DETERMINANTS OF SUCCESS OF GOLFERS ON THE SUNSHINE TOUR: A MULTI-EQUATION ANALYSIS

GAVIN FRASER, FERDI BOTHA AND ROBERT FRASER[1]

Abstract

There has always been a fascination with which production variables result in optimal efficiency and success. Production functions have been around for a long time but only in the mid-1980s were they used to determine the success of professional golfers. Empirical studies have shown that there is a strong relationship between golfers' scores and, hence, earnings and particular golfing skills. Initial studies of golfing success assumed a direct relationship between the success of golfers and their skill level and specified single equation models. This approach was questioned and a subsequent study developed a structural model of golfing success that used a multi-equation approach. This study will adopt the multi-equation approach in determining which golfing skills result in success for professional golfers on the Sunshine Tour. The study uses statistics taken from the Sunshine Tour website for the 2011 – 2013 seasons for the top golfers to reveal the determinants for success. Many studies in this area have covered the PGA Tour and the European Tour, however there has been no studies done on the Sunshine Tour. The study confirms the indirect relationship between the golfer's scores, ranking and earnings. A low average score is associated with a better overall ranking and hence higher earnings.

JEL Codes: D24, Z21, L83

Keywords: Sunshine Tour; golfing skills; earnings; determinants of success; multi-equation analysis

1. INTRODUCTION

Production functions are a microeconomic phenomenon which takes a group of variables and analyses the input flows and compares it to the end result, the output. Whilst it may not seem obvious, production functions are widely used in a sporting framework. As with any production function there is a transformation process which takes place between the input variables and the outcome. Many different variables are used for production functions in sport, these can include taking players' technical abilities (goals scored, wickets taken, putts per round etc.), players' features (height, weight, age), teams' spending and revenue, and other aspects such as weather conditions to name but a few. Some of these variables go through the transformation process during the performance of the athlete or team. During a match, race, or round the athletes' inputs of their abilities are being transformed into the resulting output which is the performance at the end of the fixture (Lewis, 2003).

There has been extensive fascination with trying to predict sporting fixtures and determining which individual factors result in the most success (Callan and Thomas, 2007). This has led to growth and greater emphasis in the field of sports economics. When it comes to sport the performance of a team is dependent on the collective performances of the individuals in the team. This means that players could experience a successful result even if they perform below standard due to other members in the team performing well. This is not the case with individual sports as there is no one to 'pick up the slack' if the individual does not perform well. The performance is all on the individual and whether he/she performs well determines the amount of success experienced. With individual sports including racing, athletics, tennis and golf there is a 'team' in essence behind a particular athlete (coaches, dieticians, nutritionists, etc.). Nevertheless, when competing the result all depends on the performance of the athlete, and to a certain degree the performance of the competition.

In the case of sports, the 'human' aspect changes the dynamic of the production function. Production functions are more likely to change over time which is not the case with such aspects as machinery in the production process as the resulting performance is generally consistent and predictable. The human aspect is a widely debated topic in economics, particularly when it comes to consumer behaviour, as approaches consider people to be predictable and reliable when in reality it can be the opposite (Hendry and Mizon, 2013). With production functions constantly changing it is important to narrow it down to specific categories, as a production function in one sport may not be applicable for another sport. Also, different forms of a particular sport could be further narrowed down, for instance; cricket could be split into test

[1] Department of Economics and Economic History, Rhodes University, Grahamstown, South Africa.

match performances and limited-overs performances (Brock, Fraser and Botha, 2012). This means that production functions in sport are usually unique and are only applicable in certain criteria.

1.2 South Africa and Golf

Sports economics is somewhat of a lesser practiced area of economics in South Africa (Brock, Fraser and Botha, 2012). This is even more the case when it comes to analysing golf in South Africa from an economic perspective as there have been no studies done on the major Sothern African golfing tour, the Sunshine Tour. This is in contrast with its American counterpart, the US Professional Golf Association (PGA) Tour. There have been a number of studies done on the various golfing tours in America, these include the PGA tour (Davidson and Templin, 1985), the Ladies PGA tour and the Seniors PGA tour (Moy and Liaw, 1998). There have also been studies done on the other major tour for professional golfers, the European tour (Oliviera, 2013).

It is understandable the lack of research done on South African sports economics with, particularly, the Sunshine Tour being less marketable and popular than its global counterparts. Less top ranked players are registered for Sunshine Tour events due to the profitability the other tours offer, this is evident with the highest ranked South African player in 2013 (Charl Schwartzel) only playing 4 events on the tour. That being said, the Sunshine Tour is still a very popular tour with golfers from all over the globe being represented. Furthermore it is a tour where many prospective South African golfers can make their name.

With no previous empirical studies on this area of research, this paper will be the first of its kind. The paper examined the determinants to success of professional golfers on the Sunshine Tour. Furthermore, a comparison of the performances of players who are cardholders on the European Tour to players who are not cardholders on the European Tour was done. There are many different methods of study and given that no previous studies had been done on the Sunshine Tour it meant that any method would be helpful.

Founded in the 1960s and previously known as the South African Tour or the FNB Tour, the Sunshine Tour rebranded itself in 2000 in attempt to increase its popularity and broaden its appeal. This resulted in the tour being extended to events outside of South Africa as it became a men's professional golf tour based in Southern Africa. The Sunshine Tour still holds most of its events in South Africa, however, events also take place in Namibia, Swaziland, Zambia and Zimbabwe.

The Sunshine Tour currently co-sanctions 6 events with the European Tour, these include the Africa Open, The Alfred Dunhill Championship, the Joburg Open, the Nelson Mandela Championship, the Tshwane Open and the South African Open. This serves as the launch of the European Tour each season. The South African Open is the second oldest open championship in the world, after the British Open.

2. LITERATURE REVIEW

The use of production functions has been used in a number of studies in many different sports. Such sports include; cricket (Brock, Fraser and Botha, 2012), baseball (Marburger and Reynolds, 1994), ice-hockey (Kuosmanen, 1998), basketball (Scherzer, 2010) and soccer (Haas, 2003), among others. The literature review acknowledges that whilst golf is an individual sport and the aforementioned are regarded as team sports, there is still relevance when it comes to studies which use production functions. There have been many studies which have used a production function to measure and analyse athletes' ability from a statistical standpoint, such studies will be mentioned below. Such measures are sometimes applicable to other studies despite being previously used in a different sport.

Production functions can be very similar despite the sport, with only a few inputs being changed. For example a study done on ice-hockey used the 'non-parametric' method of data envelopment analysis (DEA) to analyse teams' performances in the National Hockey League (Kuosmanen, 1998). Barros and Garcia (2006) used a similar framework, however, their study analysed the technical efficiency of 12 clubs in the English Premier League. This analysis method has been extended to individual sports as a study viewed the career efficiency of 229 professional tennis players playing on the Association of Tennis Professionals (ATP) tour from 1991 to 2012 (Halkos and Tzeremes, 2012). The model used one input (career matches) and nine different outputs (statistics which determine performance) of players' performance on the three different types of court surfaces (grass, clay and hard courts).

Production functions can be used to investigate an individual's performance in a team sport; however, it has mostly been used to view a combined collective impact of a team. While production functions for individual sports, such as singles tennis and golf, only look at the results of a particular person's effect. The result in individual sports comes down to the player, whereas in team sports, players could have a poor performance and still have favourable results because of other players' performances on their team. This means that overall results in individual sports are a better and more accurate reflection of a player's efficiency and performance (Halkos and Tzeremes, 2011).

To date there has been a number of production function analyses pertaining to the success of golfers. With regards to golf, the development of the production function method has been rather slow and repetitive. From the mid-1980s to early 2000s, most production function models were identical in the approach which was used (a single equation approach). Whilst many researchers acknowledged flaws in the approaches used they failed to change or develop their own approach. This meant that much of the research done was the same with the exception of the tour (PGA, LPGA, Senior PGA or European tour) or the use of different variables in the models (sand saves or scrambling, for example). The single equation approach models did have results which were proven to be quantitatively reliable and to be accurate with actual results (Watkins, 2008).

Davidson and Templin (1986) have played a major role when it comes to analysing the performances of professional golfers, with many other researchers using their framework as a focal point. Davidson and Templin (1986) took stepwise single-equation regression models to analyse the effects of a player's golfing abilities (shot-making skills) on their earnings (earnings reveal tournament success). The reduced form production function can be generalised as follows:

$$\text{OUTPUT}_i = f(\text{SKILL}_i, \text{EXPER}_i) \quad (1)$$

where OUTPUT_i is the ith player's tournament success, SKILL_i provides a vector of shot making skills for player i, EXPER_i is player i's years of experience.

Using their framework they simply look at why some professional golf players are more successful than their counterparts. Their assumptions were that success stems from an individual's superior skills along with being adept in an array of skills (driving distance and greens in regulation, for example) leads to better tournament scoring and therefore greater success. The findings from using the official statistics from the 1983 PGA tour supported their initial assumptions. While many subsequent studies followed the procedures used by Davidson and Templin, it should be noted that Davidson and Templin were one of few who used a stepwise regression procedure to determine their final specification. Subsequent studies which were similar to that of Davidson and Templin include Belkin, Gansneder, Pickens, Rotella, and Streigel (1994), Moy and Liaw (1998), Shmanske (2000), and Rishe (2001).

In more recent times, a multi-equation approach has been employed which takes into account a number of aspects to a golfers success. One of the main differences being that shot-making skills were previously viewed to have a direct effect on tournament earnings, whereas in the multi-equation model shot making skills are regarded to indirectly affect tournament earnings. The notion of a multi-equation approach was set by Scully (2002) who acknowledged it but failed to implement it. The multi-equation approach was finally used by Callan and Thomas (2007) and to some extent Watkins (2008). Callan and Thomas argued Davidson and Templin's method by stating that stroke play abilities do not have a direct effect on a player's tournament earnings. They go on to say that instead shot making abilities have a direct effect on a player's score which in turn has an effect on a player's ranking and earnings. Thus concluding that shot making ability has an indirect effect on golfers' earnings. The relatively newer multi-equation approach has been used far less than that of the single equation approach, for obvious reasons. Both models have positive points; with the single equation approach being simpler to use and the multi-equation approach being more accurate due to its complexity.

From the review, it is evident that studies can help reveal the determinants of success in golf with quite some accuracy as shown by many studies having similar results and conclusions.

3. EMPIRICAL STUDY

3.1 Theoretical model

The basic premise of the multi-equation model is that the golfer's shot making skill does not directly determine the total earnings for the season. The shot making skill of each player determines the intermediate outcomes of the average score and the ranking, which in turn determines the aggregate earnings. In order to participate, players need to qualify for the limited number of places in a tournament. The players need to perform at a satisfactory level during the event to post low scores that improve their ranking and eventually lead to earnings. The prize money for each tournament is shared between the golfers who 'make the cut', which is usually the lowest 70 cumulative score over the first two rounds of the tournament. The prize money earned will be determined by the players with the lowest scores after completing the final two rounds. The players' skills and experience over the duration of the tournament, with the proviso that they survive the cut, determines their position in the field, their ranking and finally their financial gain from prize money.

The underlying structural model, as proposed by Callan and Thomas (2007), for the production of the season's aggregate earning for each player is as follows:

$$\text{SCORE}_i = f(\text{SKILL}_i, \text{EXPER}_i), \tag{2}$$

$$\text{RANK}_i = g(\text{SCORE}_i, \text{EVENTS}_i), \tag{3}$$

and

$$\text{EARN}_i = h(\text{RANK}_i, \text{TOURNC}_i) \tag{4}$$

where SCORE_i is player *i*'s average tournament score, SKILL_i provides a vector of shot making skills for player *i*, EXPER_i is player *i*'s years of experience, RANK_i captures a player *i*'s overall tournament ranking, EVENTS_i is the number of tournaments entered by player *i*, EARN_i is player *i*'s aggregate tournament earnings, and TOURNC_i is the number of tournaments player *i* completes over the season.

3.2 Empirical Model and Data

The empirical specification model is the following recursive system of estimating equations for each player in the data set:

$$\text{SCORE}_i = \alpha_0 + \alpha_1\text{DRIVEACC}_i + \alpha_2\text{GIR}_i + \alpha_3\text{PUTT}_i + \alpha_4\text{EXP}_i + \alpha_5\text{EURCD}_i + \alpha_6\text{ROUNDS}_i + \varepsilon_{1i} \tag{5}$$

$$\ln(\text{RANK}_i) = \beta_0 + \beta_1\text{SCORE}_i + \beta_2\text{EVENTS}_i + \varepsilon_{2i} \tag{6}$$

and

$$\ln(\text{EARN}_i) = \delta_0 + \delta_1\ln(\text{RANK}_i) + \delta_2\text{TOURNC}_i + \varepsilon_{3i} \tag{7}$$

where SCORE_i is the average of player *i*'s score of all official rounds; DRIVEACC_i is the percentage of player *i*'s drives on the fairway; GIR_i is the percentage of greens player *i* hits in the regulation number of shots; PUTT_i is the average number of putts for GIR for player *i*; EXP_i is the number of years player *i* has been a professional golfer; EURCD_i is a dummy variable where player *i* is a European Tour card holder or not; ROUNDS_i is the number of rounds played by player *i* per tournament during the season; RANK_i is the average ranking for player *i* over the course of the season; EVENTS_i is the number of events entered by player *i* during the season; EARN_i is player *i*'s aggregate earnings over the season; TOURNC_i is the number of tournaments entered by player *i* in the season; and ε_{1i}, ε_{2i} and ε_{3i} are the respective disturbance factors for each equation. The *a priori* expectations of the relationships between the dependent and independent variables are shown in table 1.

Table 1: *A priori* expectations of the relationships between the dependent and independent variables

Variables	Score	Earnings
Driving accuracy	-ve	+ve
GIR	-ve	+ve
Putts	+ve	-ve
Experience	-ve	+ve
Rounds per tour	-ve	+ve
European cardholder	-ve	+ve
Events		+ve
Tournaments completed		+ve

The study looks at the 2011 - 2013 season ending statistics of the top golfers on the Sunshine Tour. Summary statistics of all the variables are provided in Table 2. All the statistics are collected from the official Sunshine Tour website. The majority of the foreign professional are not included in the study as the full spectrum of statistics was not available for them. Therefore the players in the included are mainly southern African players. Some of the variables utilised in other studies, such as the 'sand save' percentage and driving distance, will be omitted from the findings. This is due to these statistics only being available for a small proportion of the professionals on the circuit. Although the sand save percentage potentially could have a significant impact on the earning ability of a player, the lack of sufficient data the variable was removed from the model. In addition, it was not used in the majority of previous studies. Driving distance had a similar problem with no data being available for 2012 and 16 and 27 players having their driving distance recorded for 2011 and 2013 respectively. It is also doubtful how representative the distances are as they were recorded over a small number of holes. As the majority of the available driving distances were in excess of 280 metres, it was considered that driving accuracy would be more important in shooting low scores.

Table 2: Summary statistics

Variable	Mean	Standard deviation	Minimum	Maximum
Rank	72.38	45.64	1	207
Log(Rank)	3.98	0.92	0	5.33
Average score	71.93	1.38	67.81	75.83
Driving accuracy (%)	59.60	8.48	25	89.29
GIR (%)	67.84	5.71	44.44	86.11
Putts (ave)	1.80	0.06	1.64	2.13
Tournaments	17.67	7.24	1	26
Tournaments completed	9.32	5.16	1	24
Earnings (R)	452 256.00	676 923.20	2 880	5 699 739
Log(Earnings)	12.29	1.24	7.97	15.56
Experience (yrs)	8.45	6.88	0	33
Rounds per tour	2.64	0.65	1.08	4.00
European cardholder	0.25	0.43	0	1

What is evident from Table 1 is that there is a great deal of variation in some variables across players but in others the variation is relatively small. While the mean annual earnings is R452 256, the minimum earnings is R2 880 compared to the maximum of almost R5.7 million. The average score is almost 72 shots per round, but the difference between the highest and lowest is only 8 shots, while driving accuracy (average 59.6) varies between 25% and 89% and greens in regulation, which has a mean of 68% of greens being hit in the required number of strokes, varies from 44% to over 86%.

4. EMPIRICAL RESULTS (work in progress)

Two regression approaches are used. The first is a single-equation OLS regression that estimates the determinants of annual earnings. This approach is similar to that adopted by the original studies determining the success of professional golfers where they assumed a direct link between the output (earnings) and the determinant variables. The second approach is analogous with the method promulgated by Callan and Thomas (2004) and uses three-stage-least squares (3SLS) to estimate a set of simultaneous equations. The purpose of running the two approaches is to be able to compare the results of the models.

The earnings equation is specified as:

$$\ln(earnings_i) = \beta_1 + \beta_2 driveacc_i + \beta_3 gir_i + \beta_4 putts_i + \beta_5 experience_i + \beta_6 ecard_i + \beta_7 tournc_i + \beta_8 roundst_i + \beta_9 year_i + \varepsilon_i$$

Table 3 reports the OLS results regressing earnings on a range of potential determinants. The *a* priori expectation of the relationships between earnings and the independent variables are as expected, except for driving accuracy. With the exception of driving accuracy and experience, all the variables showed a significant relationship with the level of earnings. The fact that driving accuracy is not significant is surprising, given that greens in regulation and putts for GIR are highly significant. It is expected that accurate drives from the tee would make approach shots to the green easier. Hitting more greens in regulation is related to higher earnings, whereas each additional putt made from hitting the green in regulation is associated with a massive 532% drop in earnings.

Table 3: OLS regression predicting the log of earnings

Variables	Log of earnings	
	Coefficient	Standardised Beta
Driving accuracy	-0.0043 (0.0055)	-0.0293
GIR	0.0424*** (0.0099)	0.1953
Putts from GIR	-5.3236*** (1.0030)	-0.2375
Experience	-0.0032 (0.0045)	-0.0176
Tournaments completed	0.0786*** (0.0084)	0.3269
Rounds per tour	0.6863*** (0.1098)	0.3604
European cardholder	0.7172*** (0.1194)	0.2501
Intercept	16.1160*** (1.6829)	-
R^2	0.7616	
Observations	335	
F-statistic	105.7***	

Note: Year dummies included but not shown. Robust standard errors in parentheses. p < 0.001***, p < 0.05 **.

Table 4 contains the 3SLS results. While driving accuracy and experience have no significant impact on the average score, hitting more greens in regulation decreases a player's average score. Similarly, sinking more putts from greens hit in regulation has a large effect on the average score: each additional putt made from GIR increases the average score by about 8.34 strokes. Players holding a European tour card score on average 0.87 strokes lower than non-cardholders. The average score has a large impact on a player's ranking: For each additional stroke made on average, a player's ranking worsens by about 59%. Finally, a better ranking is related to higher earnings. Specifically, earnings rise on average by 2.13% for each improved ranking position. Completing on additional tournaments is associated with a 4.45% increase in earnings.

What the 3SLS results seem to suggest is that in order to maximise earnings, golfers need to improve their ranking and complete more tournament by making the cut. Ranking is improved by playing in more tournaments and, unsurprisingly, shooting lower scores on average. In order to attain a lower average score, a golfer needs to hit the greens in regulation and minimise the number of putts made from hitting the green in regulation. Since driving accuracy is not a significant predictor of the average score, iron play and the short game, in particular putting, seem to predict a player's score (i.e. long game is not as important as the short game). Players also need to complete as many tournaments as possible, which implies making the cut more often. This in turn would mean that a player plays more rounds per tournament, and the more rounds are played the greater the chances of lowering the average score over the number of rounds.

Table 4: 3SLS regression results

	Dependent variable		
Variable	Average score	Log(Rank)	Log(Earnings)
Driving accuracy	-0.0038 (0.0055)		
GIR	-0.0644*** (0.0085)		
Putts from GIR	8.5568*** (0.7834)		
Experience	-0.0036 (0.0061)		
European cardholder	-0.8127*** (0.1172)		
Rounds per tournament	-0.7288*** (0.0743)		
Average score		0.5876*** (0.0119)	
Tournaments		0.0197*** (0.0022)	
Ranking			-0.0212*** (0.0009)
Tournaments completed			0.0467*** (0.0076)
Constant	63.3062*** (1.5183)	-38.6344*** (0.8454)	13.3964*** (0.1145)
R^2	0.7086	0.9108	0.6918
Chi-squared	874.1***	3276.3***	825.8***

Note: Robust standard errors in parentheses. Year dummies included but not shown.
$p < 0.001$***.

Work in progress – to be competed
Detailed write-up of the results of the regressions
Determination of the value of marginal product of the skill variables.
Comparison of the two approaches in terms of their VMPs

5. CONCLUSION

To be completed.

LIST OF REFERENCES

BARROS, C.P. and GARCIA, P. (2006). *Efficiency Measurement of the English Football Premier League with a Random Frontier Model*. Portugal: ISEG.

BELKIN, D.S., GANSNEDER, B., PICKENS, M., ROTELLA., R.J. and STREIGEL, D. (1994). Predictability and stability of professional golf association tour statistics. *Perceptual and Motor Skills, vol. 78: 1275-1280.*

BROCK, K., FRASER, G. and BOTHA, F. (2012). A Production Function for Cricket: the South African Perspective. *J.Stud.Econ.Econometrics, 36(2): 1-15.*

CALLAN, S.J. and THOMAS, J.M. (2007). Modeling the Determinants of a Professional Golfer's Tournament Earnings: A Multiequation Approach. *Journal of Sports Economics 8: 394-410.*

DAVIDSON, J.D. and TEMPLIN, T.J. (1986). Determinants of success among professional golfers. *Research Quarterly for Exercise and Sport, 57: 60-67.*

HAAS, D.J. (2003). Technical efficiency in Major League Soccer. *Journal of Sports Economics, 4(3): 20-215.*

HALKOS, G. and TZEREMES, N. (2012). *Evaluating professional tennis players' career performance: A Data Envelopment Analysis approach*. Greece: MPRA.

HENDRY, D.F. and MIZON, G.E. (2013). *Unpredictability in Economic Analysis, Econometric Modeling and Forecasting*. Working Paper. Oxford University: Department of Economics.

KUOSMANEN, T. (1998). *Efficiency of hockey teams in the NHL*. Helsinki, Finland: Helsinki School of Economics and Business Administration.

LEWIS, M. (2003). Moneyball: The Art of Winning an Unfair Game. United States of America: W.W. Norton & Company Inc.

MARBURGER, D. and REYNOLDS, M. (1994). Are baseball players paid their marginal products? *Managerial and Decision Economics, 15: 443-457.*

MOY, R.L. and LIAW, T. (1998). Determinants of professional golf tournament earnings. *American Economist, 42: 65-70.*

RISHE, P.J. (2001). Differing Rates of Return to Performance: A Comparison of the PGA and Senior Golf Tours. *Journal of Sports Economics, 2(3): 285-296*

SCHERZER, F. (2010). *On the Value of Individual Athletes in Team Sports*. Germany: HHL.

SCULLY, G. (2002). The Distribution of Performance and Earnings in a Prize Economy. *Journal of Sports Economics, 3(3): 235-245.*

SHMANSKE, S. (2000). Gender, Skill, and Earnings in Professional Golf. *Journal of Sports Economics, 1(4), 385-400.*

TIAN, G. (2007). *Implementation of Pareto Efficient Allocations*. United States of America: Texas A&M University.

WATKINS, J.R. (2008). Drive for show, putt for dough: rates of return to Golf skills, events played, and age on the PGA Tour. *Michigan Journal of Business, 1(1): 35-59.*