MIT Sloan Sports Analytics Conference 2012
March 2-3, 2012, Boston, MA, USA

Optimal Design of the PGA Tour; Relegation and Promotion in Golf

Martin L. Puterman
Sauder School of Business, University of British Columbia,
Vancouver, BC Canada V6T 1Z2

Qingchen Wang
Centre for Computational Statistics and Machine Learning, University College London
London, England, WC1E 6BT

Email: martin.puterman@sauder.ubc.ca; qingchen.wang@yahoo.com

Abstract

This paper investigates how best to reconfigure PGA TOUR end-of-season "tour card" allocation rules in order to enhance tour competitiveness. It provides an empirically based mathematical/simulation framework to investigate alternatives to current allocation rules in which players were represented by their age, a career skill level transition curve and the expected number of tournaments they compete in. For each rule, 20 replicates of 1000 seasons were simulated. At the end of each simulated season, players were reassigned to tours on the basis of the end of season money lists. The relative quality of the rules was compared on basis of the number of players correctly assigned to PGA TOUR and NW based on their skill levels. PGA TOUR competitiveness was maximized when the top 150 players on the PGA TOUR money list retain full PGA TOUR status, the top 30 players on the NW money list are promoted to the PGA TOUR, and Q-School does not promote players directly into the PGA TOUR; a considerable change from current practice. The approach described in this paper is suitable for investigating other end of season playoff structures under consideration by the PGATOUR and also other professional sports tour structures.

1 Introduction

This paper investigates how best to configure the end-of-season "tour card" or exemption rules used by the PGATOUR, the Nationwide Tour (NW), and Q-School so as to enhance competitiveness. Presently, at the end of each season, the top 125 players on the PGA TOUR money list retain full PGA TOUR status, the top 25 players on the NW money list are promoted to the PGA TOUR and the top 25 finishers (plus ties) from Q-School receive PGA TOUR status, while the remainder of the players either are relegated to NW status or have limited status. As noted by Puterman and Wang [1], this may be regarded as a special case of relegation and promotion (RP) – a concept used in most European sports leagues to maintain the competitive quality of the league by replacing weaker teams in upper tiers with stronger teams from lower tiers at the end of each season. Recently, this similarity has also been noted by NY Times golf writer Karen Crouse [2], who in describing Q-School wrote on December 4, 2011: "Results, not reputations are rewarded, making the PGA Tour the nearest thing in American sports to the promotion and relegation system in English soccer".

The basic premise of this paper is that the RP rules underlying the PGA TOUR, NW, and Q-School structure do not achieve the PGA TOUR objective of ensuring that the PGA TOUR has the best players and is as competitive as possible. Consequently some players on the NW tour might actually be better than those on the PGA Tour. The reasons for this belief are threefold:

- Objectives are not clearly stated.
- This structure has evolved over time and been adjusted on an ad hoc basis on several occasions.
- Golf traditionalists are resistant to change.

Even if the PGATOUR objectives were clearly stated, an experimental approach of changing the system from time to time will not achieve them. To do so requires system modelling and evaluation. That is what this paper seeks to achieve. It develops a mathematical model of the PGATOUR, NW, and Q-School system, uses statistical methods to estimate model parameters and simulation to compare alternative RP structures. It enables evaluation of changes with respect to a range of metrics which are also developed in this paper. It extends the author's previous research on relegation and promotion (Puterman and Wang [1]).

2 Problem Definition

In this study, our goal is to evaluate the ability of different relegation and promotion methods to assign players into the correct tiers of the PGA TOUR. The method used in 2011, which was described above, combines seasonal performance with an open tournament to promote players. Two errors are possible:

- Players are promoted or relegated who shouldn't be.
- Players are not promoted or relegated who should be.

Furthermore, if players who earned PGA TOUR cards via the Nationwide Tour generally performed well in subsequent years, while players who earned PGA TOUR cards via the Q-School tournament generally performed poorly in the next year, a structural change to remove direct PGA TOUR promotion from Q-School may be warranted. In this paper, we develop a mathematical model and use simulation to optimize the existing relegation and promotion structure. These methods can also

be used to investigate the potential of some new end of season playoff structures under consideration by the PGA TOUR and apply to other golf tours and sports as well.

Model Components: We adapt the model of Puterman and Wang [1] to this setting. Its key features are teams, intrinsic (unobservable) skill levels (ISLs), ISL dynamics and match determination models. In it, a season consists of a sequence of matches determined by a schedule which specifies the number of matches and the extent of inter-divisional play. After completion of a season, a team's performance in its division is used to determine whether it is relegated or promoted or remains in the same division. In their model, the PGA TOUR would correspond to Division 1 and the Nationwide Tour to Division 2.

In the golf setting, which we study here, players replace teams and stroke play tournaments replace matches. Also we must account for player aging and career skill trajectories when modeling ISL dynamics. We assume no inter-divisional play although in reality NW players infrequently compete on the PGA TOUR and vice versa. Also there are non-tour players who compete in some events.

Players: A *player* is represented by four entities: his age each season, his ISL, his tournament participation percentage and his career progression curve which we represent by

$$ISL_{Age} = a + \beta\, Age + \gamma Age^2 + \varepsilon_{Age}$$

where a , β and γ are player specific parameters, Age is his age at the start of the season and ε_{Age} represents unexplained variation modeled as a $N(0, \sigma^2)$ random quantity.

Some comments about this model are in order. The ISL represents a player's true but unobservable ability *if he competed on the PGA Tour*. Since it was observed that on average first year scores of NW players who competed on the PGA Tour were .95 higher (See Appendix 1), NW scores were adjusted accordingly.

We assume that we observe a realization of the ISL through the player's average adjusted end of year scoring per round average. In 2011, it varied between 69 and 73. We assume that the ISL evolves from year to year according to a quadratic curve with a player's ability peaking in his early 30's. We estimate the model parameters for each player for which we have sufficient data. Note that Berry et. al. [3] and Connolly and Rendleman [4] consider other models of player dynamics. Figure 4 in Berry et. al. [3] and our empirical investigation suggest that a quadratic model may well describe dynamics. Further it is well suited to our simulation methodology. Recent technological advances and better conditioning suggest that players' skills may decline more slowly than suggested by a quadratic and a piecewise linear, a piecewise quadratic or a spline model may be more suitable.

Evaluation Metrics: In light of two possible errors described above, and the pre-eminence of the PGA TOUR, we choose *the number (or percentage) of players on the NW TOUR whose start of season ISL exceeds that of a player on the PGA Tour* as our primary metric. Since we assume 180 players on the PGA TOUR, this quantity can vary between 1 and 180. In practice this quantity is unobservable, but our simulation approach allows us to compute it each season. To compute this quantity, we rank all players by skill, and count the number of Nationwide players with higher ISLs than that of any PGA TOUR players. Although in reality, it is possible for a lower skilled player to outperform a higher skilled player, and since relegation and promotion is only based on performance, it is likely to assign some players

incorrectly. In the long run some RP schemes would on average assign more players into correct tiers than others. We evaluate relegation and promotion methods and optimize their parameters based on this long run average.

An alternative metric which we have used previously and has been used by other authors in other settings is the standard deviation of skill levels on each tour. This was a useful comparator in more general circumstances when considering a league system with more than two tiers and in which the number of teams per division and the schedule structure were variables in addition to the RP rules.

Tours and tournaments: We assume two tours, the PGA Tour consisting of 180 full-time and 60 part-time players, and the Nationwide Tour consisting of 140 full-time and 80 part-time players. In addition there is an end of season Q-School tournament. A season consists of 45 PGA TOUR tournaments (156 players) and 30 NW tournaments (144 players); each tournament is made up of 4 rounds. All tournaments on each tour are identical with respect to difficulty and purse. The current PGA TOUR and NW prize structure (First prize is 18% of the total purse, 2nd prize is 10.8% of total, 3rd place is 6.8% and so on down to 70th place which is 0.2% of total), which has been in use since 1979 (http://frankosport.com/golf/Purse/index.html#PM01) is used to allocate awards and create the money list. All tournaments have the same prize structure. Since not all players play in all tournaments, we use the player's probability of entering a given tournament to sample tournament entry. We exclude the possibility of a non-tour player entering any tournament via qualification events or through various exemptions and ignore the effect of "Majors" and limited field events although we speculate on their impact below.

We assume that each year, a player's tournament round scores are independent with mean *ISL* and variance σ_t^2 where ISL denotes his start of season ISL and σ_t^2 is his between round variance. Empirically, we observed that this variance was increasing with the ISL and the data was skewed. Because of this we stratified players into four groups on the basis of the ISL ([68, 70.5], (70.5, 71.5], (71.5, 72.5], (72.5, 75]) and sampled round scores from the appropriate empirical distribution. We assume a cut is made after two rounds using rules described below.

Relegation and promotion: Our primary focus is on optimizing the parameters of the existing PGA TOUR/NW/Q-school structure. *The existing model* relegates and promotes players using the rules described in Appendix 2. A more detailed description can be found in Figure 1 of Connolly and Rendleman [5].

In this research we investigate the impact of modifying the numbers of players who retain PGA TOUR status, N_{PGA}, the number promoted from the NW, N_{NW}, and the number who receive PGA TOUR cards for the successive year from Q-School, N_Q. At present $N_{PGA} = 125$, $N_{NW} = 25$ and $N_Q = 25$ (plus ties). Thus our problem can formally be stated as that of choosing values for (N_{PGA}, N_{NW}, N_Q) so as to minimize the number of NW players who have higher ISLs than PGA TOUR players.

3 Simulation Methodology

A season/year is simulated as a series of 4-round stroke-play tournaments. At the conclusion of each season, players are relegated and promoted based on Q-School and the final money lists for the two tours. Ages and skill levels are updated using player specific career curves. Players over 50 leave the tour and new players are generated by sampling from empirical distributions of ISL, age, and career

progression curve. We simulate 1000 years using each relegation and promotion rule that is a specified value of (N_{PGA}, N_{NW}, N_Q) and performance metrics are accumulated. For statistical evaluation purposes, the first 100 years of every simulation is removed as a warm-up period and we perform 20 trials for each RP configuration. Details follow:

Parameter estimation and data: To develop player's skill level models we analyzed all PGA Tour and Nationwide (and its predecessors) end-of-season player performance between 1998 and 2010. Data was obtained from relevant sections of www.pgatour.com/r/stats, www.espn.go.com/golf and sports.yahoo.com/golf. Empirical distributions were created using MS Excel's data analysis package, regression models were fit using NCSS, and career progression curves were fit using a Matlab script.

Initialization: Create initial player pools for the PGA and NW Tours:

- Sample 180 PGA and 140 NW player's ISLs from the historical distribution of ISLs on the PGA and Nationwide tours.
- Sample 180 PGA and 140 NW player's ages in each tour from the historical distribution of ages on the PGA and Nationwide tours.
- Sample players parameters a, β, γ and σ^2 from a multivariate distribution derived from existing data. Resample if player score deviates significantly from curve value.

Simulate a Season: A season is simulated as a series of four-round stroke-play tournaments. Tournaments are simulated on a per-round basis. Tournament fields are generated by sampling full and part-time players according to their playing probabilities. If a full field is not attained, extra players are added with ISLs that are chosen to represent those of non-exempt players who have competed in the past in each tour. The mean of these ISLs was 73.4 for the PGA and 73.8 for NW.

Tournaments are simulated as follows:

- A score is sampled for each player based on his ISL and his between round variance. Sampled scores are rounded to the nearest integer.
- After scores are generated for the first two rounds of the tournament, a cut takes place following the rules employed by the PGA and NW tours (top 70 scores and ties for PGA, and top 60 scores and ties for Nationwide).
- Finally, after scores for all four rounds are generated, players are ranked by their total scores and awarded money based on their positions.
  - Ties for first place are resolved through random sampling.
  - Players are assigned prizes based on the prize money distribution.

After all tournaments have been simulated, an end of the season money list is constructed.

End of season updates: The end of season money lists are used to relegate and promote players based on the rules under consideration. Q-School is simulated by drawing the required number of graduates then probabilistically assigning them attributes to agree with historical distributions..

Players ages are incremented by 1 and the new ISL is drawn from a $N(a + \beta\, Age + \gamma Age^2, \sigma^2)$ population where the player specific parameters are used in computing the mean. All players retire from play once they reach the age of 50 (i.e., the age of 49 is the last year any player competes). Furthermore, any player who is out of the system is removed from future consideration.

Validation: We compared simulated age and ISL distributions to actuals for each tour under the current rules. Appendix 3 provides the CDFs for the PGA Tour data indicating slight deviation of the simulated from the actuals.

4 Results

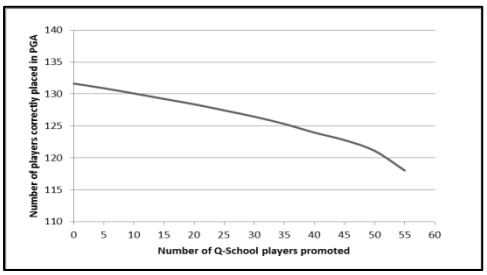


Figure 1: Number of players correctly classified on PGA Tour as a function of the number of Q-School entrants who receive cards assuming 125 players promoted from PGA.

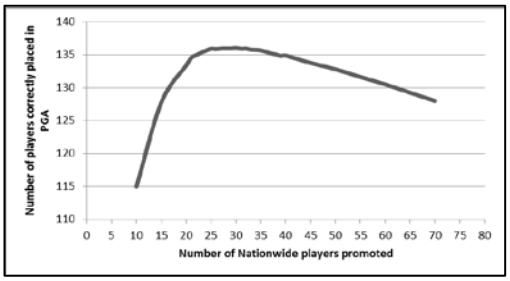


Figure 2: Number of players correctly classified on PGA Tour as a function of the number of NW players promoted assuming no players promoted from Q-School.

Figure 1 shows the impact of the Q-School promotion rules on competitiveness. It shows that as the number of Q-School promotions increases the quality of the PGA Tour decreases. Further it is maximized when no Q-School players are promoted to the PGA TOUR and instead 55 NW Tour players are promoted. We note further that under no configuration of N_{PGA} and N_{NW} is it advantageous to promote players form Q-School. We note also that the average PGA TOUR ISL is greatest in this case too. Figure 2 compares the quality of a system with no Q-School. It shows in this case, that the optimal tour quality occurs when 150 players retain their PGA Tour status and 30 NW players are promoted. Observe that the curve is relatively flat near its maximum.

5 Conclusions

This paper has developed a rigorous model and a simulation to investigate how best to structure the PGA TOUR. It has found in agreement with Connolly and Rendleman [5] that tour quality increases as the number of Q-School finalists who receive card decreases. This makes sense since Q-School is a single 6-round tournament while the Nationwide Tour consists of 27 4-round tournaments. Further we show that the PGA TOUR quality is optimized when 150 PGA players retain their cards and 30 players are promoted to the PGA TOUR from the NW Tour. However, tour quality is relatively insensitive to the number of NW players promoted in the range 25-40. Thus we recommend that the tour adopt these modifications and use the Q-School to populate the Nationwide Tour. But because of the excitement generated by Q-School, we suggest that between 1-5 players from Q-school receive PGA Tour exemptions.

Clearly the question of how best to bring new talent to the PGA TOUR remains open. We believe that the models and approaches developed herein can shed light on that topic and provide precise quantitative recommendations.

Acknowledgement

We acknowledge Ken Lovell from the PGA TOUR for his contributions and discussions regarding this research. A portion of this research was funded through NSERC Grant RGPIN 5527.

References

1. Puterman, M.L. and Q. Wang, "Optimal Dynamic Clustering Through Relegation and Promotion: How to Design a Competitive Sports League", *Journal of Quantitative Analysis in Sports:* Vol. 7: Issue 2, Article 7, 2010.

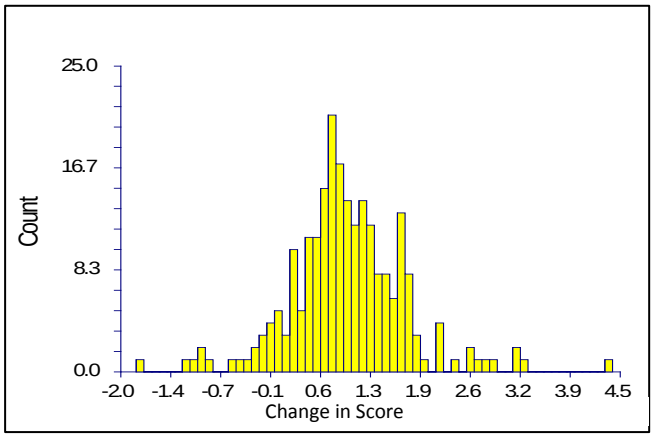
2. Krause, K. "For Golfers, 6-Round Fight with Huge Stakes", *New York Times SPORTS,* December 4, 2011.

3. Berry, S.M., Reese, C. S. and P.D. Larkey, "Bridging Different Eras in Sports", *Journal American Statistical Association*, 94, 47, 661-676, 1999.

4. Connolly, R.A. and R.J. Rendleman, "Skill, Luck and Streaky Play on the PGA Tour", *Journal of the American Statistical Association,* 103,1:74-78, 2008.

5. Connolly, R.A. and R.J. Rendleman, "Going for the Green; A Simulation Study of Qualifying Success Probabilities in Professional and Golf", *Journal of Quantitative Analysis in Sports, Volume 7, Issue 4, Article 7,* 2011.

Appendix 1



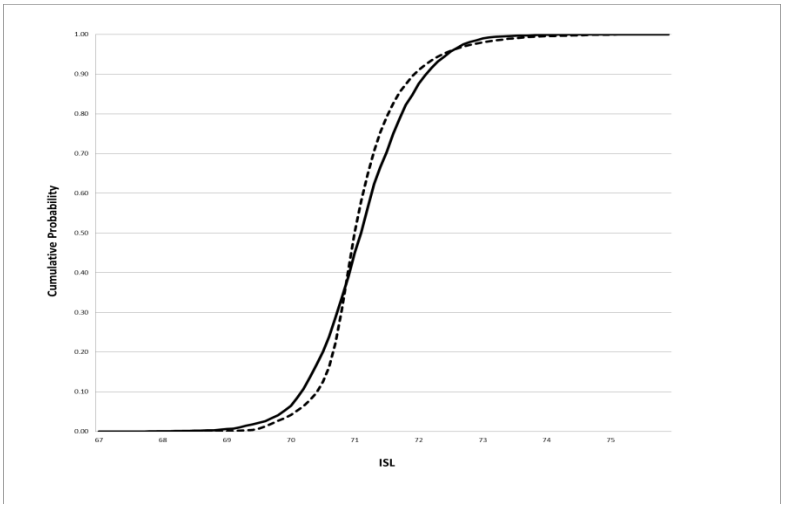
This is a histogram of the change in end-of-season average adjusted scores by players who played in at least 20 rounds of the Nationwide tour one year and played in at least 20 rounds of the PGA tour the following year. Only stroke-play events where all rounds had the same par score were considered, years range from 2004 to 2009 for the Nationwide tour and 2005 to 2010 for the PGA tour. There were 229 data points with mean 0.948, standard deviation 0.768, minimum -1.787 and maximum 4.360.

Appendix 2

PGA Tour Relegation and Promotion Rules

- The top 125 players from the PGA tour retain their PGA tour cards for the next year.
- The top 25 players from the NW tour earn PGA tour cards for the next year.
- PGA tour players ranked 126th and below and Nationwide tour players ranked 26th and below can participate in the Q-School tournament with various levels of exemption.
- Other players qualify for the final stage of Q-School through qualifying stages to make up a final field of 160 players.
- The top 25 players and ties from the Q-School tournament earn PGA TOUR cards for the next year.
- The remaining players from Q-School who reached Stage 3 earn NW tour cards for the next year (with various exemptions).
- PGA tour players ranked 126th–150th who have not qualified through Q-School, retain partial exempt status on the PGA TOUR.

Appendix 3



Simulation validation data for PGA Tour showing cumulative distribution functions of actual (Solid line) and simulated (dashed line) ISLs under current RP rules.