes two or fewer strokes to hole the ball from greenside bunkers, putts per round (*PPR*), and putts per green reached in regulation (*PPG*). Each of those measures, *GIR*, *DD*, *DA*, *SS*, *PPR* (or *PPG*), are related in theory to scoring and scoring is clearly related to monetary success.

The purpose of this paper is to provide empirical estimates of the effect of a single variable, *GIR*, on scoring average and to investigate why the predictive power of *GIR* has fallen in recent years for the LPGA tour.

LITERATURE REVIEW

There are several strains of research on professional golf performance based on the statistics compiled by One of the first studies of the statistical determinants of success in the PGA and LPGA tours. professional golf was by Davidson and Templin [2]. Utilizing data from the 1983 PGA (119 of the top 125 money winners) in a multiple regression framework, Davidson and Templin found that greens in regulation (GIR), putting (PPR), and a combined driving efficiency measure were capable of explaining 86% of the variation in scoring average for the PGA tour, with GIR the most important single variable. When the dependent variable was earnings, putting was slightly more important statistically than the other explanatory variables, based on standardized beta coefficients. Shmanske [8], also using a multiple regression framework for data from the 1986 PGA tour (the top 60 money winners), finds that putting and driving distance are the two most important skills in determining success on the PGA tour. When player money winnings per event are the dependent variable, he finds no significant role for GIR as an explanatory variable. Shmanske also attempts to estimate the greatest payoff for practice, and finds the greatest payoff is for putting practice. Belkin et al. [1] utilize PGA statistics for three years (1986-88) in correlation and step-wise regression frameworks. Their research confirms the importance of GIR and putts per round (PPR) as dominant variables in determining scoring average, with lesser, but statistically

important roles for driving distance, driving accuracy and sand saves. They conclude that their research confirms the importance of tour statistics in predicting scoring average.

A paper by Englehardt [4] concludes that the rankings of the top 10 money winners are *not* significantly correlated with *GIR* for 1993 and 1994 PGA seasons, and cites an increasingly important role for "total driving," which is the sum of the ranks in driving distance and driving accuracy. This study utilizes, however, a sample size of only 10. Moy and Liaw [5] find evidence that conflicts with that from Englehardt for the same PGA year. They find statistically important roles for driving distance, driving accuracy, *GIR*, and putting in determining earnings on the PGA tour for the 1993 season. The latter study utilizes a multiple regression framework and a much larger sample size than Englehardt. Moy and Liaw's work also includes analysis of the LPGA and the Senior PGA tours and they offer the general conclusion that a well rounded game is necessary for success in professional golf. Nero [6] using data from the 1996 PGA tour finds statistically important roles for driving accuracy, putting, and sand saves in determining money won. Interestingly, Nero does not include *GIR* in his analysis. Nero also estimates a frontier earnings function in an attempt to identify the most efficient golfers—that is those golfers who earn more than that predicted by the regression equation.

Dorsal and Rotunda [3] using data from the top 42 players on 1990 PGA tour found that *GIR* was the most important variable determining scoring average, and that driving accuracy was more important than driving distance. Their analysis included simple correlation analysis and multiple regression techniques. They also used scoring average, top 10 finishes, and money winnings as dependent variables. Pfitzner and Rishel [7] document the determinants of scoring average and money winnings on the LPGA tour in 2004. Among other conclusions, their finding indicate that driving distance and driving accuracy are approximately equally important in determining scoring average and money winnings for the LPGA.

METHODOLOGY

The research methods for this paper are regression analysis and simple tests for equality of variances.

The general regression model may be represented as:

$$SA_i = \beta_o + \beta_1 GIR_i + \varepsilon_i \quad , \tag{1}$$

where, *SA* = Scoring average (strokes per round) *GIR* = greens in regulation (percentage of greens reached in regulation or fewer strokes)

and the *i* subscript refers to the i^{th} observation (here the individual player)

DATA AND RESULTS

Summary Statistics on the LPGA Tour

Table I represents the summary statistics for the 2004 LPGA tour. The mean scoring average across all 164 players for 2004 was 72.88 strokes per round. Annika Sorenstam led the tour in scoring average with 68.7 strokes per round (minimum). Sorenstam also led in total money winnings (\$2.5 million) and money per event. The average money winnings per player for the year was approximately \$230,000 (this distribution is heavily skewed, of course), and player winnings per event averaged a little over \$10,000. The mean driving distance on the LPGA tour in 2004 was approximately 250 yards, and the drives of these tour players found the fairway about 70% of the time. Perhaps surprisingly, in 2004 LPGA tour

players were able to hole out from greenside bunkers in two or fewer strokes only thirty-five percent of the time. The tour also keeps putts per green-in-regulation (*PPG*). This latter statistic may be a better measure of overall putting efficacy, but it turns out that players who hit more greens have fewer putts per green in regulation (because they likely hit it closer on average).

Variable	Mean	Standard Deviation	Minimum	Maximum
Scoring Average (SA)	72.88	1.42	68.7	77.64
Money Winnings per Event (M/E)	\$10,369.85	\$15,426.45	\$139.19	\$141,372.61
Greens in Regulation (GIR)	63.89%	5.17%	48.9%	78.8%
Driving Distance (DD)	249.81	9.04	224	270.2
Driving Accuracy (DA)	70.16%	6.43%	48.2%	83.6%
Putts per Round (PPR)	30.13	0.58	28.75	31.68
Sand Save Percentage (SS)	35.29%	7.23%	18.2%	60.6%
Putts per GIR (PPGIR)	1.84	0.04	1.74	1.97
Money Winnings (M)	\$233,406.84	\$331,782.62	\$1,819.00	\$2,544,707.00
Number of Events (E)	21.05	4.20	10	31

Table I: Summary Statistics for the 2004 LPGA Tour

(n = 164)

Table II: Summary Statistics for the 2008 LPGA Tour

Variable	Mean	Standard Deviation	Minimum	Maximum
Scoring Average (SA)	72.93	1.28	69.70	79.44
Money Winnings per Event (M/E)	\$13,549.30	\$16,787.57	\$0.00	\$125,599.68
Greens in Regulation (GIR)	62.80%	4.12%	41.9%	71.6%
Driving Distance (DD)	246.58	9.55	221.5	269.3
Driving Accuracy (DA)	67.55%	5.84%	49.2%	79.80%
Putts per Round (PPR)	29.23	1.06	26.95	31.95
Sand Save Percentage (SS)	37.70%	8.42%	14.3%	60.0%
Putts per GIR (PPGIR)	1.83	0.04	1.74	1.94
Money Winnings (M)	\$327,676.60	\$420,767.88	\$0.00	\$2,763,193.00
Number of Events (E)	21.70	4.68	10	30

(n = 161)

For 2008, scoring average was almost exactly the same as in 2004, and greens in regulation actually fell by about one percentage point. Of more interest (as will be explained later), the standard deviations (the square root of the variance) of both scoring average and *GIR* were lower in 2008 than was the case in 2004. By 2008, Lorena Ochoa had taken over the top spot in both scoring average and money winnings.

Some Regression Results

This section reports regression results of scoring average on *GIR* with visual and statistical evidence of the decline in *GIR* as an explanation of scoring average.

Figures 1 and 2 show the results of regressions of scoring average on greens in regulation. Note that in 2004, the adjusted R^2 is .76, meaning that three quarters of the variance in scoring average is explained by *GIR*, and only four years later the adjusted R^2 is only .52. Visually, the "fit" for the 2004 data is

considerably better. Note further that the general spread in *GIR* is smaller for 2008. Put differently, it appears that the variance in *GIR* across players has decreased.



Figure 1: Scoring Average as a Function of GIR, 2004



Figure 2: Scoring Average as a Function of GIR, 2008

Explaining the Results

Previous research indicates that GIR are a much less important determinant of scoring average on the PGA tour (men's tour) than is the case for the LPGA. The corresponding \overline{R}^2 for the men's tour regression of scoring average on GIR is consistently near .25 over recent years. This comparison may be interpreted as GIR being roughly three times more important is explaining scoring average on the LPGA tour ($\overline{R}^2 = .76$, for the LPGA in 2004) than on the PGA tour. The fall in \overline{R}^2 to approximately .52 by 2008 on the LPGA may suggest that in some way the LPGA is becoming more like the PGA tour.

The first hypothesis entertained was hole position. It is not in dispute that hole positions are more difficult on the PGA tour than on the LPGA tour. On the men's tour, it is often advantageous for players to miss the green in a position where it is easier to make par (or birdie), than to hit the green in regulation where a three-putt is a likely outcome. Theory would suggest that if hole positions had become more difficult in recent years on the LPGA tour, that would be a potential explanation of the fall in \overline{R}^2 between 2004 and 2008. Though the LPGA does not archive the hole positions for tournaments on their website (nor does the PGA), I was able to interview LPGA players, caddies, and even an LPGA official in charge of hole placements. To my surprise, there was unanimous agreement that hole positions have not become more difficult in recent years. The caddies and players suggested that the same approximate positions were used for the courses played over time. (It is possible that as new courses are introduced to the tour, those hole positions could be more difficult.)

The second hypothesis entertained was an accidental discovery. As stated previously, examination of Figures 1 and 2 reveals that the variance in GIR seems to have fallen from 2004 to 2008. It is well known that at the extreme of no variance in the explanatory variable, the intercept and slope coefficients in a simple regression are not uniquely determined. In general, as the variance of the explanatory variable decreases, its statistical significance falls (recall that the variance of x is in the denominator of the standard error of the slope coefficient) and, of course, \overline{R}^2 falls as well. To interpret this more concretely for the current case, if players on the LPGA tour are more evenly matched in terms of GIR, then other factors will play more important roles in determining scoring average across players.

The hypothesis that the variance in *GIR* has fallen is easily testable. Since the LPGA does not report full statistics on the same number of players each year, the F-test for differences in variances was conducted for the top 125 players in terms of scoring average (the results here also pertain for all players for whom full statistics are reported). The result for the F-test is contained in Table III. The hypothesis of constant variance is rejected at an alpha level far less than .01. Between 2004 and 2008, the variance in GIR across the top 125 players on the LPGA has fallen approximately in half.

Table III: F-test for a Change in the Variance of GIR						
	GIR 2004	GIR 2008				
Mean	65.8544	64.4136				
Variance	12.71024258	6.616184516				
Observations	125	125				
df	124	124				
F	1.921083451					
P(F<=f) one-tail	0.00016132					
F Critical one-tail	1.345223605					

Some observers may not find this result surprising. In recent years many would argue that the level of competition on the LPGA tour among players has become keener. That is, players are more closely

matched. To be sure, certain players dominate in terms of wins and money, but more players are capable of competing for wins and top finishes on tour than was the case just a few years ago. To further reinforce this argument, it is also true that the variance across players in scoring average has fallen on the LPGA tour is recent years. It is interesting to note that scoring average has not fallen—players are scoring the same on average, but the differences across players have diminished.

The observed increased level of competition across players on the LPGA tour coincides with an increasing presence of Asian players (especially Korean) on the LPGA tour. The presence of the Asian players may well explain the fall in the variances of *GIR* and scoring average on the LPGA tour.

CONCLUSIONS

The explanatory power of greens-in-regulation (*GIR*) has fallen precipitously in recent years in regression models predicting scoring average on the LPGA tour. This research documents a simultaneous decrease in the variance in *GIR* across players on the tour. Players' abilities to reach greens in 2 or fewer strokes than par on a given hole are becoming more similar over the past several years. This is the likely explanation of the fall in \overline{R}^2 of such regressions. Players are also more evenly matched in terms of scoring average as well. That the level of competition between players on the LPGA tour has become closer is the practical conclusion of such statistical findings.

ACKNOWLEDGEMENT

This research was supported by a Walter Williams Craigie grant from Randolph-Macon College.

REFERENCES

[1] Belkin, D.S., Gansneder, B., Pickens, M., Rotella, R.J., & Striegel, D. (1994). "Predictability and Stability of Professional Golf Association Tour Statistics." *Perceptual and Motor Skills*, 78, 1275-1280.

[2] Davidson, J. D. & Templin, T. J. "Determinants of Success among Professional Golfers." *Research Quarterly for Exercise and Sport*, 57, 60-67, 1986.

[3] Dorsel, T. N. & Rotunda, R. J. "Low Scores, Top 10 Finishes, and Big Money: an Analysis of Professional Golf Association Tour Statistics and How These Relate to Overall Performance." *Perceptual and Motor Skills*, 92, 575-585, 2001.

[4] Engelhardt, G. M. "It's Not How You Drive, It's How You Arrive: the Myth." *Perceptual and Motor Skills*, 80, 1135-1138, 1995

[5] Moy, R. L. and Liaw, T. "Determinants of Professional Golf Tournament Earnings." *The American Economist*, 42, 65-70, 1998.

[6] Nero, P. "Relative Salary Efficiency of PGA Tour Golfers." *The American Economist*, 45, 51-56, 2001

[7] Pfitzner, C. B. and Rishel, T. D. "Performance and Compensation on the LPGA Tour: A Statistical Analysis," *International Journal of Performance Analysis in Sport*, University of Wales Institute, Cardiff, Volume 5, Number 3, December 2005.

[8] Shmanske, S. "Human Capital Formation in Professional Sports: Evidence from the PGA Tour." *Atlantic Economic Journal*, 20, 66-80, 1992.