Who are the best snooker players and what sets them apart from the rest?

James Adam Patel Collingwood

A thesis presented for the degree of

Doctor of Philosophy

Department of Management Science

Lancaster University

United Kingdom

December 2021

Who are the best snooker players and what sets them apart from the rest?

James Adam Patel Collingwood

Abstract

Sports analysis is a growing area of research, but snooker has received limited attention, with no systematic analysis of the results or progression of snooker matches previously carried out.

The first part of this research considers four different methods of quantifying the relative ability of professional snooker players. Models based on frames won and world ranking points earned are developed alongside two pairwise comparison models using formulations devised by Bradley-Terry and Elo. The predictive ability of the resultant player ratings is evaluated, with detailed analysis of the results produced to assess the relative strengths and limitations of the models. A recent change in the design of the World Championship qualifying rounds is modelled to determine the likely impact on players at different levels of the professional game.

The second part of this research evaluates different measures of player performance. Post-match video analysis is used to record every shot played during 734 frames within 46 matches in two top-level professional tournaments. The statistics currently produced contain elements of subjectivity, which limits their reliability and availability. Alternative, objective measures are constructed which are shown to reflect the dynamic nature of the game more effectively and could theoretically be generated from the automated scoring system.

Finally, a Monte Carlo simulation model of a snooker frame is developed based on the probability that a ball is potted on different shots, depending on the stage of the frame and the length of the player's current visit. Simulated frames are shown to accurately reflect the progression of observed frames and examples are provided to show how the model could be used to evaluate the impact of alternative choices of shot. Scaling factors are introduced to demonstrate how different levels of player ability could be represented and what effect this has on the progression of a frame.

Acknowledgements

I am very grateful for the support provided by my supervisors, Professor Michael Wright and Dr Roger Brooks. You have both shown a keen interest in my work and have provided valuable feedback and advice as the research has developed. I have also appreciated your patience and encouragement during times where I have encountered difficulties and my progress has slowed.

I would also like to thank everyone in the Department of Management Science and across the Lancaster University Management School for the support they provide to postgraduate research, and for the funding provided from EPSRC.

The decision to take time out of my career to study for a PhD was not any easy one and was only possible due to the support and encouragement of my wonderful wife, Suchi. Moving out of London provided us with a refreshing change of scene, which we have both benefited from. If that wasn't a significant enough change for us, mid-way through my studies we were also blessed with the arrival of our delightful son, Neesh, who provided a welcome distraction during the darkest days of the Covid-19 pandemic. I am also grateful to my colleagues in the Department for Work and Pensions (DWP), who allowed me to take a break from my Civil Service career and who supported my recent return to work.

I have had many enjoyable moments through playing snooker, starting with *Cuestars* competitions when I was younger, through to competing alongside fellow students in the Lancaster University Pool & Snooker Club. I have met some great people and had many fantastic experiences and I hope that this research will make some contribution to the future development of the sport.

Declaration

I declare that the work in this thesis has been produced by myself and has not been submitted elsewhere for the award of any other degree.

James Collingwood

Contents

List of Tables

List of Figures

1 Introduction

I played various sports when I was at school, but snooker was the one which gave me the most enjoyment and is the one which I have stuck with throughout my life. It is the analytical nature of the game which appeals to me. On each shot the player weighs up the merits of different options, balancing the potential gain from executing a shot successfully against the likely cost of not achieving the intended outcome. Having committed to a choice of shot the player then needs to execute it to the best of their ability. Whatever the outcome, perhaps the biggest challenge comes in processing, evaluating and absorbing the consequences of the shot before the next one is played. All sports share this to an extent, but it is particularly noticeable in snooker due to the pace that it is played at and the indirect way in which the two players interact.

At university I found that a large proportion of my fellow players were also studying mathematics and related subjects. The type of analytical thinking nurtured by these disciplines would appear to be well suited to aspiring snooker players. At a time when there has been a surge in interest in applying analysis to inform decisions in a whole range of sports (popularised most notably by Lewis in *Moneyball*, 2004), it is therefore disappointing that snooker is yet to attract the same amount of attention. Snooker is a sport which should lend itself well to such analysis, with the two players taking turns to visit the table, making it easier to assess the outcome of each shot and the interaction between successive shots. Despite this, in Wright's review of the contribution made by

Operational Research over 50 years in sport (Wright, 2009), snooker was only mentioned in passing and little has changed in the subsequent decade.

The main aim of this research is therefore to shine a light on the rich potential which snooker offers to an analyst; and the type of insights that such analysis may provide to interested parties within the sport. It barely touches the surface of what could be done but hopefully serves as a starting point to inspire others to consider ways in which data and analysis can be used more effectively within the sport.

1.1 Thesis Outline

The ultimate aim driving this research is to understand what differentiates the best snooker players from the rest. Can we identify key measures of performance which explain how matches are won and lost? How do the relative strengths of the players affect the progression of a snooker match (made up of a pre-determined number of frames) and how likely is it that a player will win the current frame from a given position?

A core element of this is determining who are the best players in the world and quantifying the relative difference in ability between players. How likely is it that one player will beat another, and how does this change as the match progresses?

The three objectives of this research are:

1. To compare and evaluate different methods of quantifying the relative ability of professional snooker players.
2. To identify and evaluate ways of measuring different aspects of a player's performance during a match.
3. To develop a model of a snooker frame which can be used to estimate the probability of either player winning as the frame progresses and how this is affected by the choice of shot at key moments.

1.1.1 Quantifying the relative ability of professional players

The first part of this thesis considers different methods for rating and ranking professional snooker players. The official World Rankings provide one means of rating a player's performances based on the amount of prize money they have won; a player's win percentage offers another. Unlike these aggregate measures, mathematical models can also account for the strength of opposition faced. A further consideration relates to the time period that is used to measure a player's performance. The official World Rankings assign the same weight to all prize money won within the last two years. Restricting the ratings to a single year may reflect a player's current form more accurately.

Establishing a robust method for quantifying the relative ability of players enables us to produce a reliable expectation about how likely it is that each player will win the next frame / the next match. Aside from providing an insight into how reliable the official World Ranking of players appears to be, this could also form the basis for further analysis exploring different aspects of professional snooker.

1.1.2 Measuring the performance of players within a match

The second part of this thesis considers measures used to assess each player's performance during a match and seeks to establish which of these convey the most meaningful information about how the match was won and lost.

Existing statistics are limited to some basic information about how many points each player has scored and how effectively they have executed different shots. More sophisticated measures are considered which reflect the dynamic nature of the game; each shot played has a consequence for the next shot and this also needs to be captured in any assessment of a player's performance.

1.1.3 Modelling the progression of a frame

The third part of this research uses Monte Carlo simulation to model the progression of a single frame based on the probability of potting a ball on each shot as the frame develops. The model is designed to estimate the probability of each player winning the frame depending on the current score and number of balls remaining, which in turn provides a basis for evaluating the likely outcomes from selecting one shot or another.

The default version of the model is based on two top-level professionals of equal ability, but further adaptations are used to understand how the progression and outcome of a frame varies:

- Depending on the typical level of play of the two players (e.g. whether both are top-ranking or middle-ranking professionals)
- Depending on the relative strengths of the two players; their break-building and tactical prowess.

This creates a framework for linking the first two elements of this research, looking at how differences in the performance of two players affects the progression of a frame and how the top players consequently achieve better results than the rest.

1.1.4 Chapter Summaries

The remainder of this chapter provides background to the research carried out.

- Section 1.2 summarises key elements of the game of snooker, including a description of the game (§1.2.1), the Professional Snooker Tour under which all professional snooker matches are played (§1.2.2) and the World Ranking system used to rank professional snooker players (§1.2.3).
- Section 1.3 summarises the data available, which covers the results of all professional matches played (§1.3.1) and the world ranking of players leading into each tournament (§1.3.2). A limited amount of information is

available from the automated scoring systems used during all professional and some amateur tournaments (§1.3.3), while statistics from some high-profile events are also available (§1.3.4). Post-match video analysis has also been used to record data from over 30,000 shots played during 46 top-level professional matches (§1.3.5).

- Section 1.4 reviews existing literature relating to snooker (§1.4.1) as well as wider research covering the rating and ranking of competitors within sports (§1.4.2), the development of sports performance measures (§1.4.3) and models representing the progression of sporting contests (§1.4.4).

Chapters 2-4 present three papers which have been submitted to academic journals as part of this research.

- Chapter 2 - *Evaluating the effectiveness of different player rating systems in predicting the results of professional snooker matches*. Published in the European Journal of Operational Research (Collingwood et al., 2022).

- Chapter 3 – *The analysis and development of performance measures in snooker*. Originally submitted to the International Journal of Performance Analysis in Sport in June 2020 (Collingwood et al., 2020).

- Chapter 4 – *Simulating the progression of a snooker frame*. Originally submitted to the European Journal of Operational Research in February 2021 (Collingwood et al. 2021).

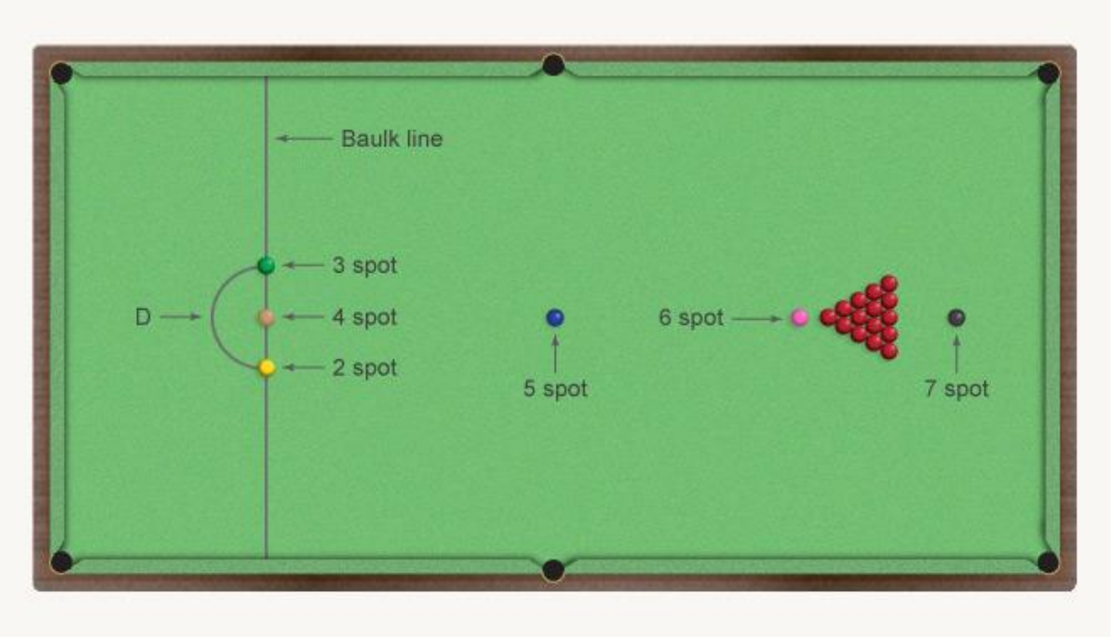
Chapters 5 then summarises the main findings of this research and its contribution in respect of the application of Operational Research analysis to snooker and sport in general. Potential extensions of this research are also discussed.

1.2 Snooker – The Basics

The following section is intended to summarise all aspects of snooker which are referred to in this thesis. Core elements of the game itself are described in Section 1.2.1, such as the points values of the balls used and the order in which they are potted. The structure of the professional game is summarised, with snooker matches organised under a single Professional Tour (§1.2.2). The amount of prize money won in designated ranking events determines a player's official World Ranking, which reflects their status within the game (§1.2.3).

1.2.1 Description of the Game

Snooker is a cue-sport played on a table measuring 12 feet by 6 feet with pockets in each corner and in the middle of the longest sides, as shown in Fig 1.1.



Source: From *Billiards, snooker, pool and darts* by Clayworth, 2013 (https://teara.govt.nz/en/diagram/38477/table-layouts-snooker). Copyright 2013 by Te Ara – The Encyclopedia of New Zealand

Figure 1.1: Top-down view of a snooker table

Players take turns to use a *cue* to hit the white *cue ball* towards *object balls*, which when *potted* in a pocket earn the player a set number of points. The object balls comprise of 15 *reds* (each worth 1 point) and 6 *colours* (yellow = 2, green = 3, brown = 4, blue = 5, pink = 6 and black = 7). The *reds* are initially racked together in a triangle between the pink and black at one end of the table. Each colour has its own *spot* on the table where it begins the frame and is returned to after it is potted. The blue is in the middle of the table, while the yellow, green and brown (known collectively as the *baulk colours*) are located at the opposite end of the table to the *pack* of reds.

For each shot a player can either attempt to *pot* an object ball or play *safe* with the aim of preventing their opponent from potting a ball on their next turn. If a ball is potted then that player continues with their next shot, otherwise play switches to their opponent.

On each new *visit* to the table, a player must first aim to hit (and ideally pot) one of the reds, which can be played in any order and are not replaced once potted. After potting a red, the player nominates one of the colours and has a single shot at it, with the colour replaced on its spot if it is successfully potted. This sequence of *red*, *colour*, *red*, *colour* continues until there are no reds left on the table. At this point the colours are potted (without replacement) in ascending order of value.

Throughout this thesis, any visit where a player legally pots at least one ball is referred to as a *scoring visit*. The total number of points scored during a scoring visit is referred to as a *break*. The “maximum” break under normal conditions is 147 (15 reds, each followed by the black, and then all 6 colours).

A single *frame* of snooker is won by the player who scores the most points - primarily through potting balls, although penalty points are also awarded the opposing player commits a *foul* shot (e.g. if the wrong ball is potted or the cue ball hits another ball before the object ball). A frame ends when the final black is potted, although most frames are conceded once the difference in points between the two players becomes too great for one player to win. A match is won by the first player to win a pre-determined number of frames.

Players need to be conscious of the number of *points remaining* on the table (i.e. the maximum number of points which could be scored from the remaining balls, assuming that the black is potted after each red). If one player trails by more than this amount, then they are said to "require *snookers*". More precisely, they need to try and create a situation from which their opponent will give away penalty points. The best way to achieve this is to *snooker* their opponent, so that they are unable to hit the object ball directly (i.e. there is another ball in a direct line between the cue ball and the object ball). Everton (2014) is recommended for a more detailed description of the game.

1.2.2 The Professional Game

The World Professional Billiards and Snooker Association (WPBSA) is the governing body for the sport. Among other things, the WPBSA is responsible for setting and communicating the official rules of the game (WPBSA, 2019).

The Professional Tour is run by World Snooker, with over 20 tournaments played each season. The season typically starts in June and runs through to the World Championship, which is scheduled to finish on the first Bank Holiday Monday in May.

Each season there are nominally 128 places on the Professional Tour and the Tour is open to male and female players. During the 2020/21 season all professional players were male but for subsequent seasons it has been announced that Tour cards will be awarded to the two highest ranked players on the Women's Tour, currently Reanne Evans and Ng On Yee.

The UK has traditionally been the 'home' of the sport, but it is also growing in popularity in China and the Far East and in some parts of Europe. During the 2020/21 season, the majority of players were from the UK (82), with China providing 23 players. Other countries represented included Australia, Brazil, Germany, Iran, Morocco and Thailand.

There are various ways to qualify for the Professional Tour, typically by winning an international qualifying competition or through a high ranking on an amateur tour. At the end of the season a "Qualifying School" is held, with the strongest performers from a series of competitions earning the final spots on the Tour. Any player qualifying through these routes is awarded a new 2-year Tour card.

Most tournaments held on the Professional Tour are open to all current professionals, with results counting towards the World Rankings. Players are not obliged to participate in every event and amateur players are invited to take the place of any professional who chooses not to enter.

There are also a small number of non-ranking events which the leading players are invited to participate in. The highest profile of these is the Masters, which is contested by the 16 highest-ranked players and is televised by the British Broadcasting Corporation (BBC) in January of each year.

Alternative league-based formats are sometimes used, but during the seasons analysed all ranking events were straight knock-out competitions. The standard format is for all 128 players to start from the same round, with the 16 highest-ranked players seeded to avoid each other until the latter stages (i.e. the top two seeds are in separate halves of the draw; the top four seeds are in separate quarters etc.).

The World Championship is the most notable exception to this design with the 16 highest-ranked players automatically progressing to the final stages, to be joined by 16 qualifiers. The structure of the qualifying rounds for the World Championship is discussed in more detail in Section 2.5.

Towards the end of the season, participation in a series of ranking events is restricted to the players who have earned the most prize money during the season so far. These events are the World Grand Prix (32 players), the Player Championship (16 players) and the Tour Championship (8 players).

A full list of ranking events played during the 2016/17, 2017/18 and 2018/19 seasons is provided in Appendix I. Around half were held in the UK, with each tournament taking place over the course of 1 or 2 weeks. China, India and mainland Europe all hosted tournaments, although it was generally only the final stages (i.e. the “Last 32” onwards) which were played in these countries with the first two “qualifying rounds” played in the UK at an earlier date.

The length of each match varies between tournaments – and often within a tournament. Matches are usually referred to as the “Best of $\boldsymbol{N}$” frames, with the contest ending when one player has an unassailable lead – i.e. is the first player to win $\boldsymbol{n}$ frames, where $\boldsymbol{n} = \frac{(\boldsymbol{N}+1)}{2}$. The early rounds of a tournament are typically the “Best of 7”, “Best of 9” or “Best of 11” with all frames played during a single session. The final match of each tournament is usually either the “Best of 17” or “Best of 19” and is played over two sessions. Matches played during the World Championship are longer; with the early rounds the “Best of 19” and finishing with a “Best of 35” frames final played over 4 sessions.

1.2.3 World Rankings

The official snooker World Rankings are based on the amount of prize money earned by current professionals in ranking events over the previous 2 years. Players earn prize money depending on how far they progress through a tournament; from £0 for first match losers, up to £500,000 for the winner of the 2019 World Championship. Amateur players can receive prize money for any tournament they enter but are not assigned a World Ranking.

The prize fund for each tournament varies depending on the amount received from sponsorship and the sale of television rights. During the 2018/19 season, the World Championship had a total prize fund of £2,231,000 compared with just £100,000 for the Paul Hunter Classic. As a result, the value of each match varies considerably in terms of its influence on a player ranking. Winning the final of the 2019 World Championship earned Judd Trump £300,000 (the difference between the prize for the winner and the prize for the runner-up); while winning the final of the 2018 Paul Hunter Classic earned Kyren Wilson just £10,000 – the same amount as a player winning a match in the first qualifying round of the World Championship that year.

In principle the World Rankings are updated on a rolling basis following the completion of each tournament, although there are around 10 key updates (cut-off points) each season which determine the seedings for subsequent tournaments.

At the end of each season, the 64 highest-ranked players automatically retain their place on the Professional Tour for the following season. New entrants are awarded a place on the Tour for 2 years, so roughly half of those finishing outside the Top 64 will be half-way through this period and will continue for another season, while the other half will lose their place.

It is possible (and fairly common) for a player to lose their place only to immediately earn it back again through the Qualifying School. In these cases, any prize money they earned over the previous two seasons is not counted towards their World Rankings. All “newly qualified” professionals therefore start with 0 ranking points, regardless of their previous experience.

1.3 Data

The data which has been used in this research comes from a variety of sources. Data relating to the results of matches (§1.3.1) and the World Rankings (§1.3.2) are widely available. A limited amount of information is published from the automated scoring systems used in the professional game (§1.3.3) and from statistical analysis produced during televised matches (§1.3.4). Much of the analysis in this thesis has been based on a separate collection of data carried out by the author using post-match video analysis (§1.3.5).

1.3.1 Match Results

The results of all professional snooker matches are available online, with Snooker.org and CueTracker.net the sites used to collate the information used in this study. These sources have been used to capture the results of all frames and matches played in ranking events from the 2005/06 season right through to the current date, with most of the analysis in this thesis based on those played between the 2013/14 and 2018/19 seasons (e.g. Årdalen, 2019a and Florax, 2019a). For each tournament, CueTracker.net also records the number of points scored in every frame by each player (not used in this research) and the sequence in which the frames were won (e.g. Florax, 2019b).

1.3.2 World Rankings

Historical information relating to the World Ranking of players is more patchy and less reliable. CueTracker.net presents the ranking position and prize money earned by players at the start and end of each season (e.g. Florax, 2019c), while Snooker.org lists the seedings of players at each cut-off point during the season (e.g. Årdalen, 2019b). These have been used to determine the prize money ranking of each player at every cut-off point in 2016/17, 2017/18 and 2018/19.

Note that this may differ from a player's ranking at the time a match was actually played – e.g. where there is a gap between the qualifying rounds and final stages of a tournament. This approach was easier to corroborate against the available information and enabled a single ranking to be specified for each player across the whole of the tournament.

1.3.3 Automated Scoring Systems

An automated scoring system is used by the World Snooker Tour at all professional events. After every shot, the official scorer (who might also be the match referee) will enter how many points were scored. This automatically updates the score displayed at the venue, while the feed (for professional events) is also used to update World Snooker's live scoring service (World Snooker Tour, n.d.) as well as being sent to bookmakers for use on their websites. For commercial reasons the outputs from this are not publicly available and this research has had no access to information from this source.

Some amateur tournaments also use this technology and the outputs from two such events are available online.

- The World Snooker Federation (WSF) Open is one of the highest-profile international amateur events played annually. The winner of the event is awarded a place on the professional tour for the following season. Scoresheets from the 2020 WSF Open are published online (World Professional Billiards and Snooker Association, 2020a) and have been analysed within this research.

- The World Seniors Snooker Championship is contested by players over the age of 40 – excluding professionals ranked inside the top 64. The 2020 event was played at the Crucible Theatre and scoresheets are also available online (World Professional Billiards and Snooker Association, 2020b). An example of the data available from this source is provided in Appendix II.

Scoring apps are also available for players to record their own matches.

1.3.4 Performance Measures

Real-time statistics are produced for televised matches. These are sometimes referred to by the commentators during a match and displayed on screen, but they are not systematically collated or disseminated after the match. The only data that have been made available publicly are from matches played in The Masters and the final stages of the World Championship. These were produced for the BBC by Alston Elliot and were released via their twitter account (no longer available). From this source, statistics were extracted from 146 matches played over 7 tournaments (the 2016, 2017, 2018 and 2019 Masters and 2017, 2018 and 2019 World Championships).

The main figures produced are a combination of scoring statistics (Frames Won, Total Points Scored, Balls Potted and Highest Break) and success rates for different types of shot (Pot, Long Pot, Rest Pot and Safety). These are analysed in detail in Chapter 3, with Table 3.1 containing the official statistics produced for the 2018 World Championship final.

Scoring apps have been developed to produce additional statistics. In particular, the creators of mysnookerstats (MSS) have produced a rating for players based on the proportion of pots which are followed by another (Guest, 2010). This is discussed in Chapter 4, with Table 4.1 showing how the MSS Rating is believed to equate to the overall playing standard. The equivalent of their MSS Rating is described as a player's ***scoring power*** within this research as it is effectively a measure of a player's break-building prowess and the number of points they score from each opportunity.

1.3.5 Post-Match Video Analysis

In addition to the publicly available information, I also used post-match video analysis to generate shot-by-shot data from all 31 matches played during the 2018 World Snooker Championship finals; and a further 15 matches played during the 2019 Masters. In total this comprised of 32,601 shots played over 734 frames of snooker. A list of matches analysed is provided in Appendix III.

The information collected included the player taking the shot, the type of shot played (typically a pot or a safety shot) and the outcome of the shot – whether any balls were potted along with any points scored or penalty points conceded.

Additional data were captured for a subset of matches to replicate the success rates produced for the official statistics. These included whether the rest was used to play the shot and an assessment of whether an attempted pot could be considered 'long'. A subjective assessment of whether a safety shot was executed successfully was also made for shots played during the 16 matches contested during the 1st round of the 2018 World Championship.

Any notable features of the shot, perhaps relating to the difficulty of the shot (e.g. if the player was intending for the cue ball to make contact with another ball after striking the object ball - known as a *cannon* - in order to develop or gain position on the next ball), or any unusual outcomes (e.g. 2 reds potted in the same shot) were also recorded within notes relating to each shot.

From these data it was possible to derive information about the status of the frame at the start of each shot: the number of points currently scored by each player, the number of balls (and points) remaining on the table, whether the player was starting a new visit (along with the number of shots since the last pot) or continuing an existing visit (along with the number of shots played and points scored during the visit).

A sample of the information recorded and how it was used to generate match statistics is provided in Appendix IV.

1.4 Literature Review

The following section first looks at existing academic research relating to snooker (§1.4.1). The following sub-sections then consider research which has been carried out in other sports in relation to rating and ranking players (§1.4.2), measuring sports performance (§1.4.3) and modelling sporting contests (§1.4.4).

1.4.1 Snooker

A search for academic literature relating to snooker produces relatively few results. Only a couple of papers analyse the results of professional matches and these only consider a limited number of matches from the World Championship. The mathematics relating to a shot and the implication this has for shot choice is another area with rich potential which has only briefly been explored by a couple of authors. Technical and psychological aspects of the sport are of interest to players and coaches and have been the focus of some academic research.

Expected Results

There are just two academic papers focusing specifically on the results of snooker matches. Clarke et al. (2009) analysed results from the World Snooker Championship between 2004-2007 to assess how well the criteria of fairness, balance and efficiency were met by the design of the tournament. Norman (2015) considered the winners of the World Championship from 1977-2014 in discussing the fairness of the tournament design in light of changes introduced to the Professional Tour from 2010/11.

Neither paper formally considered the relative ability of individual players in their analysis, with Clarke et al. basing their expected results on their finding that across all matches analysed the higher-ranked player won 60.9% of matches played, equating to 53.1% of frames. Neither the round the match was played in nor the difference in the seeding of the two players were found to be significant in determining the likely winner.

It should be noted that in these years, a player's seeding was determined by their world ranking at the start of the season and took no account of matches played since then. While the seedings were broadly representative of one player's ability relative to another, they would not reflect any recent improvements in performance and could not be considered a precise estimate of each player's current ability.

Under the assumption that frames are independent from one another, Clarke et al. used the binomial formulation to estimate the probability of a player winning the match, $\boldsymbol{P}$, based on their chances of winning a single frame, $\boldsymbol{p}$, and the number of frames required to win the frame, $\boldsymbol{n}$:

$$\boldsymbol{P} = \sum_{i=\boldsymbol{n}}^{2\boldsymbol{n}-1} \binom{2\boldsymbol{n}-1}{i} \boldsymbol{p}^{i}(1-\boldsymbol{p})^{2\boldsymbol{n}-1-i} \qquad (1)$$

They compared the actual distribution of 253 match scores against the expected distribution based on the higher-ranked player in each match having a probability of 0.531 of winning a frame. A chi-squared test indicated a significant difference between the two distributions; with a larger number of very one-sided matches than would be expected. They speculated that lower-ranked players may have given up hope of winning the match too quickly if they fell behind.

They were using a very simplistic model of a snooker match though by utilising a single probability to predict the frames won by the higher-ranked player. This ignores the presence of any variation in ability between players, which if it exists (as is likely) would lead to their model under-estimating the occurrence of more extreme results.

Haigh (2009) showed that by setting $\boldsymbol{p} = 0.5 + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ represents a small advantage for one player, an application of Stirling's formula can be used to approximate equation (1) to show that the stronger player's advantage extends to $2\boldsymbol{\epsilon}\sqrt{\boldsymbol{n}/\pi}$ for a match of length $\boldsymbol{n}$. Further analysis shows that this tends to overestimate the player's advantage but is reasonably accurate for $\boldsymbol{\epsilon} \leq 0.05$, with an error of < 0.01 for $\boldsymbol{n} < 13$.

A chart showing the probability of players of different relative abilities winning a match of varying length is shown in Appendix V, with estimates calculated from equation (1).

Shot Choice

The mathematics relating to a shot and potential implications for shot choice have been analysed. Key factors in determining the difficulty of a pot are:

- the distance between the cue ball and object ball,
- the distance between the object ball and the pocket,

- the angle formed by the path of the cue ball and the path of the object ball in order for the object ball to fall into the centre of the pocket (or alternatively, the angle between the path the cue ball would take to hit the object ball in its centre ("full ball") and the path it would need to take to hit the object ball into the centre of the pocket),
- the angle at which the object ball will enter the pocket (or alternatively, the effective width of the pocket perpendicular to the path of the object ball required for it to fall into the centre of the pocket).

Eastaway and Haigh (2011) calculated the margin for error for different shots; looking at how far from the correct angle the object ball could be struck before the pot would be missed. They concluded that the most difficult pots were those where there were relatively large distances between the cue ball, object ball and pocket, with the object ball needing to be hit at a very fine angle towards a narrow pocket. Attempting to pot the black off its spot with the cue ball starting on either the yellow or green spots is about as challenging as it gets.

Percy (1994) demonstrated how Bayesian methods of predictive inference could be applied to snooker in to develop optimal strategies for choosing between different pots; albeit acknowledging the difficulties in obtaining enough data to enable a posterior distribution to be generated. The gain or loss associated with each shot in his model was also limited to the points value of the object ball; further work would be required to factor in additional gains (or losses) from the player (or their opponent) potting a ball on subsequent shots.

Technical and Psychological Analysis

Beyond this, academic research relating to snooker has tended to focus on more technical, technological or psychological aspects of the sport.

A systematic approach to training was devised by Chung et al. (2014), based on assessing and improving a player's level of skill relating to five key elements of the game. Tests were proposed to evaluate a player's control of power, judgement of angles, generation of top / back spin, control of side spin and cue alignment.

A feasibility study on using visualisation to capture and present data obtained through video analysis of a player's practice sessions was produced by Höferlin et al. (2010), initially focussing on capturing any unwanted spin applied to the cue ball as the shot is played.

The visual-perceptual and cognitive abilities of different levels of player have been tested with expert players showing greater capacity for evaluating the situation and a greater depth of forward planning in their shot selection (Abernethy et al., 1994). The stressors and coping strategies of elite players have also been explored (Welsh et al., 2018).

1.4.2 Rating and Ranking Players

This section provides an overview of methods for rating and ranking players and how they are used within sport. Two commonly used methods for exploiting data based on paired comparisons of competitors and are summarised, along with methods for comparing the predictive ability of models.

Overview

Various approaches are taken to produce a numerical rating for the ability or performance of different competitors or teams within a sport. Any such rating system can then be used to produce an ordered ranking of the participants.

Accumulative systems award points according to the result achieved in each event, which incentivises competitors to take part in as many events as they can. Where the points on offer varies this also incentivises organisers and sponsors to provide additional backing for their event to maintain / increase its prestige and attract the top players.

Adjustive systems update each competitor's rating based on how far the actual result deviated from the result that would be expected given the difference in ratings between the competitors.

Aside from official ranking systems, analysts have developed a range of methods to rate players and teams based on the results of matches, some of which are described by Langville and Meyer (2012) who looked at how they have been used to rate and rank teams, individuals and products in a variety of contexts.

Elo

The Elo model is one of the most well-known and most used type of adjustive system, initially devised in the 1960s by Arpad Elo to determine the rating of Chess players relative to one another (Elo, 2008). This utilised a simple method for generating a new rating for a player (R_n) based on their old rating (R_o) and the difference between the actual result (W) and the expected result (W_e):

$$R_n = R_o + K(W - W_e)$$

Where K is the weight assigned to that particular result (Elo, 2008, p25).

Stefani (2011) reported that in some systems such as Women's football, K varies according to the type of match being played. A friendly match would therefore be assigned a lower weight than a World Cup match. In other systems such as Chess, K varies according to the player's level of experience. The rating of a more experienced player is expected to reflect their ability more accurately, so a lower weight is applied when updating their ratings. A higher weight is applied to the results of less experienced players on the basis that their previous rating may be less representative of their true ability, with their level of performance more likely to change over time.

Expected results are based on the difference between the ratings of each player (d) and are commonly expressed as $P(d)$:

$$P(d) = \Phi\left(\frac{d}{\sigma\sqrt{2}}\right)$$

Elo initially assumed that each player's performance was normally distributed, although he also discussed the feasibility of basing the system on the logistic distribution instead.

The standard deviation of the player's performance is usually set at 200, with Elo selecting this value as it ensured that the spread of ratings generated by his system would be similar that of the system it replaced. In this sense it is designed to reflect the spread of players at a similar level of ability, rather than necessarily reflecting how an individual's performance is likely to vary. Under Elo's system, σ is fixed, with the same estimate used for all players. As such, its effect is cancelled out when determining the probability of one player defeating another.

Glickman (1999) has since implemented a variation on Elo's method which allows for different assumptions about a player's ratings deviation. Players competing less frequently would be assumed to have a higher ratings deviation, reflecting that there is less information available on which to base their rating. Subsequent adjustments to their rating would also be slightly larger than for players who had played more frequently.

Glickman (2016) suggested using a ratings deviation of 350 for an unrated player, down to a minimum of 30 for players who compete more regularly. A 95% confidence interval for a player with rating r and ratings deviation RD can be computed as: $(r - 1.96RD,\ r + 1.96RD)$. The confidence interval for an unrated player with a rating of 1500 is therefore (814, 2186), while for a regular competitor with RD = 30 it would be (1441, 1559).

The ratings deviation is an attempt to model the uncertainty in a player's true ability based on the frequency with which they have competed, rather than being directly related to variation in their actual performances. Glickman (2013) proposed a modification to his system which included an additional parameter to represent the volatility of a player's performance.

Bradley-Terry

Agresti (2013) described the logit form of a paired comparison model first proposed by Bradley and Terry (1952). This states that $\log\frac{\pi_{ab}}{\pi_{ba}} = \beta_a - \beta_b$ where π_{ab} is the probability that a (with a rating of β_a) 'prefers' (i.e. beats) b (with a rating of β_b).

For all pairs of competitors, the model is based on the number of previously observed occasions where a was preferred to b, denoted n_{ab} and the number of occasions where b was preferred to a, denoted n_{ba}. Assuming that outcomes of contests are independent of one another then for each pair of players, n_{ab} has a binomial distribution with $n = n_{ab} + n_{ba}$ and $p = \pi_{ab}$. A logit model can be fitted using maximum likelihood estimation to produce ratings for each player.

One additional constraint is required to specify the rating of one individual (e.g. $\beta_1 = 0$). The ratings of all other players are then derived relative to this.

The standard version of the Bradley-Terry model assigns the same weight to all historical contests, regardless of whether they took place at the beginning or the end of the period analysed. McHale and Morton (2011) proposed using an exponential decay function to weight results so that more recent results would have a stronger influence over a player's rating and found that using a half-life of 240 days improved the predictive ability of their model for rating tennis players.

Comparing Models

A standard way to compare the reliability of models used to rank players is to assess their ability to predict the outcomes of matches played. The simplest measure is a model's ***Prediction Accuracy*** (i.e. what proportion of subsequent matches were won by the higher-rated player).

This merely looks at a model's ability to identify the higher-rated player, so ***scoring rules***, which evaluate the probability assigned to each possible outcome, are generally favoured instead (Gneiting & Raftery, 2007). These typically penalise a model more heavily if they confidently predict an outcome which fails to materialise.

Additional measures of the level of ***Calibration*** and ***Discrimination*** exhibited by a model are designed to identify any biases in the models created (Kovalchik, 2016) and are described in more detail in Chapter 2. Kovalchik based her calibration ratio on expected and actual wins for the tennis player with the higher World Ranking, but I have preferred to focus on the higher rated player according to the model being analysed.

The Bradley-Terry models in her analysis were poorly calibrated, achieving a much higher log-loss score than the other models. Inspection of her code reveals that the Bradley-Terry ratings produced actually represent the probability of one player winning a single game against another. These would need to be scaled up to provide an estimate of one player winning a match against the other – as described by McHale and Morton (2011) in their original formulation of the model.

1.4.3 Measurement of Sports Performance

This section considers research relating to measures of sports performance. Statistics relating to sports performance serve different purposes and can be categorised into different types of measure. To be considered an appropriate performance indicator a statistic should possess key measurable qualities. Performance indicators should also reflect the dynamic nature of the contest they are reporting on.

Types of Measures

There are various aspects of a sporting contest which we may wish to analyse to help understand how the match has progressed and how a particular result has arisen. *Scoring measures* present basic information about key events during the match, such as the number of goals scored, or games / sets / frames won. *Quality measures* capture more detailed aspects of a player's or team's performance such as passing accuracy or strike rate. These may capture technical skills, tactical decisions or biomechanical factors (Hughes and Bartlett, 2002).

Qualities of Performance Measures

In developing appropriate measures of performance, O'Donoghue (2015) describes the qualities which a performance indicator should possess.

Validity refers to the relevance and importance of the aspect of performance the indicator is measuring. The scale of measurement should be known and recognisable and it should be possible to interpret each indicator through evaluation against a 'gold standard'. This could be the outcome of the match or a recognised benchmark for success based on data from a comparable peer group or past performances of the teams / individuals studied. A complete set of indicators should ideally cover all aspects of performance.

Objectivity means that the values taken by an indicator should be independent of a given observer's opinion. Human judgement may be required when categorising the data collected but the indicator should be clearly defined to ensure that observers produce consistent records. The *reliability* of each indicator should be tested by checking levels of inter-operator consistency.

Dynamic Interactions

In game sports such as snooker, the outcome of a match is not determined purely by the individual performance of each player. Lames and McGarry (2007) argued that such contests are characterised by a dynamic interaction between the competitors and that the approach used to analyse performance should take this into account. Performance fluctuates both during and between matches, making performance traits (and therefore the indicators themselves) inherently unstable. Individual performance measures should therefore be interpreted with caution.

Similarly, O'Donoghue (2009) showed that outcomes and styles of play in tennis are influenced by both the quality and type of opponent faced, and that different players are influenced by the same opponent types in different ways.

Lames and McGarry (2007) highlighted the merits of depicting game sports as Markov processes, using an example from a Volleyball match to show how a transition matrix summarising the progression of points within a match could be used to help identify key areas where one team dominated. Sensitivity analysis can also be carried out to understand how improvements in one area of performance might affect the overall result.

Examples of Performance Analysis in Sport

In *Moneyball*, Lewis (2004) described how **baseball** teams were starting to use statistical analysis to change the way that they evaluated players, challenging preconceptions widely held within the game. Recruitment strategies and match tactics were tailored accordingly, bringing success to teams who were quick to adopt these and take advantages of inefficiencies in existing practices.

Moneyball introduced sports performance analysis to a wider audience, although the analysis it described had been developed over many years. The Society for American Baseball Research (SABR) was founded in 1971, leading to the emergence of “Sabremetricians” who have become increasingly influential in informing the way that baseball teams operate (Schell, 2011). A wide range of official statistics are now publicly available, which are updated as games progress. Raw data from matches have also been collated and are able to download, allowing a wider pool of analysts to mine the data and enabling the continual evolution of baseball analysis (SABR, n.d.).

In the UK, analytical teams are now embedded into many professional **football** teams, with a vast amount of data now available for them to mine. Opta Sports provide data to a number of clubs (Arastey, 2018) and to aid consistency of data collected across the world they have developed and published a set of ‘Event Definitions’ to formally define the statistics produced (StatsPerform, n.d.). *The Numbers Game* (Anderson and Sally, 2014) and *Football Hackers* (Biermann,

2019) describe how the increased availability of data has changed the way that analysis of football is carried out.

One of the challenges in analysing football matches is that is a relatively low number of goals are scored. Brechot and Flepp (2020) explained how there is a tendency to place too much weight on analysis of goals scored when evaluating the performance of a team given the high degree of randomness associated with this. A measure of *expected goals* is preferred, derived by quantifying the scoring chances created by a team over the course of a game. This captures a larger amount of a team's performance, increasing the robustness of the analysis produced.

This more systematic, "Moneyball" style, approach has also been applied to the evaluation of individual football players, considering the different types of skills which are required in different positions of the pitch (Hughes et al, 2011).

A key innovation in **golf** has been the development of *ShotLink*, an advanced ball-tracking system which provides detailed information about every shot played. Arastey (2020) described how the system has evolved over time to provide insights which benefit players, coaches and viewers. Broadie (2012) used the data from this system to devise an approach for quantifying the success of every shot played in terms of the number of *Strokes Gained* compared with the performance of an average PGA Tour player. This yielded a more robust method for comparing the relative ability of each player in different aspects of the game (putting, long tee shots, approach shots etc.).

1.4.4 Modelling Sporting Contests

Sporting contests can be viewed as complex systems in which competitors interact. A common modelling approach that is used in this research is Monte Carlo simulation and its previous use in sporting settings is reviewed. Alternative approaches are also discussed which use more deterministic methods to identify optimal solutions to specific problems.

Monte Carlo Simulation

Monte Carlo methods are often used to simulate complex systems, allowing for greater depth of study than would be available through deterministic mathematics alone. Eckhardt (1987) described how Stan Ulam and John van Neumann first developed the idea of using computer sampling techniques to explore the behaviour of neutron chain reactions in fission devices, predicting the explosive power of the weapons being designed. The term ‘Monte-Carlo’ was applied to this approach, reflecting similarities with games played at the famous Monte Carlo casino which was frequently visited by Ulam’s uncle.

Monte Carlo analysis is commonly used to simulate the results of sporting contests; typically to model the outcome of a larger tournament with the purpose of evaluating expected results (Koning et al., 2003), different tournament designs (Scarf et al., 2009) or scoring rules (Scarf et al., 2019).

There are fewer examples of Monte Carlo simulation being used to model the constituent parts of a sporting contest. Freeze (1974) described an early model of baseball matches by simulating the outcome of individual pitches based on the ability of the batter and pitcher, along additional factors such as the fielding team’s defensive rating and the batting team’s running ability. Freeze used the model to assess the extent to which variations in a team's batting order affected the expected outcome of a match.

Swartz et al. (2009) used Monte Carlo simulation to model runs scored in a one-day cricket match; using characteristics of the batsman and bowler, as well as factors such as the number of balls bowled and the number of wickets lost, to simulate the outcome of each delivery. They posed a range of questions relating to team selection and batting order which could be addressed using the model.

Broadie and Ko (2009) used simulation to model the outcomes of golf shots based on parameters relating to both the golfer and course layout. They were particularly interested in isolating the effect of changing one skill parameter (e.g. driving distance) while holding others constant.

Alternative Approaches

Deterministic approaches are sometimes preferred, with simplifying assumptions enabling an exact solution to be produced. Markov models can be used to reflect the transitions from one moment in a contest to another based on the likely outcome of the next event. Markov models developed for tennis typically assume that points played on serve are independent and identically distributed throughout the match. Spanias and Knottenbelt (2013) described how this approach can be used to produce a point-level model for predicting the outcomes of tennis matches.

Dynamic programming has been used to determine optimal scoring rates in cricket (Clarke, 1988). The core of the problem is that the higher the rate that a team tries to score at, the more likely they are to lose a wicket. Using assumptions about the relationship between run-rate and probability of dismissal, it is possible to identify the target run-rate which would result in the largest expected total given the number of balls and wickets remaining. Clarke's formulation did not account for the characteristics of individual batsmen or bowlers, although he noted that these could be included.

2 Evaluating the Effectiveness of Different Player Rating Systems in Predicting the Results of Professional Snooker Matches

The first objective of this research is to evaluate different methods for quantifying the relative ability of professional snooker players.

The official World Rankings are based on the amount of prize money won by players over the previous two years. An alternative means of quantifying a player's performance is to calculate the percentage of frames won over the same period. Bradley-Terry and Elo models have also been developed to capture additional factors such as the strength of opponents faced.

In this chapter, the four different approaches to measuring the ability of professional snooker players are evaluated. Section 2.1 summarises relevant existing literature while Section 2.2 sets out the potential methods for rating snooker players which are then tested for their ability to predict future performances. Model predictions are evaluated in Section 2.3, with results compared using the measures described by Kovalchik (2016) to consider both the accuracy and any bias within the predictions.

As this type of analysis has not been applied to snooker results before, relatively simple versions of each model have been selected. This allows for a comparison of the effects of different factors which are likely to influence the outcomes of matches. Section 2.4 looks in more detail at the predictions produced by different models to understand how they differ from one another, highlighting limitations of each approach and potential areas for development.

(1) With limited information relating to ***'new' players***, the resultant predictions for matches involving such players are less certain – particularly for the World Ranking model, which is likely to under-estimate their potential.

(2) The Bradley-Terry model accounts for the ***strength of opposition*** faced when assessing a player's performance, so we would expect this to generate stronger predictions than the Win Percentage model.

(3) ***Current form*** is likely to affect player performance. The Elo model is designed to be the most sensitive to changes in a player's form, while models based on 1 year and 2 years of results and are also compared.

Variations in the qualifying rounds for the World Championship are discussed in Section 2.5, with the 2-year Win Percentage model used to assess the likely impact of changes on players at different levels of the World Rankings. Note that this section is separate from the published version of the paper (Collingwood et al., 2022) but provides an example of how ratings systems could be used to inform decision-making.

Concluding remarks relating to the development of these models are presented in Section 2.6.

2.1 Introduction

The application of analysis in sporting contexts is a popular and growing field of research and Wright (2009) reviewed the specific contribution of Operational Research to sport over a period of 50 years. One sport that has received relatively little attention so far is snooker, despite it offering up plenty of opportunities for analysis.

Whether the focus of analysis has been directly on forecasting the results of matches, or more on the study of tactics or strategy, or on operational matters such as scheduling or tournament design, a key starting point is to determine a baseline expectation for what will happen during a sporting contest. In a contest between two players we are therefore interested in measuring the relative ability of the participants and the resultant expectation that one player will defeat the other.

2.1.1 Snooker

A description of the game is provided in Section 1.2.1.

There are relatively few academic papers focusing specifically on the results of snooker matches. Clarke et al. (2009) analysed results from the World Snooker Championship between 2004-2007 to assess how well the criteria of fairness, balance and efficiency were met by the design of the tournament. Norman (2015) also considered the winners of the World Championship from 1977-2014 in discussing the fairness of the tournament design in light of changes introduced to the Professional Tour from 2010/11.

Neither paper formally considered the relative ability of individual players in their analysis, with Clarke et al. basing their expected results on their finding that across all matches analysed the higher-ranked player won 60.9% of matches played, equating to 53.1% of frames. This ignores the variation in ability between players and would therefore under-estimate the occurrence of more extreme results.

A method for modelling the ability of individual players should generate more accurate predictions. This could also provide the basis for more detailed analysis of how matches and tournaments are expected to progress and how changes to the structure of individual tournaments - or the Professional Tour as a whole - might affect different players.

2.1.2 Rating & Ranking Players

Various approaches are taken to produce a numerical rating for the ability or performance of different competitors or teams within a sport. Any such rating system can then be used to produce an ordered ranking of the participants.

Stefani (2011) produced a comprehensive study of the official rating systems used within 159 sports and noted that objective rating systems broadly fell into two categories.

Accumulative systems award points according to the result achieved in each event, which incentivises competitors to take part in as many events as they can. Where the points on offer varies this also incentivises organisers and sponsors to provide additional backing for their event to maintain / increase its prestige and attract the top players. Out of the 99 sports in Stefani’s study with a published rating system, 84 used some form of accumulative system.

Snooker is included in Stefani's list of accumulative systems and its official World Rankings are now based on the amount of prize money won by 128 professional players in designated ranking events over the past two years.

Accumulative systems can be effective predictors of future success (Stefani, 1998) but there are a couple of features of the Snooker World Rankings which are likely to limit their effectiveness in this respect:

- Results from all ranking events are counted towards the rankings, so players who miss some events (whether through choice, injury or other factors) may be ranked lower than their true ability. Similarly, players joining the Professional Tour start with 0 ranking points, so it isn't until after 2 years that their ranking is on the same basis as more experienced players.
- There is no account taken of the margin of victory or the strength of opposition faced, and the prize money earned for winning a match varies considerably across tournaments and between different rounds of the same tournament.

Adjustive systems update each competitor's rating based on how far the actual result deviated from the result that would be expected given the difference in ratings between the competitors. Stefani (2011) noted that most of the sports using an adjustive system required the competitors to interact in order to control an object, which could be a ball (e.g. Rugby Union and Netball) or a playing piece (e.g. Chess and Go). It therefore makes sense for the impact of a result on a participant's rating to account for the strength of opposition they were facing.

The Elo model is one of the most well-known and most used type of adjustive system, which could be used to rate snooker players. It was devised in the 1960s by Arpad Elo who developed the system using data from historical matches to determine the rating of Chess players relative to one another (Elo, 2008). This utilised a simple method for generating a new rating for a player (R_n) based on their old rating (R_o) and the difference between the actual result (W) and the expected result (W_e):

$$R_n = R_o + K(W - W_e)$$

Where K is the weight assigned to that particular match (Elo, 2008, p25).

Expected results are based on the difference between the ratings of each player (d) and are commonly expressed as $P(d)$:

$$P(d) = \Phi\left(\frac{d}{\sigma\sqrt{2}}\right)$$

Elo initially assumed that each player's performance was normally distributed, with a standard deviation of σ. He also discussed the feasibility of basing the system on the logistic distribution instead, and this is the system now used by the United States Chess Federation:

$$P(d) = 1/\left(1 + 10^{-d/2\sigma}\right)$$

The standard deviation of a player's performance is usually set (in common with Elo's original system) at 200 (Stefani, 2011, p13).

Vaziri et al (2018) noted that the sequence of matches played influenced the Elo ratings. In their study - ranking teams following a round-robin competition - this was an undesirable feature of the model. In most applications though the Elo model is used to estimate the current rating of a player or team, in which case reflecting the sequence of matches played is highly desirable.

Aside from official ranking systems, analysts have developed a range of methods to rate players and teams based on the results of matches, some of which are described by Langville and Meyer (2012) who looked at how they have been used to rate and rank teams, individuals and products in a variety of contexts.

O'Brien and Gleeson (2021) used a form of Markov model based on Google's *PageRank* algorithm to produce all-time rankings for snooker players. Although this type of model does account for the quality of opposition faced, it still nevertheless seems to be biased towards players who have contested more matches. John Higgins (with a win record of 70% from 1,301 matches) is therefore ranked ahead of Ronnie O'Sullivan (75% wins from 1,133 matches).

One type of model that could be applied to rating snooker players is known as a Bradley-Terry model, after the authors who first described the logit form of a paired comparison model (Bradley and Terry, 1952).

As described in Agresti (2013), the Bradley-Terry model states that $\log \frac{\pi_{ab}}{\pi_{ba}} = \beta_a - \beta_b$ where π_{ab} is the probability that a (with a rating of β_a) 'prefers' (i.e. beats) b (with a rating of β_b).

Assuming that outcomes of contests are independent of one another then a logit model can be fitted using maximum likelihood estimation to produce ratings for each player.

McHale and Morton (2011) used a Bradley-Terry model to forecast the results of tennis matches based on the probability of winning games within a match. They developed their model to allow for player strengths varying over time, using an exponential decay function to give greater weight to recent matches.

Rather than a logit model, Baker and McHale (2014 and 2017) based their paired comparison models for tennis players on the beta distribution and estimated strength parameters at different times for each player, interpolating between these using barycentric rational interpolants.

An alternative approach is to assess performance against a range of metrics in order to come up with an overall rating for a player or team. Dixon and Coles (1997) produced separate attack and defence parameters for English football teams in order to predict the results of matches. Broadie (2012) produced a ranking of golfers based on the number of strokes gained across three different aspects of their play (Long game, Short game and Putting).

There is a limited amount of information available which covers all professional snooker players. Within matches, statistics relating to Pot Success and Safety Success are produced for broadcasters – but only for televised matches - and even then, they are not systematically retained after the event. It is therefore not yet possible to formally develop a model on this basis for all players.

The proportions of matches and frames players have won are presented on the website Cuetracker.net (Florax, n.d.) and Win Percentages are sometimes referred to during TV coverage of tournaments, although they aren't currently used to rank players or link past performance to expected future performance. Barrow et al (2013) described how the winning percentage of a player can, however, be used to generate a simple rating of players.

Based on this literature review I have selected four different methods which may be suitable for rating the performance of professional snooker players:

(1) The amount of prize money won by players, as reflected in the official World Rankings

(2) The Win Percentages of each player

(3) A Bradley-Terry model

(4) An Elo model

2.1.3 Comparing methods

A simple way of assessing different rating models is to compare their ***Prediction Accuracy*** (i.e. what proportion of subsequent matches were won by the higher-rated player). Stefani (2011) noted that a well-structured predictive system should be able to out-perform a random prediction by around 17%. Snooker matches are either won or lost, so a random prediction would be correct on 50% of occasions and a well-structured model should therefore achieve a success rate of 67% or higher.

This merely looks at a model's ability to identify the higher-rated player. We also need to look in more detail at the probabilities assigned to different outcomes. We would expect relatively few upsets where the model has identified a strong favourite for each match, but where there is less to choose between two players, we would expect the results to be more unpredictable.

Additional measures set out by Kovalchik (2016) are designed to identify any biases in the models created. The ***Log-loss*** score is designed to penalise incorrect predictions made with high probability, while the ***Calibration*** measure compares the number of expected and actual wins for the higher-rated player and a higher ***Discrimination*** factor would indicate that model predictions are more certain in matches where the higher-rated player prevailed. The formulae for these measures, as well as the Prediction Accuracy, are set out in Table 2.1.

Table 2.1: Descriptions of the measures used to assess the accuracy and reliability of model predictions and the formulae used to create them.

Measure	Description	Formula
Prediction Accuracy	Proportion of matches won by the higher-rated player	$\frac{\sum_i y_i}{N}$
Log-loss score	A scoring rule which applies a higher penalty the more incorrect the forecast	$-\frac{\sum_i [y_i \ln(p_i) + (1 - y_i)\ln(1 - p_i)]}{N}$
Calibration	Ratio of expected wins over actual wins for the higher-rated player	$\frac{\sum_i p_i}{\sum_i y_i}$
Discrimination	A measure of the model's ability to discriminate between results by comparing mean predictions where the higher-rated player won / lost the match	$\frac{\sum_i p_i y_i}{\sum_i y_i} - \frac{\sum_i p_i(1 - y_i)}{\sum_i (1 - y_i)}$

N denotes the number of matches played; $y_i = 1$ if the higher-rated player won the match and $y_i = 0$ if the higher-rated player lost the match; and p_i is the modelled probability that the higher-rated player will win.

2.2 Data and Models

Information about the results of professional matches played are readily available from a variety of sources. The websites Snooker.org and CueTracker.net both record results of snooker matches and have been used to collate the information used in this analysis (e.g. Årdalen, 2019a and Florax, 2019a).

Historical information relating to the World Ranking of players is more patchy and less reliable. CueTracker.net presents the ranking position and prize money earned by players at the start and end of each season (Florax, 2019c), while Snooker.org lists the seedings of players at each cut-off point (Årdalen, 2019b). These have been used to identify the prize money ranking of each player at every cut-off point in 2016/17, 2017/18 and 2018/19.

My analysis has used these data to estimate the probability of one player winning a frame against another based on their relative rating according to four different types of model. Where required, model parameters are determined using results from the 2016/17 season. Predictions are then generated for the 2017/18 and 2018/19 seasons and analysed in Section 3.

2.2.1 World Rankings

The first model (**WR**) is based on the relative ability of each player according to the official World Rankings. The probability of one player winning a frame against another is estimated using the difference in the amount of ranking points (i.e. prize money) earned by the two players in the previous 2 years leading up to the event. The log of ranking points was taken as the distribution of prize money is heavily skewed towards the highest-ranked players.

The logit model generated is that the probability of Player i defeating Player j in a single frame is given by:

$$P(F)_{ij} = \frac{exp[(\ln(RP_i) - \ln(RP_j)) \times C]}{1 + exp[(\ln(RP_i) - \ln(RP_j)) \times C]}$$

where RP_i is the number of ranking points earned by Player i and RP_j is the number of ranking points earned by Player j. The model was fitted using results from the 2016/17 season to determine an appropriate value for C of 0.1657.

2.2.2 Win Percentages

The second approach utilises logistic regression models based on the difference in the overall percentage of frames won by each of the two players. Two separate models were created using this approach; **Win%_1**, which uses a player's results over the last year, and **Win%_2**, which is based on a player's results over the last two years. The 2-year model was designed to mirror the way that the World Rankings are updated, while the 1-year model was considered as potentially a truer reflection of current form.

Logit models for the probability of Player i defeating Player j in a single frame were developed:

$$P(F)_{ij} = \frac{exp\ [(WP_i - WP_j) \times C]}{1 + exp\ [(WP_i - WP_j) \times C]}$$

where WP_i is the current Win Percentage of Player i and WP_j is the current Win Percentage of Player j. The model was fitted using results from the 2016/17 season to give values for C of 3.685 for the 1-year model, and 3.764 for the 2-year model.

2.2.3 Bradley-Terry

The Bradley-Terry models take as an input the total number of frames won by every individual against each of the other players they faced in the given period. As with the Win Percentage model, two separate models were created using this approach; the first based on the results over the previous year (**BT_1**) and the second on results from the last two years (**BT_2**).

The *BradleyTerry2* package in R (Turner & Firth, 2012) was used to carry out the maximum likelihood estimation and produce relative ratings for each player at each cut-off point. The probability of Player i winning a frame against Player j is subsequently:

$$P(F)_{ij} = \frac{exp(BT_i)}{exp(BT_i) + exp(BT_j)}$$

where BT_i is the current Bradley-Terry rating for Player i and BT_j is the current Bradley-Terry rating for Player j. Note that no additional model parameters are required so no training data is used.

2.2.4 Elo

The Win Percentage and Bradley-Terry models place an equal weight on all results which inform them, so are essentially designed to rate a player's aggregate performance across the whole period. Comparing results from the 1-year and 2-year models is a first look at the merits of basing a model on more recent form, and other variations could be created which weight recent results more heavily than past results. The Elo model does this automatically.

The Elo model I developed uses results from the 2005/06 season through to the current date, with players given an initial rating of 500. The logistic formulation of the Elo model is used here, with the log parameter set at 500 and a weight of 10 given to the most recent set of results. This means that ratings typically range from around 1,000 for the top player to 100 for the lowest rated players. Results from matches played during the 2016/17 season were used to assess the suitability of the weighting factor, with similar log-loss scores achieved for a range of values.

The probability of Player i winning a frame against Player j is therefore:

$$P(F)_{ij} = \frac{1}{1 + 10^{(Elo_j - Elo_i)/1000}}$$

where Elo_i is the current Elo rating of Player i and Elo_j is the current Elo rating of Player j.

The formula used to generate a new rating for Player i (R_{iN}) based on their old rating (R_{iO}) is:

$$R_{iN} = R_{iO} + 10\left[F_{ij} - P(F)_{ij} \times \left(F_{ij} + F_{ji}\right)\right]$$

where F_{ij} was the actual number of frames won by Player i against Player j and F_{ji} was the actual number of frames won by Player j against Player i.

2.2.5 Modelling unrated players

One key decision to make for each model is how to account for players who have a limited number of past performances to base a rating on. The approaches I took are set out below and the impact on the results is discussed in more detail in Section 2.4.1.

In the ***World Ranking*** model, unrated players are professionals who have yet to win any prize money, and any amateur player (even if they have previously won prize money). I could have modelled them as having won £0 (and let $\ln(£0) = 0$) but this was found to under-estimate their ability too severely. They have instead been modelled as having won £250, which is lower than the smallest amount earned by a professional over the period studied (£500) so maintains their effective position in the rankings.

For the ***Win Percentage*** and ***Bradley-Terry*** models, an individual rating is not calculated for players who have contested fewer than 10 matches over the relevant period. Basing a rating on too few performances could lead to an unrealistic assessment, while waiting too long fails to use valuable information about a player's performance. Setting a criterion of 10 matches was found to strike a reasonable balance between the two. For players who had contested fewer than 10 matches, a nominal rating was produced based on their aggregated results (i.e. $\sum w_i \,/\, \sum n_i$ where w_i is the number of frames won and n_i the number of frames played by player i).

Matches played during the season 2016/17 which involved players who had contested fewer than 10 matches over the relevant period were excluded when deriving the parameters for the World Ranking and Win Percentage models.

In the ***Elo*** model, new players are given an initial rating of 200 (amateurs) or 300 (professionals) – which in previous years were found to broadly reflect the ability of such players. Subsequent results are then used to update this initial rating for each player. Note that within an official rating system an individual Elo rating for a player would typically not be produced until they have played a requisite number of matches - the main objective being to accurately rate a player's past performance rather than to predict future results.

2.2.6 Updating and applying the models

The World Rankings are updated on a rolling basis throughout the year, with prize money from the latest tournament replacing any money won by players in the equivalent event from two seasons ago. There are around 10 official cut-off points during the season when the latest rankings are used to determine the qualification and seedings for subsequent events. I have based the World Ranking model on the prize money counting towards the official rankings at the relevant cut-off for each event. Similarly, The Win%, BT and Elo models are all updated according to the same cut-off points to allow for a direct comparison with the World Rankings. Taking any of these models forward I would choose to update the ratings more regularly as each event is completed, but for this analysis I have preferred to follow a consistent approach for each model.

In all cases I modelled the probability of a player winning a single frame and derived the probability that the player will win the match using a sequence of independent Bernoulli trials (as described by Haigh (2009)). This is the most effective way of accounting for the variation in the length of matches played across the season, which range from a single frame in the Snooker Shoot-Out to the "Best of 35 frames" in the World Championship final. Modelling frames rather than matches also greatly expands the pool of results informing the analysis.

The assumption of independence is unlikely to be strictly accurate, however it is not likely to affect the results significantly. For the 2017/18 and 2018/19 seasons there was a very strong correlation (0.98) between the overall proportion of frames won by each player and the overall proportion of matches that they went on to win (Fig. 2.1). This suggests that it is reasonable to model the outcomes of matches as a sequence of independent frames.

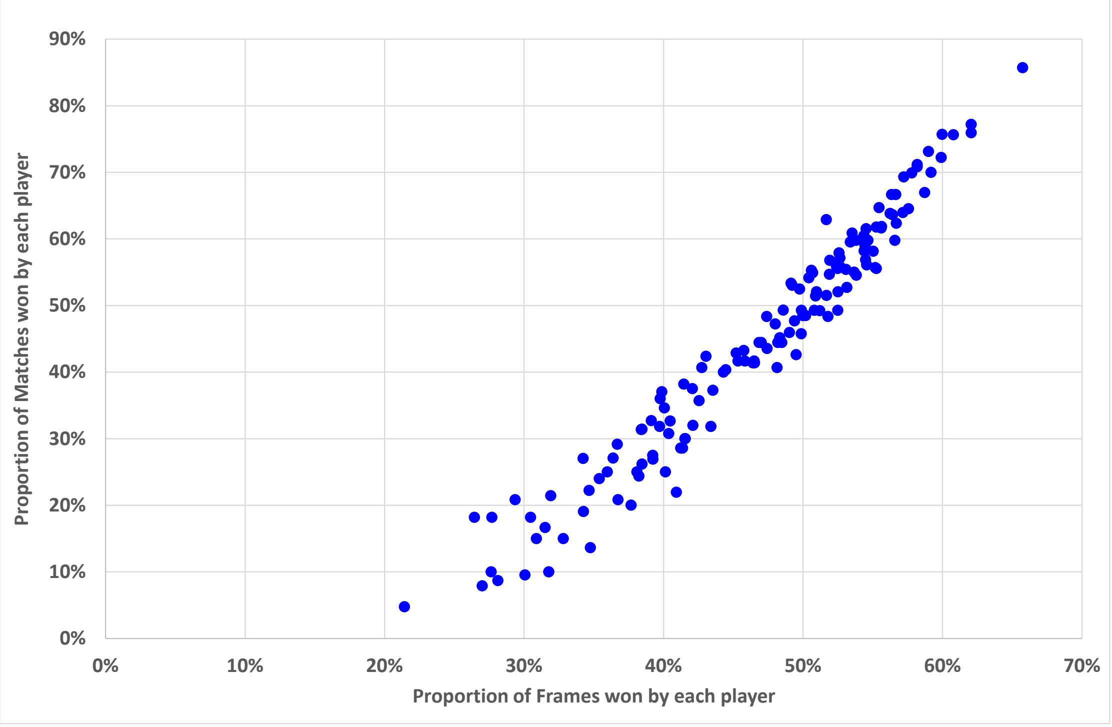


Figure 2.1: Proportion of matches won by professional players in 2017/18 and 2018/19 plotted against the proportion of frames won.

2.3 Results

This section considers the main outputs from the models, looking at 1) how players are ranked under each model at the end of the 2018/19 season, 2) the accuracy of the models in predicting the winners of matches during the 2017/18 and 2018/19 seasons, and 3) any evident bias in the predictions.

2.3.1 Player Rankings

Table 2.2 shows how the top players were ranked by each of the methods at the end of the 2018/19 season.

Table 2.2: The top 16 players in the official rankings at the end of the 2018/19 season and how they were ranked by each of the models.

	World Ranking (WR)	Win Percentage		Bradley-Terry		
		Win%_1	Win%_2	BT_1	BT_2	Elo
Ronnie O'Sullivan	1	1	1	1	1	2
Judd Trump	2	2	2	2	2	3
Mark Williams	3	13	3	10	3	16
Neil Robertson	4	3	4	3	5	1
John Higgins	5	22	5	11	4	6
Mark Selby	6	8	8	7	7	13
Mark Allen	7	9	10	5	8	8
Kyren Wilson	8	4	7	8	9	7
Barry Hawkins	9	7	13	9	12	5
Ding Junhui	10	17	11	19	11	17
Jack Lisowski	11	6	12	6	10	14
David Gilbert	12	19	23	13	21	9
Stuart Bingham	13	5	6	4	6	4
Shaun Murphy	14	41	21	37	18	21
Luca Brecel	15	11	20	20	22	20
Stephen Maguire	16	27	17	18	15	25

- There is a fair amount of consistency between the different models, with Ronnie O'Sullivan, Judd Trump and Neil Robertson ranked within the top 5 of each list.
- Mark Williams won 3 ranking events in 2017/18, including the 2018 World Championship, ensuring that he retains a high placing in the 2-year ranking lists. His performances in 2018/19 were not as strong though and so he is ranked lower on the 1-year ranking lists and Elo model.
- Similarly, John Higgins is ranked lower on the 1-year lists, although by reaching the 2019 World Championship final he retained a high placing in the official world rankings. This latter performance is reflected positively in the Elo ratings, which is the most sensitive to recent form.
- Conversely, Jack Lisowski and Barry Hawkins feature higher in the 1-year ranking lists following stronger performances during the 2018/19 season.
- Stuart Bingham was suspended during the middle of the 2017/18 season so missed out on ranking points from a number of events and slipped down the official rankings as a result. Missing events does not affect a player's rating in any of the other models, so he is ranked much higher by these.

Table 2.3 presents the probability that the #1 rated player under each model at the end of the 2018/19 season will win a frame against players of different ranks. There is a fair amount of similarity between the different methods, although the Bradley-Terry and Elo models both favour the top-rated player more heavily in matches against lower-ranked players than either the World Ranking or Win Percentage models.

Table 2.3: The probability of the #1 rated player under each model at the end of the 2018/19 season defeating opponents of different ranks in a single frame.

Opponent	WR	Win%_1	Win%_2	BT_1	BT_2	Elo
# 2	50%	51%	53%	53%	55%	53%
# 16	55%	57%	58%	64%	64%	63%
# 64	61%	63%	65%	72%	72%	72%
# 100	66%	71%	72%	79%	79%	79%

Two ranking lists can be compared by computing the mean absolute difference in ranking position for each player: $\sum_n \frac{|R_{n1} - R_{n2}|}{N}$, where R_n is the rank of player n under the two models being compared. The more similar the lists, the lower the value computed. Table 2.4 provides an indication of the level of similarity in the ranking of players under each model by comparing player ranks at the end of the 2018/19 season.

Table 2.4: The mean absolute difference in the ranking of players at the end of the 2018/19 season between each pair of models.

Model	WR	Win%_1	Win%_2	BT_1	BT_2
Win%_1	13				
Win%_2	10	7			
BT_1	14	5	9		
BT_2	11	8	4	8	
Elo	12	6	8	6	7

The largest differences are between the World Rankings and each of the other methods. This is likely to be because the design of the official World Rankings differs from the other lists in that the number of ranking points awarded varies from match to match depending on the tournament, while the other methods weight each frame equally. The 2-year versions of the Win Percentage and Bradley-Terry models are a slightly closer match given that the World Rankings are also based on 2 years of performances.

The smallest differences are between the 2-year versions of the Win Percentage and Bradley-Terry models, followed by the 1-year versions. These sets of models are closely related, with the Bradley-Terry models additionally allowing for the strength of opposition. The Elo model is more closely related to the 1-year models in the way that it reflects current form.

2.3.2 Prediction Accuracy

This analysis considers a total of 4,493 matches which took place in ranking events played during the 2017/18 and 2018/19 seasons. [NB – This excludes any qualifying rounds not involving professional players.] To allow a full comparison between methods 72 matches are not included where one or more of the methods couldn't separate the two players, typically because the match took place early in the season when both players were unrated. This left 4,421 matches where I was able to compare the performance of each of the models.

Table 2.5 shows the proportion of matches won by the higher-ranked player under each model for matches played during the 2017/18 and 2018/19 seasons, along with the Log-loss scores achieved by each model.

Table 2.5: The prediction accuracy and log-loss scores for the predictions made by each of the models for 4,421 matches played during the 2017/18 and 2018/19 seasons.

	Matches predicted correctly		Log-loss Score	
Model	**#**	**Proportion**	**Matches**	**Frames**
Win%_2	3.041	68.8%	0.592	0.669
BT_2	3,040	68.8%	0.599	0.670
Elo	3,039	68.7%	0.593	0.669
WR	3,025	68.4%	0.630	0.678
BT_1	3,009	68.1%	0.610	0.672
Win%_1	2,996	67.8%	0.597	0.670

There is very little difference between the ability of the models to predict the winner of each match, although the models based on 2 years of results performed slightly better than those based on a single year. Indeed, in the majority of cases (3,611; 82%) all of the models predicted the same winner, with an accuracy of 72%. The three most accurate models (Win%_2, BT_2 & Elo) predicted the same winner for 91% of matches.

The results are consistent with Stefani's rule of thumb that a well-structured predictive system should achieve a prediction accuracy of 67%. Longer matches are more predictable as they favour the higher-rated player. For the World Championship, where matches are the "Best of 19" or longer, the 2-year Win Percentage model had the highest prediction accuracy (77.2%) and the best Log-loss score (0.495).

Prediction accuracy is a relatively blunt tool for measuring the effectiveness of models as it just requires the model to predict the correct winner; the strength of the prediction is not assessed. The Log-loss score considers the expected probability given to each event and penalises a model more heavily if it confidently predicts a result which fails to take place. Note that the log-loss scores for a single frame are much higher than for the whole match as these outcomes are more uncertain and therefore harder to predict.

The Log-loss score is essentially the average of the log likelihoods across all matches played. The overall scores achieved by two models can therefore be compared by analysing the differences in the log likelihoods obtained. The results of paired t-tests are presented to assess the significance of any differences. The distribution of the differences in log likelihoods is typically not normal, however the relatively large sample sizes provide some mitigation for this and bootstrapping yields very similar results.

- These show that there is no evidence of a difference between the log-loss scores achieved by the 2-year Win Percentage model and the Elo model: $t(4420) = 0.32$, $p = 0.752$.
- There is evidence of a significant difference in the scores achieved by models rating players over two seasons rather than a single season, both for the Win Percentage model: $t(4420) = 2.59$, $p = 0.010$ and the Bradley-Terry model: $t(4420) = 3.40$, $p = 0.001$.
- There is evidence of a significant difference in the scores achieved by the Win Percentage models compared with the Bradley-Terry models, both for the 1-year models: $t(4420) = 4.53$, $p < 0.001$ and 2-year models $t(4420) = 3.39$, $p = 0.001$.

2.3.3 Measures of modelling bias

Table 2.6 shows that the higher the probability of one player beating another, the more likely the models were to be correct. However, in matches where a strong favourite had been identified there were more upsets than anticipated. As suggested by the high Log-loss score, the predictions of the World Ranking model were particularly suspect in this case – 625 players were expected to win their match with a probability of over 90%, but only 79% were successful.

Table 2.6: The prediction accuracy of each model, with matches grouped according to the probability of winning calculated for the higher-rated player.

Probability of winning match	WR	Win%_1	Win%_2	BT_1	BT_2	Elo
0.90 or higher	79%	89%	89%	86%	86%	90%
0.80 – 0.90	78%	80%	81%	77%	79%	80%
0.70 – 0.80	72%	75%	74%	71%	72%	70%
0.60 – 0.70	67%	66%	67%	61%	62%	63%
0.50 – 0.60	58%	52%	55%	53%	54%	55%
All	68.4%	67.8%	68.8%	68.1%	68.8%	68.7%

The Bradley-Terry models identified far more matches where the higher-rated player was given more than a 90% chance of winning (643 for the 1-year model; 602 for the 2-year model) than the Win Percentage models (305 for the 1-year model and 324 for the 2-year model).

The calibration of a model is the ratio of expected wins to actual wins for the higher-rated player, with values above 1 indicating that the model over-predicted the results of the higher-rated player. This was the case for all of the models developed, albeit not significant at the 95% confidence level for the

World Ranking or Win Percentage models. The 2-year Win Percentage model achieved the score closest to 1 (3,062 / 3,041 = 1.01).

The discrimination of a model compares the mean of predictions made where the higher-rated player won against those where the higher-rated player lost (higher scores are better). The 1-year Bradley-Terry model showed the strongest discrimination (75.2% – 68.0% = 7.27%). It was closely followed by the 2-year Bradley-Terry and Elo models, with the World Ranking model showing the least amount of discrimination in its predictions.

Table 2.7: Calibration scores (with 95% confidence intervals) and discrimination of each model, with models ordered by calibration score for matches predicted.

	Calibration Scores with 95% Confidence Intervals		
Model	**Matches**	**Frames**	**Discrimination**
Win%_2	1.01 (0.97, 1.04)	0.99 (0.98, 1.01)	6.61%
Win%_1	1.02 (0.99, 1.06)	1.00 (0.99, 1.02)	6.73%
WR	1.03 (0.99, 1.06)	1.01 (1.00, 1.02)	5.19%
Elo	1.04* (1.01, 1.08)	1.02* (1.00, 1.03)	7.11%
BT_2	1.05* (1.02, 1.09)	1.03* (1.01, 1.04)	7.26%
BT_1	1.07* (1.03, 1.11)	1.04* (1.02, 1.05)	7.27%

* Significantly different from 1.00 at the 95% confidence level

Overall, the 2-year Win Percentage model produced predictions that were of comparable quality to the Elo and 2-year Bradley-Terry models. The discrimination of these latter models was stronger, but they tended to over-estimate the results of the higher-rated players, so in cases where the higher-rated player lost the log-loss was larger.

2.4 Further Analysis

To further understand these results I identified three areas where there are differences between the models: (1) the modelling of new players, (2) whether they take into account the strength of opposition, and (3) whether they take into account recent form of the players. Understanding the strengths and limitations of each approach should provide insights into why there are differences in the quality of predictions produced and how they might be improved.

To do this I looked at how well the models are calibrated for particular subsets of matches, along with the log-loss scores. Note that calibration scores produced in this section relate to the groups of players / matches in question, rather than reflecting predictions for the higher-rated player in each match as in previous calculations. Calibration scores are also produced for the number of frames won as this provides a larger pool of results to analyse.

2.4.1 Modelling 'new' players

It is a challenge for any model to predict the outcomes of matches where one or more participant is relatively unknown, and these models do so with varying degrees of success. The Calibration scores for competitors who have contested fewer than 20 matches in the previous 2 years (Table 2.8) indicate that there are some biases inherent in each of the models, suggesting room for improvement. [Matches contested between players who had both played fewer than 20 matches in the last 2 years were excluded from the analysis.]

Table 2.8: Log-loss scores and Calibration scores for each model relating to predictions made for players who had contested fewer than 20 matches over the previous 2 years – where their opponent had played more than 20 matches (814 matches analysed). Models ordered by ascending log-loss score.

	Log-loss score	Calibration Scores, with 95% Confidence Intervals	
Model	**Matches**	**Frames**	**Matches**
Elo	0.514	0.98 (0.93, 1.03)	0.89 (0.78, 1.03)
Win%_1	0.533	1.01 (0.96, 1.05)	0.94 (0.81, 1.08)
Win%_2	0.534	0.97 (0.93, 1.02)	0.86* (0.74, 0.99)
BT_1	0.554	0.92* (0.88, 0.97)	0.79* (0.68, 0.91)
BT_2	0.560	0.89* (0.85, 0.94)	0.73* (0.63, 0.84)
WR	0.573	0.88* (0.84, 0.92)	0.67* (0.58, 0.77)

* Significantly different from 1.00 at the 95% confidence level

The models tended to under-estimate the performance of players who have contested fewer than 20 matches over the previous 2 years, particularly in relation to the number of matches they win. As anticipated, the World Ranking model is especially weak in this respect, significantly under-estimating the number of frames and matches that these players will win.

The Elo model achieved the lowest Log-loss score with some evidence of a significant difference from both the 1-year Win Percentage model: $t(813) = 2.33$, $p = 0.020$ and 2-year Win Percentage model: $t(813) = 2.18$, $p = 0.029$, while its Calibration scores were not significantly different from 1.00 at the 95% level. The Elo model benefits from using a larger pool of historical results, so matches played more than 2 years ago can still influence a player's rating. The approach I have taken also means that I do not need to wait until a player has contested a minimum number of matches before updating their rating based on their actual performances.

An additional approach used in most Elo rating systems is to apply a larger weight when updating the rating of less experienced players until their performance level stabilises (Stefani, 2011). A smaller weight is then used for more experienced players as their underlying level of performance is not expected to change significantly. There may therefore be scope for improving the Elo model so that it converges faster towards a player's 'true' rating; either by increasing the weight used, and / or by switching their rating after (e.g.) 10 matches to one which is purely based on their performances to date.

The Win Percentage models achieved lower Log-loss scores than the Bradley-Terry models with evidence of a significant difference for both the 1-year models: $t(813) = 3.01$, $p = 0.003$ and 2-year models: $t(813) = 4.24$, $p < 0.001$. They all under-estimated the number of matches won by players who have contested fewer than 20 matches in the last two years.

Further analysis indicates that the calibration of these models may improve if players were given individual ratings after 5 matches rather than 10. This has the effect of reducing the pool of 'unrated' players, thus separating out weaker amateur players (who rarely compete on the professional tour) from new professionals and stronger amateurs who have the opportunity build up their experience more rapidly. Alternatively (or in addition to this), until a player has met the criterion for matches played it should also be possible to produce a more informed provisional rating, for example based on a player's record in amateur competitions.

2.4.2 Strength of Opposition

One anticipated limitation of the Win Percentage model is that the ratings for each player don't take into account the strength of opposition faced. Winning 50% of frames against the top ranked player is clearly a more impressive result than winning 50% of frames against the lowest ranked player; but both are treated equally by the model. The Bradley-Terry model takes as inputs the results against each individual so does effectively account for the strength of each player faced.

To assess how much of an effect this may have on the ratings I looked at the weighted average win percentage of the opponents faced by each professional player in matches which feed into the ratings at every cut-off point. To help identify potential biases I then calculated the average strength of opponent faced by players at different levels of the world rankings (Fig. 2.2). The y-axis is rescaled to show the difference from 50%, which represents 'average' opposition.

The players at the top of the rankings tend to have played against stronger-than-average opposition as they progress further in tournaments and face other in-form players. Some players - notably the #1 ranked player at the end of the 2018/19 season, Ronnie O'Sullivan - will also skip some of the lower profile events and concentrate on events featuring other top ranked players.

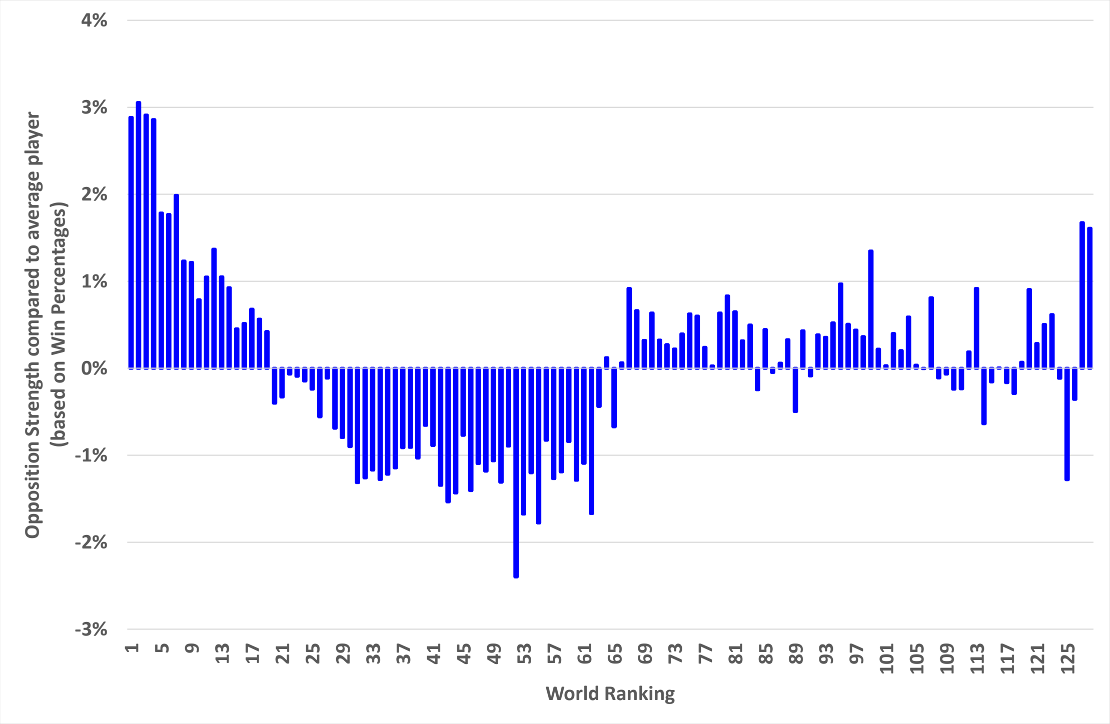


Figure 2.2: Relative strength of opposition faced over the past 2 years (difference from a 50% Win Percentage) by each player at different levels of the World Rankings (averaged over 20 cut-off points of the World Rankings during the 2017/18 and 2018/19 seasons).

Players a bit lower down the rankings tended to face weaker-than-average opposition, consistent with winning matches at the start of tournaments against lower-ranked players but losing the majority of matches they play against higher-ranked opponents.

Players in the lower half of the rankings will generally encounter strong opposition in the early stages of tournaments, so the strength of opponent they faced tended to be slightly stronger-than-average.

To investigate whether this affected the accuracy of the predictions made I considered the matches contested by players at different levels of the World Rankings. By not accounting for the strength of opposition faced we would expect the 2-year Win Percentage model to under-estimate the performance of

Top 16 players against those further down the rankings. We would also expect the model to over-estimate the performance of players ranked 17-64 when facing lower-ranked professionals. In theory the 2-year Bradley-Terry model should tend to produce more accurate predictions in these cases.

Table 2.9 compares the Log-loss scores and Calibration scores (expected / actual wins) associated with the predictions made by the 2-year Win Percentage and Bradley-Terry models for this subset of matches.

As anticipated, the Win-Percentage model did under-estimate the performances of the Top 16 when they faced players ranked 17-64, with the number of frames won significantly higher than predicted. The Bradley-Terry model did not exhibit any such bias and its log-loss score was slightly lower for this group (although not significantly so: $t(650) = 0.88$, $p = 0.380$). In 90% of these cases the Top 16 player had faced stronger opposition than their opponent had over the last 2 years and taking this into account appears to improve the relative ratings of these players.

The findings are not so clear with respect to matches involving players ranked 65 and above. Against Top 16 players the Win Percentage model was well-calibrated, but the Bradley-Terry model over-estimated the performances of the higher-ranked player and its log-loss score was higher as a result, with some evidence of a significant difference: $t(550) = 2.25$, $p = 0.025$.

Table 2.9: Log-loss scores and Calibration scores for each model relating to predictions made for matches contested by players from different levels of the World Rankings).

	Higher-ranked wins	Log-loss Score		Calibration Scores, with 95% Confidence Intervals			
		Matches		Matches		Frames	
		Win%_2	BT_2	Win%_2	BT_2	Win%_2	BT_2
Top 16 v 17-64 (651 matches)	69%	0.604	0.598	0.92 (0.84, 1.01)	1.02 (0.93, 1.12)	0.94* (0.90, 0.97)	0.99 (0.96, 1.03)
Top 16 v 65+ (551 matches)	80%	0.482	0.500	1.02 (0.92, 1.11)	1.08 (0.98, 1.18)	1.00 (0.96, 1.04)	1.05* (1.01, 1.10)
17-64 v 65+ (1,562 matches)	70%	0.583	0.592	1.02 (0.96, 1.08)	1.03 (0.97, 1.09)	1.01 (0.98, 1.03)	1.02 (0.99, 1.04)

* Significantly different from 1.00 at the 95% confidence level

Both models tended to over-estimate the performance of the higher-ranked player in matches between players ranked 17-64 and those ranked 65+. In most of these cases the player ranked 17-64 had faced weaker opposition over the last 2 years and the Bradley-Terry model gave them a slightly lower chance of winning as a result. However, in around a quarter of cases the player ranked 17-64 had faced stronger opposition and so the Bradley-Terry model gave them a higher chance of winning. Overall, there was evidence of a significant difference in the log-loss scores: $t(1561) = 2.96$, $p = 0.003$.

Note that these findings are unchanged if the analysis is restricted to matches where both players had contested at least 20 matches over the last 2 years.

2.4.3 Current Form

The Win Percentage and Bradley-Terry models weight results equally, regardless of whether they took place at the start or the end of the period being modelled. Similarly, the World Rankings essentially weight the last 2 years equally. The overall results indicate that the Win Percentage and Bradley-Terry models were improved by using two years of performances to rate a player, rather than one. At the same time, we might expect to see some instances where a player's form had improved (or dipped) to such an extent that the 2-year models do not accurately reflect their current rating relative to other players.

Fig. 2.3 shows that there are fairly notable differences between the 1-year and 2-year Win Percentages of some players at the end of 2018/19. 21.8% of players had a 1-year Win Percentage that was ± 2.5% different from their 2-

year Win Percentage, indicating that their performance over the last season had been around ± 5% different from the previous season.

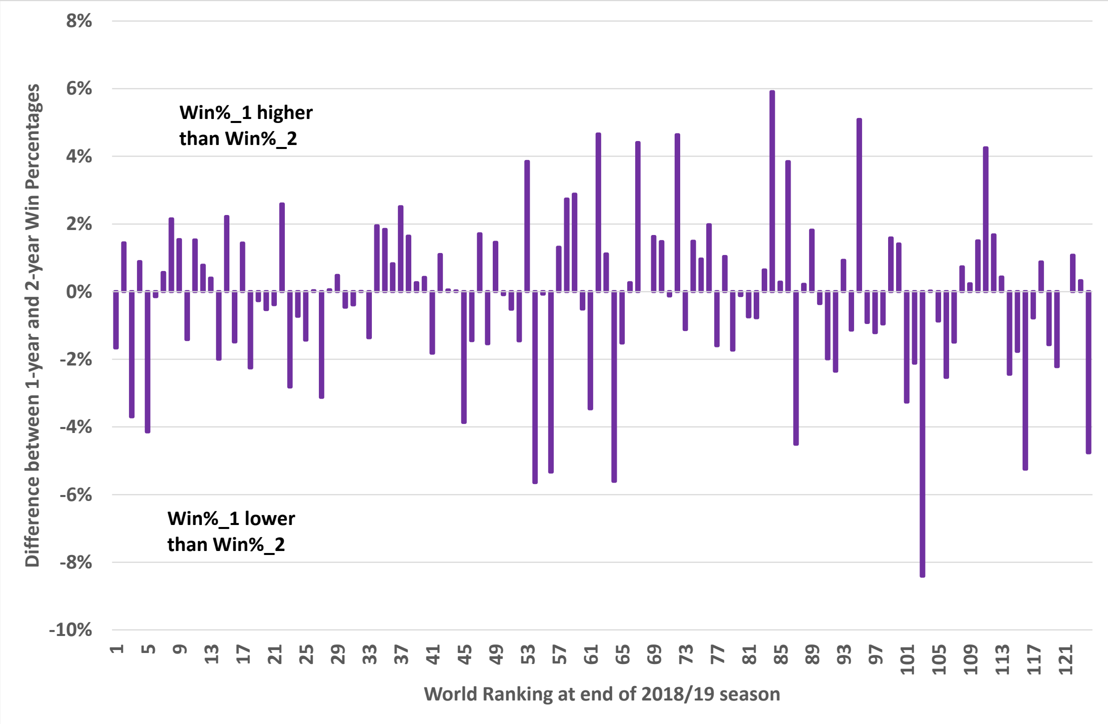


Figure 2.3: Difference between the latest 1-year and 2-year Win Percentage for professional players (Win%_1 – Win%_2), ordered by their official world ranking at the end of the 2018/19 season.

The mean absolute difference is very similar for the 64 highest-ranked players (± 1.7%) as for lower ranked players (± 1.8%), as is the median absolute difference (± 1.4% and ± 1.5%).

To assess the impact of this I looked at specific matches where the 1-year Win Percentage for one player was notably different (± 3%) from their 2-year Win Percentage. In these cases, we might expect the 2-year models to produce less accurate predictions. Matches where both players were "in-form" or "off-form" were excluded from the analysis.

Table 2.10: Log-loss and Calibration scores for predicted results involving players with a difference of ±3% in their 1-year and 2-year Win Percentages ("Off Form" = 2-year Win% is higher, "In Form" = 1-year Win% is higher)

	In-Form Players – 398 Matches			Off-Form Players – 695 Matches		
	Log-loss	Calibration		Log-loss	Calibration	
Model	**Matches**	**Matches**	**Frames**	**Matches**	**Matches**	**Frames**
Win%_2	0.619	0.93 (0.80, 1.07)	0.96 (0.91, 1.01)	0.593	1.05 (0.94, 1.17)	1.02 (0.98, 1.06)
Win%_1	0.620	1.09 (0.94, 1.25)	1.04 (0.99, 1.10)	0.607	0.86* (0.77, 0.96)	0.93* (0.89, 0.97)
Elo	0.636	1.02 (0.89, 1.18)	1.01 (0.96, 1.06)	0.605	0.91 (0.81, 1.02)	0.95* (0.91, 0.99)
BT_2	0.636	0.92 (0.80, 1.07)	0.96 (0.91, 1.01)	0.598	1.02 (0.91, 1.14)	1.01 (0.97, 1.05)
BT_1	0.647	1.10 (0.96, 1.27)	1.05 (1.00, 1.11)	0.634	0.81* (0.72, 0.90)	0.90* (0.86, 0.93)
WR	0.668	0.90 (0.78, 1.04)	0.95 (0.90, 1.00)	0.637	1.08 (0.96, 1.21)	1.04 (0.99, 1.08)

* Significantly different from 1.00 at the 95% confidence level

Table 2.10 compares the Log-loss scores and Calibration scores (expected / actual wins) associated with the predictions made by each of the models for these groups of players.

Although there was a slight tendency for the 2-year Win Percentage and Bradley-Terry models to under-estimate players who have stronger recent form, this was not significant at the 95% level and the models appear to be reasonably well-calibrated. Conversely, the 1-year Win Percentage and Bradley-Terry models tended to over-estimate the performances of players with stronger recent form – although this was again not significant at the 95% level. The Win Percentage models achieved the lowest log-loss scores, with no evidence of a difference between the scores achieved by the 1-year and 2-year models: $t(397) = 0.08$, $p = 0.939$.

The predictions made by the 1-year models did, however, significantly under-estimate the performances of players who had shown weaker form over the last year. There was evidence of a significant difference between the log-loss scores achieved by the Bradley-Terry models: $t(694) = 2.91$, $p = 0.004$, although there was no evidence of a significant difference between the log-loss scores achieved by the Win Percentage models: $t(694) = 1.71$, $p = 0.088$. This provides some support for the overall finding that incorporating an additional year of data generally appears to have improved the models.

It is also interesting to compare the performance of the Elo model in these cases – where the log-loss scores are higher than, but not significantly different from the 2-year Win Percentage model. In-Form players: $t(397) = 1.46$, $p = 0.144$; Off-Form players: $t(694) = 1.32$, $p = 0.187$. The calibration results indicate that although the Elo model appears to be less biased in cases where a player's recent form has been good, it did also under-estimate the performance of players who had shown weaker form over the last year. This again suggests that a recent improvement in form has a stronger influence over future performance than a recent dip in results.

2.5 Modelling the World Championship Qualifying Rounds

Analysis of tournament structures has previously been carried out by McGarry and Schutz (1997) who evaluated the efficacy of different designs of knock-out and round-robin competitions. Scarf et al. (2009) looked specifically at the UEFA Champions League, considering different metrics for assessing the suitability of each design.

With a reliable method for rating individual snooker players we can develop expectations for how likely it is that one player will defeat another. We can also generalise the approach to give an expectation of a player at one position in the World Rankings defeating a player at another position. This is useful if we wish to model the progression of a tournament but do not know precisely which players will be drawn against one another. In this way we can assess how any changes to the design of tournaments are likely to affect different players.

A change introduced during the 2020 season was a modification to the structure of the qualifying rounds for the World Championship. This section presents analysis estimating the impact on players at different levels of the professional game; both in terms of their chances of qualifying for the final stages and in the amount of prize money they could expect to earn from the event.

2.5.1 Structure of Qualifying Rounds

Most professional snooker tournaments are “flat” knock-out events for 128 players, with each player starting in the same round. The World Championship is the notable exception to this, with the 16 highest seeded players automatically guaranteed places in the final stages. They are joined by 16 players who have progressed through qualifying rounds played over the previous couple of weeks.

The structure of these qualifying rounds (as shown in Table 2.11) changed in 2020 so that the highest-ranked players entering qualifying (those seeded 17-32) only need to win two matches to qualify for the Crucible as opposed to three matches in 2019. The intention was to reward stronger performers over the last two seasons.

The previous structure had been in place for five years. Prior to this, the qualifying process was designed so that the top seeds would just have to win a single match to qualify.

Table 2.11: The structure of the World Championship qualifying rounds for 2020, and how this compares with the formats previously used.

	2020	**2015 - 2019**	**Previous Format[1]**
Pre-qualifying	n/a	n/a	Members of WPBSA[2] compete against lowest-ranked professionals
	128 entrants in qualifying		**80 entrants in qualifying**
Qualifying Round 1 (QR1)	Players seeded 81-112 v Players seeded 113-144	Players seeded 17-80 v Players seeded 81-144	Players seeded 65-80 v Players seeded 81-96
Qualifying Round 2 (QR2)	QR1 winners v Players seeded 49-80	QR1 winners v QR1 winners	QR1 winners v Players seeded 49-64
Qualifying Round 3 (QR3)	QR2 winners v Players seeded 17-48	QR2 winners v QR2 winners	QR2 winners v Players seeded 33-48
Qualifying Round 4	QR3 winners v QR3 winners	n/a	QR3 winners v Players seeded 17-32
16 qualifiers join the 16 highest seeded players in the final stages			

[1] Note that a variation of this format was used in 2014, with players seeded 17-32 again just having to play one match.

[2] World Professional Billiards and Snooker Association

The draws for the later rounds are dictated by the seedings, so the 17th seed (the highest-ranked player in the qualifying rounds) is scheduled to play the 80th ranked player in the penultimate round, and the 48th ranked player in the final round (assuming each of these players progress through earlier rounds), while the 18th seed will be scheduled to play the 79th and 47th ranked players, etc.

2.5.2 Historical Results

In analysing the changes to the tournament format, the primarily interest is in the impact on qualification for the final stages.

Based on results from 2004-2007, Clarke et al. (2009) noted that the higher-rated player won **60.9%** of matches completed (excluding the 1st qualifying round). Indeed, over the whole period studied in which players seeded 17-32 had to win a single match (against lower-ranked opposition), they were successful in qualifying for the final stages on **63%** of occasions (Figure 2.4) – meaning that an average of 10 top seeds qualified for the final stages.

The change to the format in 2015 meant that there was a greater difference in the ranking of players contesting each match. During these 5 years, the higher-rated player won 86% of matches in the 1st qualifying round and 68% of matches thereafter. This implies that the chances of a player seeded 17-32 qualifying for the final stages had reduced to 86% x 68% x 68% = **40% -** with 6.4 out of the 16 top seeds expected to qualify. In fact, Figure 2.4 shows that the actual proportion qualifying was **48%**, although this figure is potentially skewed by results from 2016 when 12 of the qualifiers were seeded 17-32.

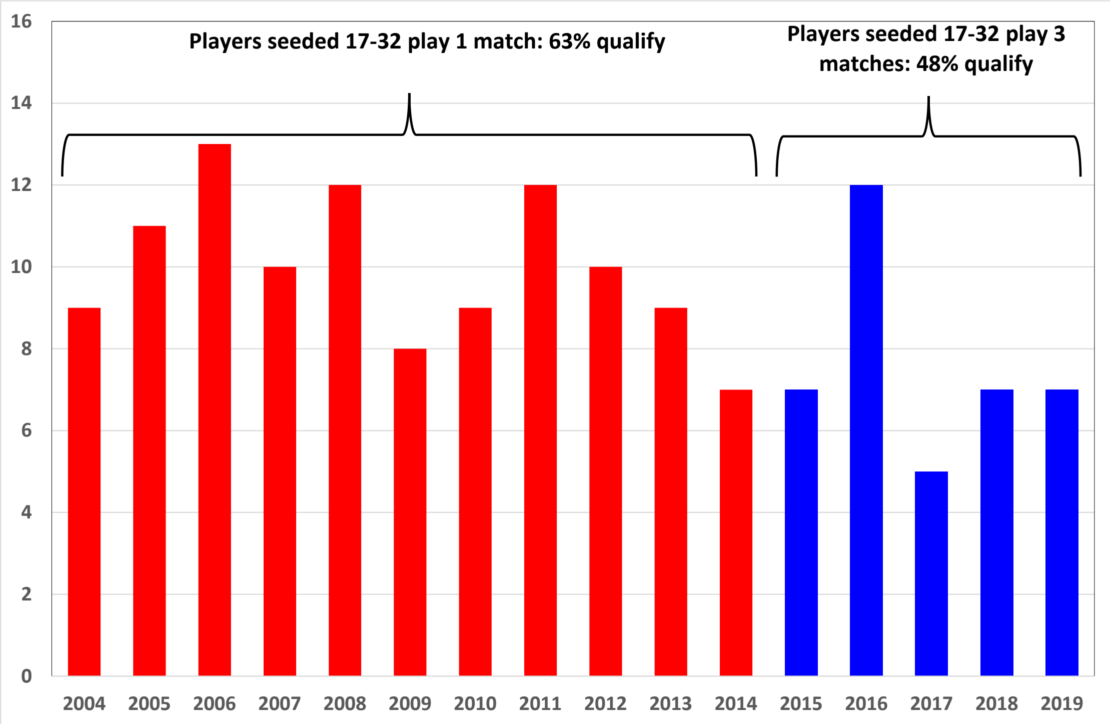


Figure 2.4: Number of players seeded 17-32 qualifying for the final stages of the World Championship each year (out of 16).

We can speculate that the changes to the format in 2020 will see the chances of players seeded 17-32 rise to 68% x 68% = **47%** (as the opposition they face in their 2 matches is likely to be comparable to previous years).

2.5.3 Expected Results

These aggregated estimates provide an indication of what will happen, but they also obscure variations in the expected outcomes - especially due to the way the draw is seeded. For example, the highest-ranked player the 17th seed could face is the 48th seed, while the 32nd seed will potentially have to play the 33rd seed in the final qualifying round.

To fully understand the impact of the changes we can learn more from modelling the results in full. As well as analysing the effect on qualification for the finals, we can also consider the effect on the expected amount of prize money earned by different players.

The amounts awarded to players losing at each stage are shown in Table 2.12. Note that players losing their first match receive £0 prize money regardless of which round they enter the tournament; players seeded 49-80 will receive nothing if they lose in the 2nd qualifying round, and players seeded 17-48 will receive nothing if they lose in the 3rd qualifying round.

For the purposes of this analysis, I have not considered any additional amounts that may be earned by qualifiers if they should progress beyond the 1st round of the World Championship finals.

Table 2.12: Prize Money awarded to players knocked-out in each of the qualifying rounds, and the minimum amount earned by players qualifying for the final stages

	2019	2020
Qualifying Round 1	£0	£0
Qualifying Round 2	£10,000	£5,000 / £0
Qualifying Round 3	£15,000	£10,000 / £0
Qualifying Round 4	n/a	£15,000
1st Round	£20,000	£20,000

I replicated the structure of the qualifying rounds of the World Championship under both the 2019 and 2020 formats and used the 2-year Win Percentage model described in Section 2.2.2 to estimate the probability of each player progressing from one round to the next. This model was preferred as it produced results for the 2019 format which were broadly comparable with actual results from the last 5 years.

To achieve this, I used data from the last three seasons to produce a regression model for the 2-year win percentage of players based on their ranking prior to the World Championship each year. The linear model (with an R-squared of 61%) yielded a typical win percentage of 56.4% for the 17th seed, down to a win percentage of 46.8% for the 80th seed, as shown in Fig. 2.5.

The error bars represent the highest and lowest win percentages across the three years, while the trendline represents the resultant regression model.

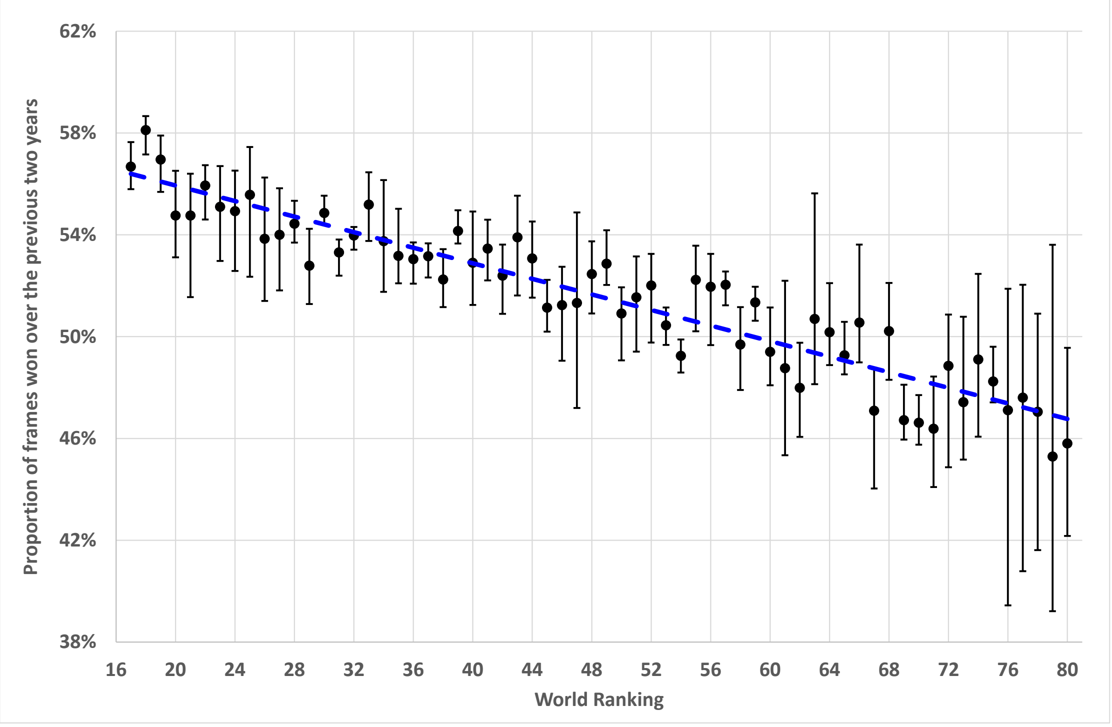


Figure 2.5: Mean 2-year win percentages of the players ranked 17-80 prior to the World Championship in 2017, 2018 and 2019.

Players seeded 81-144 are randomly drawn so I have just used the average win percentage of 39.9% for this group. In any given year the results of each player will not form such a predictable distribution, but this reflects the relative ability of a player at each level of the World Rankings over time.

Table 2.13 compares the actual and modelled outcomes for players at different levels of the rankings under the format used between 2015-2019; and how these are expected to differ under the new format introduced for 2020.

The modelled results indicate that a player's chances of qualifying for the final stages are not much different under the 2020 format. Players seeded 17-48 see their chances improve slightly as a result of playing one match fewer – although most of these players won their 1st round matches under the 2019 format anyway.

Table 2.13: A comparison of expected results from the World Championship qualifying rounds, looking at the probability of different players 1) losing their first match, and 2) qualifying for the final stages

	Probability of losing 1st match			Probability of Qualifying		
	2015 – 2019 format		2020 format	2015 – 2019 format		2020 format
Seeds	**Actual**	**Modelled**	**Modelled**	**Actual**	**Modelled**	**Modelled**
17-32	0.08	0.10	0.23	0.48	0.44	0.47
33-48	0.15	0.14	0.37	0.20	0.28	0.30
49-64	0.23	0.19	0.19	0.18	0.17	0.14
65-80	0.13	0.25	0.25	0.11	0.10	0.08
81-144	0.85	0.83	0.50	0.04	0.02	0.01

Conversely, players seeded 81-144 see their (already slim) chances reduce as they are required to play an extra match. Players seeded 49-80 play the same number of matches in both formats but are now guaranteed to face a higher-ranked player in their second match so see their chances of progressing fall slightly.

The flip side of the changes is that exactly half of the players seeded 81-144 will now win their 1st round match – and receive prize money – while under the previous format 85% went home with nothing. Their expected average earnings increase from £1,932 to £3,133 as a result. Higher-ranked players actually see a slight reduction in their expected earnings (Figure 2.6), with players seeded 17-48 in particular having a greater chance of losing their 1st match and going home with nothing (Figure 2.7).

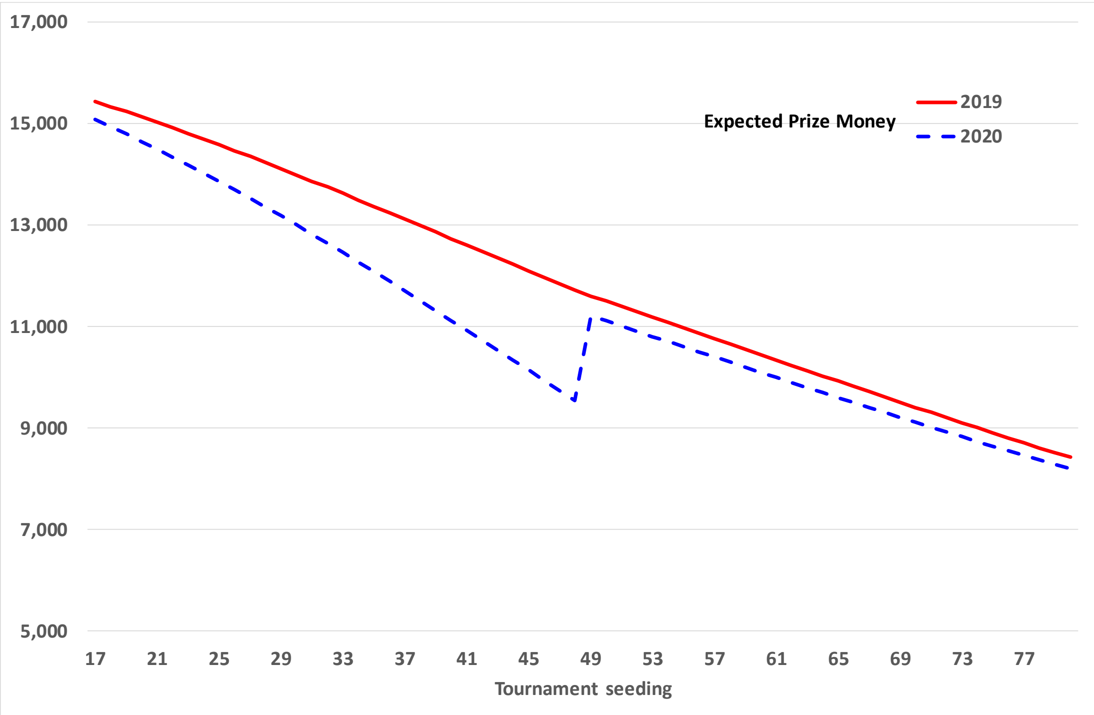


Figure 2.6: Expected amount of prize money won by players seeded 17-80 under the 2019 and 2020 qualifying formats for the World Championship

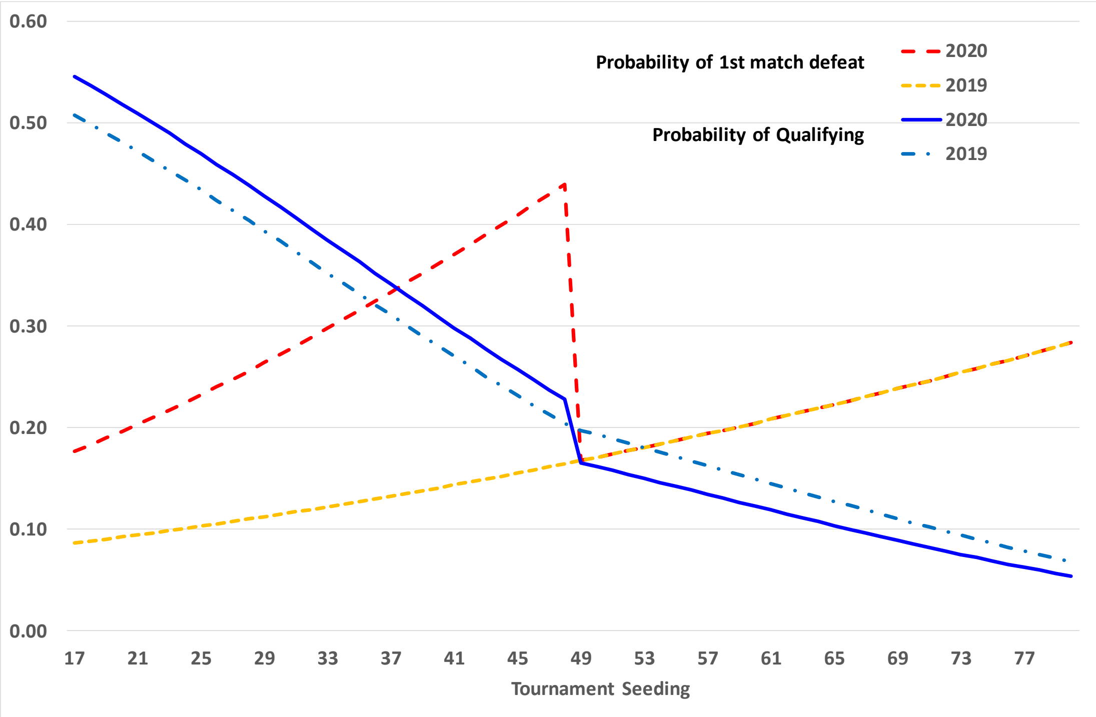


Figure 2.7: Estimated probability of 1) a 1st match defeat, and 2) qualifying for the final stages for players seeded 17-32 under the 2019 and 2020 qualifying formats for the World Championship.

2.5.4 Summary of Findings

The new structure introduces a clear difference in expectations for players just inside the Top 48 and those just outside. Those just inside benefit from having to win fewer matches to qualify, but with a greater risk that they will lose their 1st match and not win any prize money. Although they would have to win one more match to qualify than players ranked above them, the 49th seed is less likely to lose their first match than any other player in the qualifying competition and over time is consequently expected to win more prize money than players seeded 40-48.

One further factor not considered in the modelling is that lower-ranked players may benefit from having already won a match before taking on a higher-ranked opponent who hasn't played yet. It will be interesting to see if there is any evidence of such players winning more frames than expected, particularly in the early stages of these matches.

On balance, the changes introduced for 2020 seem very reasonable, with a limited impact on a player's chances of qualifying. Players seeded 17-48 receive a small advantage from having to play one fewer match – which they typically won anyway. Players seeded 81-144 now compete against each other for the right to take on higher-ranked opponents, and a higher proportion will win a match and receive prize money from the tournament as a result.

The total amount of prize money won is more unpredictable as it depends on how many of the higher-ranked players lose their first match, but World Snooker can expect to pay out around £28,800 extra as a result of the changes – from a total pot for the tournament of around £2.5million.

2.5.5 Postscript

The 2020 World Championship took place in July & August 2020, delayed by over three months due to the Coronavirus pandemic. There was also a slight modification to the qualifying rounds to reduce the amount of time players spent at the venue. Originally intended to be the "Best of 19 frames" played over two sessions, matches played during the first three qualifying rounds were reduced to the "Best of 11 frames", each played in a single session. The same format was followed for the 2021 World Championship qualifiers in April 2021.

Table 2.14 shows the number of players qualifying for the final stages from different levels of the world rankings, based on their seeding for the event. Modelled results are also shown based on a) the intended format of "Best-of-19" frame matches in each round and b) the actual format used of "Best-of-11" frame matches for the first three qualifying rounds. Note that these are the long-term expected results using the methodology described in Section 2.5.3, rather than being based on the actual rating of individual players in each year.

Table 2.14: The actual number of qualifiers for the final stages of the World Championship in 2020 and 2021 coming from different levels of the World Rankings compared with modelled outcomes

	Actual Qualifiers			Modelled Qualifiers	
Seeds	**2020**	**2021**	**Mean**	**Intended Format**	**Actual Format**
17-32	4	7	5.5	7.6	7.2
33-48	6	3	4.5	4.9	4.9
49-64	2	5	3.5	2.2	2.2
65-80	4	1	2.5	1.2	1.4
81-144	0	0	0	0.1	0.5

The outcomes in 2021 were reasonably close to the expected outcomes given the shortened format of matches played. There were more surprises in 2020 when 8 of the players seeded 17-32 lost their first qualifying first match (compared with 4 the following year).

A notable feature from both years is that there were more qualifiers than expected from lower down the rankings (those seeded 49-80). They may have benefited from playing a match before taking on a higher-ranked opponent, particularly given the shortened matches. In 2020, the effects of the suspension of the Professional Tour and limited use of practice facilities during the period of lockdown may also have compounded this. We cannot conclude anything from just two years of results, but it will be interesting to see if this is repeated in subsequent years.

2.6 Concluding Remarks

With a prediction accuracy in the 67-69% range, the models all met the benchmark for a well-structured model and there was relatively little to choose between them. Overall, the Bradley-Terry and Elo models showed greater discrimination than the Win Percentage models but tended to over-estimate the performances of higher-rated players and did not produce more accurate results.

Using frames won as the basis for the model means that there is a large pool of data used in producing the ratings for each player (with most professionals contesting over 200 frames per season). Playing conditions are broadly consistent from one tournament to another and the physical demands on players are relatively low, so performances are likely to be fairly stable. Win Percentages have therefore shown to be a reliable measure of a player's relative ability, albeit with some potential room for improvement in using this as a basis for modelling future performances.

Each of the models underestimated the performances of relatively inexperienced players and would be improved from having a more informed estimate of their ability, perhaps using results from amateur competitions. As anticipated, this was the main limitation of the model based on the World Rankings as it utilises the least amount of information about players who are new to the professional tour.

Models based on two years of results were found to be more reliable than those based on a single year of results. Although there was some indication that a recent improvement in form may be worth accounting for, dips in performance appear to be less significant, with the 1-year models and Elo model (which places more weight on recent results) underestimating the performances of players with win percentages that had fallen over the last year. The concept of form has not been extensively studied, although an analysis of results in golf tournaments showed that a player's performance in their last 6 tournaments had a significant influence over their finishing position in their next one (McHale and Forrest, 2005). Similar analysis applied to snooker could indicate whether the Win Percentage and / or Bradley-Terry models would be improved by weighting recent results more heavily.

One anticipated limitation of the Win Percentage model is that it does not account for differences in the strength of opposition previously faced by players. I found some evidence to suggest that it under-estimated the performance of the highest-ranked players, who tend to progress further in competitions and therefore compete against stronger opponents. This was not so apparent for matches involving lower-ranked players though, suggesting that other factors may be more relevant in establishing the relative ability of these players. The Bradley-Terry model did not produce more accurate predictions so an alternative approach would be to adapt the Win Percentage model instead – such as the Ratings Percentage Index described by Barrow et al, 2013.

Variations on the Elo model may also help to improve predictions for players at different stages of their career. This could be achieved by applying different weights according to the player's level of experience (Stefani, 2011) or allowing for differing amounts of variation in their performance (Glickman, 1999).

The assumption that frames are independent from another yields reliable results although there is likely to be some element of dependence, particularly where the higher-rated player takes a substantial lead in a match. Further analysis could be carried out to explore whether / when expectations should change within a match - if a match goes down to the final frame, do we still believe that one player is more likely to win, or is the outcome now closer to 50 / 50?

As well as predicting the outcomes of individual matches, these models could also be used to predict expected outcomes from tournaments. A single edition could be simulated by using the rating of players at the start of the event.

Alternatively, Section 2.5 demonstrated how expected outcomes in the long run could be modelled by looking at the mean rating of players at each rank. The impact of changes to the World Championship qualifying rounds were shown to affect players in different ways. Higher-ranked players would see a slight increase in their chances of qualifying for the final stages, while lower-ranked players were more likely to win a match and receive prize money under the new design.

A further extension to this would be to model an entire season of professional events to consider the results and level of performance required for players to progress higher up the rankings.

3 The Analysis and Development of Performance Measures in Snooker

The second objective of this research is to identify and evaluate ways of measuring different aspects of a player's performance during a match.

This chapter reviews the statistics that are produced for televised matches, evaluating their objectivity and validity. Alternative measures which capture the dynamic progression of a frame more effectively are also considered. A limited amount of data is currently available but there is potential for generating statistics from the scoring system used for all matches on the professional tour.

Section 3.1 summarises existing literature relating to performance analysis in sport, while Section 3.2 sets out the data used in this research, including published match statistics (§3.2.1), data collected using post-match video analysis (§3.2.2) and data generated from the official scoring system (§3.2.3).

In Section 3.3, the validity and objectivity of the existing statistics are discussed, while in Section 3.4 alternative measures are considered. The relevance of statistics relating to the number of ***scoring visits*** made by each player and the points scored during each scoring visit is demonstrated. Dynamic measures capturing the ***Scoring Potential*** of each player (the proportion of visits in which a ball is potted) and their ***Scoring Power*** (the proportion of pots followed by another) are shown to have a strong relation with match outcomes. Concluding remarks are presented in Section 3.5.

3.1 Introduction

In *Moneyball*, Lewis (2004) describes how Billy Beane took advantage of the wealth of baseball data available to make more educated decisions over who to sign and how to assess and use the strengths of different players within a match. He had been inspired by analysis produced by Bill James which questioned the validity of traditionally used statistics. In recent years, many sports have gone through a similar process whereby the increasing availability and accessibility of data have inspired analysts to develop enhanced methods of evaluating the performance of competing players.

Snooker is a sport which lends itself well to detailed analysis. The two players take turns visiting the table, making it easier to assess the execution and outcome of each shot. Shot choice plays a key part in the game as players weigh up the potential benefits from an attacking shot against the increased risk of handing their opponent an easy opportunity to score points. Despite this, there is very limited data collected and no work has been carried out to evaluate potential indicators of performance.

A description of the game is previously provided in Section 1.2.1, with existing literature relating to snooker discussed in Section 1.4.1.

3.1.1 Performance Analysis in Sport

There are various aspects of a sporting contest which we may wish to analyse to help understand how the match has progressed and how a particular result has arisen. ***Scoring measures*** provide information about key events during the match, such as the number of goals scored, or games / sets / frames won. ***Quality measures*** capture more detailed aspects of a player's or team's performance such as passing accuracy or strike rate. These may capture technical skills, tactical decisions or biomechanical factors (Hughes and Bartlett, 2002).

In developing appropriate measures of performance, O'Donoghue (2015) describes the qualities which a performance indicator should possess.

Validity refers to the relevance and importance of the aspect of performance the indicator is measuring. The scale of measurement should be known and recognisable and it should be possible to interpret each indicator through evaluation against a 'gold standard'. This could be the outcome of the match or a recognised benchmark for success based on data from a comparable peer group or past performances of the teams / individuals studied. A complete set of indicators should ideally cover all aspects of performance.

Objectivity means that the values taken by an indicator should be independent of a given observer's opinion. Human judgement may be required when categorising the data collected but the indicator should be clearly defined to ensure that observers produce consistent records. The ***reliability*** of each indicator should be tested by checking levels of inter-operator consistency.

Suitable indicators can be identified by contrasting the performances of winning and losing teams. Three methods are compared by Fitzpatrick et al (2019a) - the point-biserial correlation between the performance characteristic and the match outcome (win / loss), a paired t-test comparing the performance of the winner and loser, and the proportion of matches in which the winner outperformed the loser on a particular characteristic. This latter method was used to assess the relative importance of different aspects of performance in tennis matches played on clay and grass (Fitzpatrick et al, 2019b).

In game sports such as snooker, the outcome of a match is not determined purely by the individual performance of each player. Lames and McGarry (2007) argue that such contests are characterised by a dynamic interaction between the competitors and that the approach used to analyse performance should take this into account. Care needs to be taken in interpreting individual performance measures as performance fluctuates both during and between matches, making performance traits (and therefore the indicators themselves) inherently unstable. Similarly, O'Donoghue (2009) shows that outcomes and styles of play in tennis are influenced by both the quality and type of opponent faced, and that different players are influenced by the same opponent types in different ways.

Lames and McGarry (2007) highlight the merits of modelling game sports as Markov processes, using an example from a Volleyball match to show how a transition matrix summarising the progression of points within a match could be used to help identify key areas where one team dominated. Sensitivity analysis can also be carried out to understand how improvements in one area of performance might affect the overall result.

As the quantity and type of data increases, analysts have taken more creative steps to evaluate the performance of competitors. The development of ball-tracking data in golf allowed Broadie (2012) to devise an approach for quantifying the success of every shot played - the number of *Strokes Gained* compared with the performance of an average PGA Tour player. This yielded a more robust method for comparing the relative ability of each player in different aspects of the game (putting, long tee shots, approach shots etc.).

In football, matches are now analysed using measures such as the *Expected Goals* created by each team - rather than the actual number of goals scored (Brechot & Flepp, 2020). This increases the amount of action within a match that is evaluated, reducing the influence of randomness in actual goals scored, and producing a more reliable assessment of a team's true performance and the contribution of individual players.

3.2 Data

Three sources of data are used in this study: official data produced by Alston Elliot for the BBC, data that I have collected through post-match video analysis, and data captured on the *frame sheets* generated from a tournament's scoring system. This section describes the data gathered from each of the sources and how I have used them in my research.

The statistics covered in this chapter treat each shot equally. Some shots are played after a frame has effectively been 'won' (i.e. when the difference in the score is greater than the total value of the balls remaining). There may be an argument for counting these separately, but this is beyond the scope of this research.

3.2.1 Official Statistics

Real-time statistics are produced for televised matches. These are sometimes referred to by the commentators during a match and displayed on screen, but they are not systematically collated or disseminated after the match. The only data that have been made available publicly are from matches played in The Masters (an invitational event open to the 16 highest-ranked players) and the final stages of the World Championship (involving the 16 highest-ranked players plus 16 qualifiers). These were released by Alston Elliot via their twitter account (no longer available). From this source I extracted statistics from 146 matches played over 7 tournaments (the 2016, 2017, 2018 and 2019 Masters and 2017, 2018 and 2019 World Championships).

The main figures produced are a combination of scoring statistics (Frames Won, Total Points Scored, Balls Potted and Highest Break) and indicators of quality in the form of success rates for different types of shot. Table 3.1 presents the official statistics produced for the 2018 World Championship Final, won by Mark Williams.

Table 3.1: Official statistics produced by Alston Elliot for the 2018 World Championship Final

Measure	Mark Williams	John Higgins	Notes
Frames Won	**18**	**16**	
Total Points Scored	1,930	1,784	Includes penalty points conceded by their opponent
Balls Potted	556	485	
Highest Break	118	131	The highest break made in a single visit
Pot Success	**91%** (556 / 614)	**92%** (485 / 529)	Successful pots as a proportion of all attempted pots
Long Pot Success	**62%** (29 / 47)	**57%** (13 / 23)	Attempted pots which are over half the length of the table
Rest Pot Success	**94%** (30 / 32)	**83%** (24 / 29)	Pots attempted using the rest, an implement which allows players to take shots beyond their physical reach
Safety Success	**83%** (116 / 140)	**74%** (114 / 155)	The proportion of safety shots deemed to have been executed successfully

3.2.2 Data collected from post-match video analysis

Many professional matches are subsequently made available to view on youtube.com. I used this footage to manually record data from all 31 matches played during the final stages of the 2018 World Championship and the 15 matches played during the 2019 Masters; a total of 32,601 shots played over 734 frames.

For every shot played I recorded the Player, the Type of shot (Pot or Safety shot) and any Points scored (or conceded). Additional data were collected for a subset of matches to enable the production of my own estimates of the full suite of shot success rates. A sample of the data recorded is provided in Appendix IV.

3.2.3 Data generated from the scoring system

An automated scoring system is used at all professional tournaments, as well as some of the leading amateur tournaments. After every shot, the official scorer (who might also be the match referee) will enter how many points were scored. This automatically updates the score displayed at the venue, while the feed (for professional events) is also used to update World Snooker's live scoring service (World Snooker Tour, n.d.), as well as being sent to bookmakers for use on their websites. For commercial reasons the outputs from this are not available for professional events, but the frame sheets from the 2020 World Snooker Federation (WSF) Open – a high profile amateur event whose winner qualifies for the professional tour – are available online (World Professional Billiards and Snooker Association, 2020).

A sample of the information captured is shown in Appendix II.

I extracted data from 56 matches played during the WSF Open, covering all matches featuring the 8 quarter-finalists (excluding the Last 64 tie between Iulian Boiko and Nickolas Neale, where the frame sheets were unavailable). For each shot played, one of the following actions is recorded on the frame sheet, along with additional information about the shot:

- Ball potted + the player who potted it and the points scored as a result
- Foul committed + the penalty points given away and the player receiving them
- Play switching to the opponent + the player now at the table

It became apparent that the first and last shots of a frame were sometimes not recorded – if no points were scored then this has no impact on the frame score. Video recordings were available for some matches to confirm what was happening, so I subsequently added shots where it appeared that these had been missed.

3.3 Existing Statistics

This section analyses separately the existing scoring measures and quality indicators, considering their validity, objectivity and reliability.

3.3.1 Scoring Measures

Scoring measures are intended to provide an overview of the match and to give a sense of how well each player has performed. The current statistics of Total Points Scored, Balls Potted and Highest Break are all clearly defined and measured so we can focus our assessment on their validity.

The number of points scored is clearly critical to the outcome of a frame but ***Total Points Scored*** over a match is a less relevant measure to collect and isn't a meaningful indicator of a player's performance. To enable a comparison across matches, it would be slightly more interesting to present the average number of points scored per frame instead.

There is a direct link between ***Balls Potted*** and Total Points Scored so it is unsurprising that the two measures are very highly correlated (with a coefficient of 1.00). There is therefore no additional value in knowing how many balls were potted by each player.

Prize money is often awarded for the highest break made in a tournament so for this reason the ***Highest Break*** made by each player in a match has some relevance. It is not a good indicator of a player's overall performance though as it just reflects a player's most profitable visit to the table. Presenting the number of 100+ and 50+ breaks made by each player would be more useful in this regard.

3.3.2 Quality Measures - Validity

The shot success rates in snooker are primarily used to compare the performance of the two players within a match. To be valid indicators they should therefore be correlated with the outcome of the match, or at the very least there should be a recognised benchmark for top level performance.

Table 3.2 shows the mean success rates for the winners and losers of every match, with a paired t-test showing that there was a significant difference for each of the measures except for the Rest Success rate. From this we can also see that around 75% of all shots were attempted pots and 25% were safety shots. Long Pots (10%) and Rest Pots (5%) represent relatively small proportions of the total pots attempted, equating to 3.2 long pots and 1.7 rest pots per frame.

The ***Pot Success*** rate had the strongest correlation with match outcomes. This was even clearer when comparing success rates with the proportion of frames won – both for individual players and the difference in performance between the winner and loser. A success rate of 90% is often referred to by commentators as being the typical standard required to win a match against a top professional, which is supported by these data. One limitation is that shot complexity is not taken into account; a low success rate may reflect the type of chances a player has had more than the quality of their potting performance.

Table 3.2: Success rates achieved by winners and losers of matches and correlations with match outcomes. Data available from official statistics produced by Alston Elliot for 146 matches played during The Masters and World Championship.

	Pot Success	Long Pot Success	Rest Success	Safety Success
Overall Success Rates (Total Shots attempted)	90% (76,216)	59% (7,325)	88% (3,822)	80% (25,765)
Mean Performance of Match Winners (and standard deviation)	91% (2%)	62% (12%)	87% (10%)	81% (5%)
Mean Performance of Match Losers (and standard deviation)	87% (5%)	56% (13%)	88% (13%)	78% (6%)
T statistic[1]	8.65	4.50	-0.10	5.77
p-value[1]	<0.001	<0.001	0.919	<0.001
Matches won by the player with the higher success rate (with 95% confidence interval)	77% ± 6.8%	64% ± 7.8%	43% ± 8.0%	65% ± 7.9%
Point-biserial correlation with match outcome (Win / Loss)	0.46	0.25	-0.01	0.28
Correlation with proportion of frames won in the match by each player	0.66	0.33	0.01	0.36
Correlation between the difference in success rates and margin of victory	0.75	0.32	0.08	0.39

[1] Paired t-tests of the success rates observed for match winners and losers were carried out across all 146 matches (145 available matches for Rest Success)

There was a weaker correlation between match outcomes and ***Long Pot Success***, although a clear majority of matches were won by the player with the higher success rate. In theory, this measure considers a subset of more difficult pots, which may be key to a player establishing or extending a scoring visit. In practice, this covers a wide range of shots with varying degrees of difficulty and given the relatively small number of long pots attempted each match we cannot really say whether the level recorded for a player is better or worse than we would expect.

There are even fewer shots played with the rest, so it is unsurprising that ***Rest Success*** rates were found to have no influence over the outcome of a match. Collecting statistics on this is still relevant though as it is a very specific element of the game that players need to master. An individual player's success rate will vary from match to match but over a longer period of play the statistics indicate that we can expect a top professional to pot close to 90% of shots with the rest.

Although not currently utilised, the official statistics allow us to analyse the amount that different players use the rest. For some players this averages more than once per frame, for others it is closer to once every two frames – some players prefer to take shots with their 'opposite hand' (e.g. a naturally right-handed player would take the shot left-handed) or extend their cue slightly to allow them to reach further. Rather than simply looking at Rest Pot success it would be more pertinent to collect statistics on each of the different approaches taken by players to execute shots beyond their normal physical reach.

Safety play is a key element of the game and there is some correlation between the ***Safety Success*** rate currently collected and match outcomes, with the data suggesting that top professionals can be expected to achieve a success rate around 80%. Success rates do vary more from match to match, which may suggest that they are affected by the playing style and ability of the opponent – or could reflect the level of subjectivity in determining whether a shot has been executed successfully or not.

3.3.3 Quality Measures – Objectivity and Reliability

The success rates cannot be produced automatically; human judgement is required to determine whether a Pot or a Safety shot has been attempted, whether a Pot can be considered ‘Long’ (or if the rest has been used) and whether a safety shot has been executed successfully or not. Alston Elliot do provide training and guidance to data collectors to help ensure consistency, but comprehensive definitions for these measures are not publicly available.

To understand the scale of human judgement required I compiled my own statistics for the 2018 World Championship and the 2019 Masters. I did not expect to be able to replicate the official statistics precisely but was keen to understand some of the challenges involved and potential limitations in the production and interpretation of the statistics. The percentage difference between my observations and the official statistics provides some indication of the extent to which the subjectivity of observers may affect the statistics produced.

Table 3.3 compares the data I collected with the official statistics from the two tournaments. It is noticeable that I recorded more shots in both tournaments than the official statistics - I have not been able to definitively establish why this is the case. It could be that some types of shot are not included in the success rates – possibly shots re-taken following a foul (as the outcome of the shot has not yet been fully resolved), or fouls incurred while potting a ball (the ‘pot’ itself was successful so these aren’t strictly ‘missed pots’) – so I have classified these separately in my records. Alternatively, it could be that the final shot in some frames is not recorded in the official statistics where it has no impact on the match (as I found in my analysis of the data from the scoring system).

Table 3.3: A comparison of official data recorded by Alston Elliot and data collected through post-match video analysis for 1) the 2018 World Championship (1st round matches) and 2) 2019 Masters

	2018 World Championship 1st round (16 matches)		
	Official data[A]	My data[B]	% Difference[1] (Range)[2]
Total Shots	11,537	11,601	1% (0%, 2%)
Attempted Pots	8,499	8,569	1% (0%, 3%)
- Successful	7,571	7,570	0% (0%, 0%)
- Missed	928	981	6% (-8%, 20%)
- Fouls	n/a[4]	19	
Pot Success[3]	89%	89%	-1% (-2%, 1%)
Attempted Long Pots	837	698	-18% (-38%, 9%)
- Successful	518	380	-31% (-55%, -3%)
- Missed	319	318	0% (-34%, 38%)
Long Pot Success[3]	62%	54%	-7% (-18%, 0%)
Attempted Rest Pots	444	437	-2% (-11%, 15%)
- Successful	385	378	-2% (-11%, 18%)
- Missed	59	59	0% (-40%, 18%)
Rest Success[3]	87%	86%	0% (-3%, 2%)
Attempted Safeties	3,038	3,032	0% (-4%, 3%)
- Successful	2,442	2,408	-1% (-13%, 12%)
- Unsuccessful	596	545	-9% (-33%, 31%)
- Re-taken shots	n/a[4]	79	
Safety Success[3]	80%	82%	1% (-7%, 7%)

	2019 Masters (15 matches)		
	Official data[A]	**My data[B]**	**% Difference[1] (Range)[2]**
Total Shots	6,160	6,197	1% (0%, 2%)
Attempted Pots	4,676	4,645	-1% (-2%, 1%)
- Successful	4,139	4,136	0% (0%, 0%)
- Missed	537	490	-9% (-22%, 3%)
- Fouls	n/a[4]	19	
Pot Success[3]	89%	89%	1% (0%, 2%)
Attempted Long Pots	413	340	-19% (-49%, 5%)
- Successful	212	185	-14% (-67%, 18%)
- Missed	201	155	-26% (-73%, 0%)
Long Pot Success[3]	51%	54%	3% (-10%, 13%)
Attempted Rest Pots	202	197	-3% (-13%, 7%)
- Successful	171	165	-4% (-14%, 0%)
- Missed	31	32	3% (0%, 15%)
Rest Success[3]	85%	84%	-1% (-4%, 0%)
Attempted Safeties	1,484	1,552	4% (0%, 13%)
- Successful	1,132	n/a[4]	
- Unsuccessful	352	n/a[4]	
- Re-taken shots	n/a[4]	53	
Safety Success[3]	76%		

[1] For frequencies the % difference has been calculated as $(B - A)/[(A + B)/2] \times 100$, while for percentages the % difference has been calculated as $(B - A)$. Note that the convention for inter-observer comparisons is to calculate the absolute % difference but in this case the direction of any differences is of interest.

[2] The range in the % difference across all matches analysed is shown in brackets

[3] Success rates for all measures are the number of successful shots as a proportion of all successful and unsuccessful attempts

[4] n/a = not available (official data), or not captured (my data)

For more difficult attempts at pots, players will have the additional aim of keeping the balls safe in case they miss. These are described as *shots-to-nothing* as there is no attempt made to retain good position for the next object ball. In such cases where a ball is not potted, some subjectivity is required in determining whether these should be categorised as safety shots or missed pots. For the 2018 World Championship I tended to record these as missed pots, counting more than the official statistics. I changed my approach for 2019 The Masters so that I was more inclined to classify shots-to-nothing as safety shots – and subsequently counted more safety shots than the official statistics.

Despite this, the outcomes of an attempted pot are clearly identifiable, so I was able to match the overall ***Pot Success*** rates quite closely to the official figures. Similarly, the ***Rest Success*** measure is relatively easy to replicate, with observed differences due to human error. [Note that the percentage difference for rest shots is still quite large for some matches, but this typically only reflects a difference of one or two shots.]

The ***Long Pot Success*** rates proved more difficult to reproduce, with a relatively wide variation in the figures produced. It was particularly noticeable that I counted far fewer long pots than the official statistics. Although some of this may be down to perceptual error, I may have used a narrower definition of a long pot by only looking at the distance between the cue ball and the object ball. A broader interpretation would also consider the distance between the object ball and the pocket. Even within the official statistics there is considerable variation between tournaments - ranging from 51% during the

2019 Masters to 71% at the 2018 Masters – suggesting that different criteria were used and / or different judgements were made by observers.

My overall ***Safety Success*** rate for the 1st round of the 2018 World Championship was relatively close to the official figure, although there was much wider variation for individual matches. Deciding whether a safety shot had been executed successfully or not was the most time-consuming of judgements to make and I chose to not to capture this for subsequent matches.

3.4 Alternative Measures

It should be possible to use technology which exists in other sports to track the location of balls on the table. With this information an expected success rate could be derived for each attempted pot, while a player's positional prowess could also be assessed.

Such technology is not currently available, so the immediate aims of this research were to set out additional scoring measures which help to explain how matches have progressed; and to consider outcome-based quality measures which better reflect the dynamics of a snooker match. This yields a more objective set of measures, opening up the potential for using information captured by the scoring system to vastly increase the quantity of performance data available.

3.4.1 Scoring Visits

As the primary aim of each player is to score enough to points to win the frame it makes sense to highlight the number of ***scoring visits*** each player has made and their associated scoring rate.

The mean number of scoring visits made per frame (by both players) across the two professional tournaments analysed was 3.7, with the full distribution shown in Figure 3.1. Unsurprisingly the average number of scoring visits per frame was much higher (6.3) for the amateur event – the 2020 WSF Open - with players needing more opportunities to score enough points to win a frame.

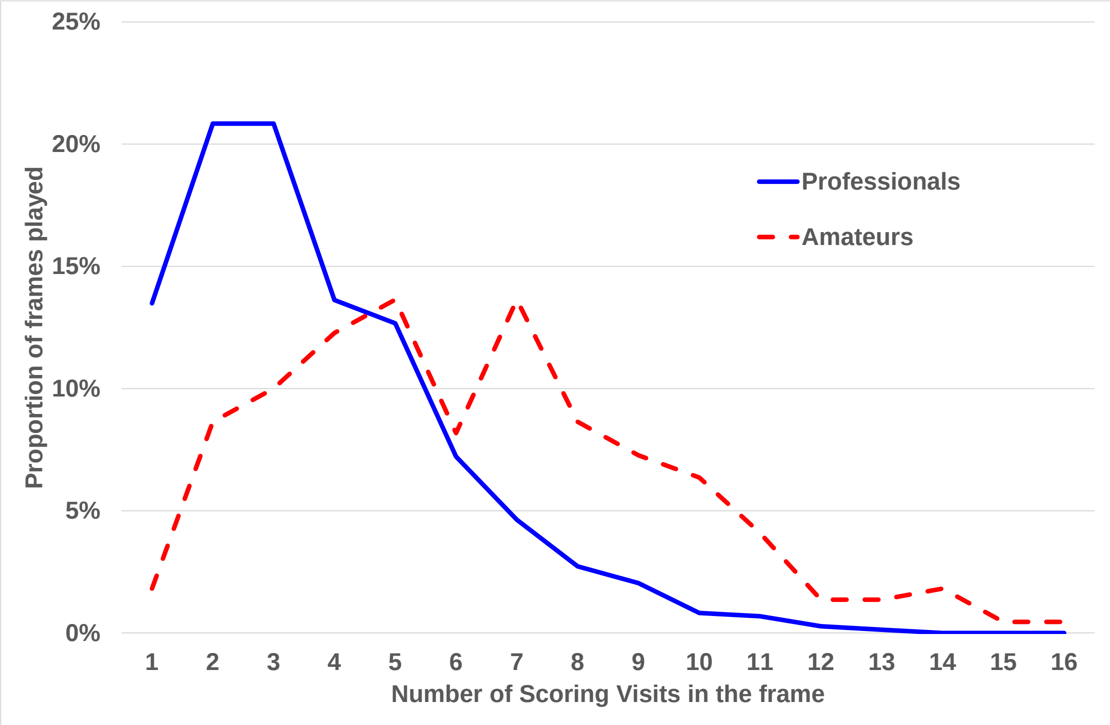


Figure 3.1: Distributions of the number of scoring visits made per frame during the 2018 World Championship and 2019 Masters (Professionals) and 2020 WSF Open (Amateurs)

Presenting the following two scoring measures during a match would indicate whether one player has had more opportunities to score points and / or has scored more heavily from their opportunities. Both measures are objective and could be produced using data generated by the scoring system.

- **Number of Scoring Visits** – which should be divided by the number of frames played to produce a figure that can be compared across matches.
- **Mean 'average' points scored per Scoring Visit** – total number of points scored from potting balls divided by the number of scoring visits.

Figure 3.2 shows that there is a negative correlation between the two measures; the more points scored at each visit then the fewer visits required to end the frame. The two measures therefore need to be interpreted in conjunction with one another rather than in isolation.

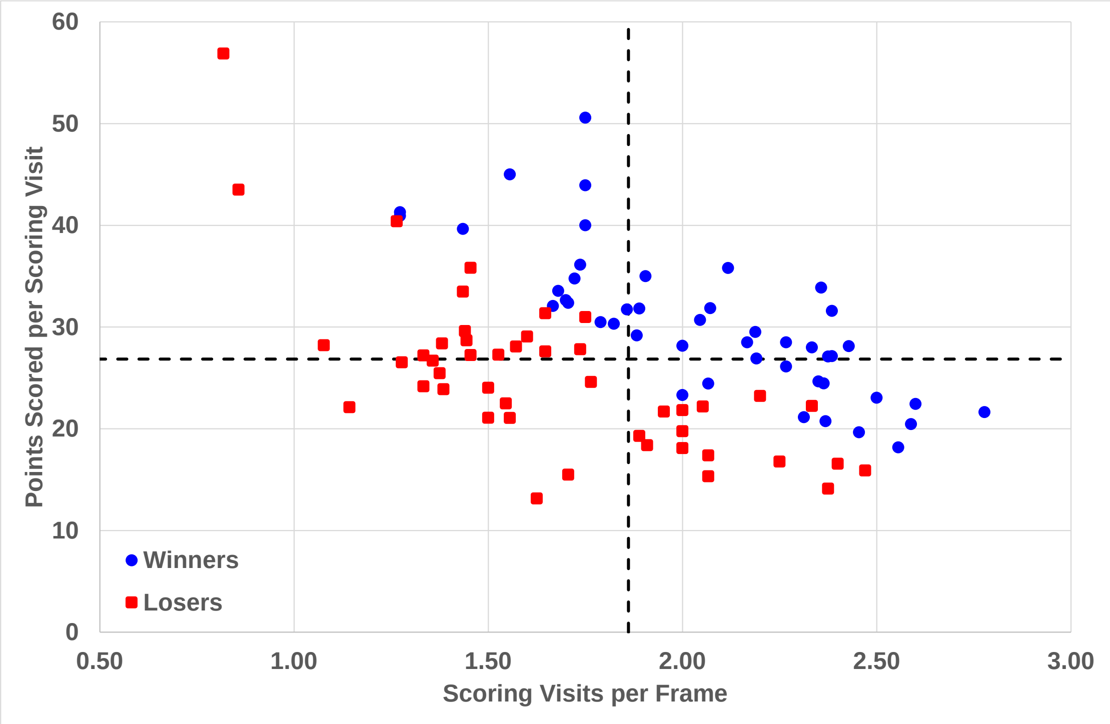


Figure 3.2: The mean number of points scored per scoring visit plotted against the mean number of scoring visits per frame, for the winners and losers of 46 professional matches

During the 2018 World Championship final, Mark Williams made almost 50% more scoring visits than John Higgins (Table 3.4). Higgins scored very heavily when he did get in (his average of 40 points per scoring visit was the highest of the tournament) but it wasn't quite enough to make up for the relatively limited number of chances he was getting.

Table 3.4: Alternative summary statistics for the 2018 World Championship Final collated from post-match video analysis

Measure	Mark Williams	John Higgins	Tournament Average
Frames Won	18	16	
Scoring Visits	64	43	
Scoring Visits per frame	1.9	1.3	1.9
Mean Points per Scoring Visit	29	40	26

3.4.2 Quality Measures – Match Dynamics

Thinking about a match in terms of scoring visits also helps us to describe a simple model for the dynamics of a snooker match. There are essentially two phases of the game – firstly, a player seeks to create a scoring opportunity for themselves (while restricting chances for their opponent); secondly, once they have an opportunity, they want to score as many points as they can. The opportunity presented on the first shot of a visit is largely dependent on the outcome of the opponent's shot; thereafter it is primarily down to the player's ability to build a break.

Distinguishing between these two states we can set out a simple transition matrix for a snooker match, as shown in Table 3.5. In this model, P_{iN} is the probability that Player i pots a ball given that they are taking the first shot of a new visit. P_{iC} is the probability that Player i pots a ball given that this is a continuation of their visit (having potted a ball in their previous shot).

Table 3.5: A simple transition matrix describing the progression of a snooker match from one shot to the next

	Player 1:		**Player 2:**	
	New Visit	**Continuation**	**New Visit**	**Continuation**
Player 1: New Visit	0	P_{1N}	$1 - P_{1N}$	0
Player 1: Continuation	0	P_{1C}	$1 - P_{1C}$	0
Player 2: New Visit	$1 - P_{2N}$	0	0	P_{2N}
Player 2: Continuation	$1 - P_{2C}$	0	0	P_{2C}

P_{iN} can be generated directly from the scoring system. This can be viewed as a player's ***Scoring Potential*** as it represents the proportion of visits to the table in which the player potted at least one ball. It is really a composite measure of one player's ability to deny their opponent opportunities to score points (through good safety play), and the other player's ability to create an opportunity (e.g. through a good pot) so isn't a pure indicator of either player's performance. It is, however, a reasonable measure of how play has developed during a match and offers a good starting point for further development.

P_{iC} is slightly different to a player's Pot Success rate as it does not include all attempted pots and also depends on a player's ability to retain good position for the next pot. It is a measure which can be generated from the scoring system and is produced by downloadable scoring apps such as *mysnookerstats* (Guest, 2010). This can be viewed as a player's ***Scoring Power*** as it reflects the player's ability to make the most of their opportunities to score points. It is directly related to the mean points scored per scoring visit; albeit based on balls potted in succession rather than points scored.

Table 3.6 summarises the measures of Scoring Potential and Scoring Power for match winners and losers, looking separately at data available from professional and amateur levels.

A clear majority of matches were won by the player with the higher outcome rates, while the paired t-test showed a significant difference between the rates achieved by the winners and losers of each match.

There was a significant difference between the Scoring Power of the top professionals (89%) and the leading amateurs (79%), which equates to potting around 10 balls per scoring visit compared with 5. The measure of Scoring Potential is much more even, although comparing rates across matches (let alone tournaments) is less meaningful given the composite nature of this measure.

Table 3.6: Analysis of the Scoring Potential and Scoring Power recorded during 46 matches played across the 2018 World Championship and 2019 Masters (Pro), and 56 matches played during the 2020 WSF Open[1] (Amateur)

	Scoring Potential		**Scoring Power**	
	Pro	**Amateur**	**Pro**	**Amateur**
Overall rates (Total shots analysed)	27% (10,147)	25% (5,482)	89% (21,487)	79% (6,329)
Mean Performance of Match Winners (and standard deviation)	30% (6%)	31% (10%)	89% (3%)	82% (6%)
Mean Performance of Match Losers (and standard deviation)	24% (4%)	22% (7%)	87% (4%)	70% (13%)
T statistic[2]	5.90	5.86	4.28	7.03
p-value	<0.001	<0.001	<0.001	<0.001
Matches won by the player with the higher success rate (with 95% confidence interval)	85% ± 10.4%	79% ± 10.7%	78% ± 11.9%	88% ± 8.7%
Point-biserial correlation with match outcome (Win / Loss)	0.50	0.43	0.40	0.54
Correlation with proportion of frames won in the match by each player	0.54	0.54	0.49	0.57
Correlation between the difference in outcome rates and margin of victory	0.38	0.66	0.48	0.37

[1] Note that the match lengths for the WSF Open (Best of 5 frames in early rounds, with the final the Best of 9) are much shorter than The Masters (Best of 11 frames before the final, which is Best of 19) and the World Championship (Best of 19; rising to Best of 35 in the final)

[2] Paired t-tests of the success rates observed for match winners and losers were carried out across 45 matches (Pro) and 55 matches (Amateur)

3.4.3 Quality Measures – Further Development

This analysis has demonstrated that meaningful statistics relating to scoring visits, along with measures of Scoring Potential and Scoring Power could be generated from the scoring system, with the potential to provide valuable insight into the relative performance of individual players and how matches are won and lost at all levels of the professional game.

With a larger pool of data there may be ways in which these measures could be refined to indicate if / how shot outcomes vary at different stages of the break and / or different stages of a frame, but there is clearly a limit to what we can infer without additional manual input helping to join the dots in the data.

Using this framework as a basis, the most useful information to add manually would be whether the player was attempting a pot or playing a safety shot. As discussed in Section 3.3. this introduces a small amount of subjectivity regarding how shots-to-nothing are classified, so it would be necessary to agree a definition between observers and users of the data. This approach allows for a wider range of measures to be produced, as shown in Table 3.7.

The Pot Success measure used here is comparable with the one produced for the official statistics (I identified slightly fewer missed pots). An additional split is provided to distinguish between pots attempted at the start of a new visit, and those as part of an ongoing break. This first of these splits is preferred to the existing Long Pot Success rate as it is more objective and focusses more directly on a player’s attempt to establish a scoring opportunity.

Table 3.7: Alternative statistics for the 2018 World Championship final collated from post-match video analysis

Measure	Mark Williams		John Higgins	
Frames Won	18		16	
Scoring Potential	34%	(64 / 189)	23%	(43 / 188)
Scoring Power	90%	(492 / 547)	93%	(442 / 474)
Pot Success	91%	(556 / 612)	92%	(485 / 527)
- New visit	73%	(64 / 88)	68%	(43 / 63)
- Continuation	94%	(492 / 524)	95%	(442 / 464)
Scoring Visits[1]				
- Ending with Missed Pot	58%	(32 / 55)	69%	(22 / 32)
- Ending with Safety	42%	(23 / 55)	31%	(10 / 32)
Outcomes of Safety Shots				
- No balls potted	87%	(122 / 141)	77%	(117 / 152)
- No chance left	79%	(111 / 141)	63%	(96 / 152)
Outcomes of Missed Pots				
- No balls potted	57%	(32 / 56)	31%	(13 / 42)
- No chance left	41%	(23 / 56)	24%	(10 / 42)

[1] Note that the bases for these measures are scoring visits which ended with either a missed pot or safety shot. This excludes scoring visits which ended with a successful pot (e.g. on the final black of the frame).

The number (and proportion) of scoring visits which ended with a missed pot and those ending with a safety shot (implying that the player lost position) can be identified. The consequences of missed pots and safety shots can also be analysed by looking at whether the opponent potted a ball (or had a chance to pot a ball on their shot). The cost of each missed pot or failed safety shot could also be quantified by identifying how many points were subsequently scored by the opponent.

These additional measures confirm that Mark Williams left far fewer chances for his opponent following his safety shots (30 v 56). Despite Williams missing more pots than Higgins (56 v 42 according to my data), Higgins had relatively few opportunities to score from them. Indeed, this was a feature of Williams' performances throughout the tournament. He has a reputation for being able to identify attacking shots which offer an opportunity to score while minimising the risk should he miss.

Appendix VI sets out a complete transition matrix with these additional breakdowns, populated with data collected from the 2018 World Championship final. Additional states of 'Break-off' and 'Frame End' are added to capture the start and end of each frame.

3.5 Concluding Remarks

Snooker is a long way behind other sports in the way performance data is used to understand the game and evaluate the strengths and weaknesses of its participants. The biggest issue is the lack of data available, but we have seen how this could be addressed by producing basic statistics for every professional match using the scoring system. The current set of measures - which rely on human judgement to identify the type of shot attempted – could not be replicated but this approach would yield information about how each match was won and lost and the relative strengths of each player.

In the longer term, it is anticipated that greater use of technology will enable analysts to extract more detailed information about the relative difficulty of every shot played and the expected success rates for different types of shot. In the short term there are clear improvements which could be made to the current set of measures which are produced.

The existing scoring measures would be enhanced by presenting data on the number of scoring visits made by each player and the average number of points scored on each of these visits. Together these provide an indication of whether one player took more opportunities to score points than the other and / or whether one player scored more heavily while at the table. This information could all be extracted from the feed generated by the scoring system.

The existing quality measures each have their merits but are limited by the level of human judgement required to produce them. They do not cover all elements of the game, with positional play not captured at all. Shot choice is also a key element of the game but none of the existing measures capture these tactical decisions. As a simple starting point, it would be possible to compare the number of pots and safety shots attempted by different players. If this was available over a larger set of matches this may reveal some information about the playing style of each player.

Considering the game as a series of dynamic interactions between the two players points towards the development of measures based on the outcomes of successive shots rather than simply looking at how an individual shot was executed. The measures of Scoring Potential and Scoring Power described in this analysis can be produced based on data available from the scoring system.

Analysis of a player's safety game is more complicated as we cannot automatically identify when a safety shot was played – so there will always be a place for more subjective measures. Thinking about the progression of a frame in a dynamic way at least indicates how existing measures relating to safety shots and long pots could be enhanced.

4 Simulating the Progression of a Snooker Frame

The third objective of this research is to model a snooker frame to explore how the relative strengths of the two players affects the way the frame progresses and our expectation of likely outcomes. This expands on the analysis presented in Chapters 2 & 3 by developing a framework to understand how differences in performance affect the progression of a snooker match and why the top players achieve better results than the rest.

This chapter sets out the analysis produced to create and validate a shot-by-shot simulation of a frame. A series of inputs are used to determine the probability of potting a ball on each shot based on the stage of the frame and the status of the current visit. The main output of the model is an estimate of the chances of either player winning the frame, along with additional statistics reflecting the progression of a frame based on the outcome of each shot played and every scoring visit made.

Section 4.1 summarises existing analysis of snooker and the use of Monte Carlo simulation to model sporting contests. Section 4.2 describes the conceptual model (§4.2.1) and the data used (§4.2.2) to generate the model inputs (§4.2.3 & §4.2.4).

Section 4.3 summarises analysis carried out to verify and validate the results of the model, comparing the simulated outcomes with the observed data with respect to the outcomes of shots played at different stages of the frame (§4.3.1), and the progression of the frame (§4.3.2). The estimated probabilities of a player winning based on the state of the frame at the start of a scoring visit are also compared with the actual outcomes of frames played (§4.3.3). The deciding frame of the 1994 World Championship is presented as a case study, showing how the model can be used to chart the progression of a frame by looking at how the chances of each player winning fluctuated from one shot to the next (§4.3.4).

Section 4.4 describes sensitivity analysis carried out by applying a scaling factor to the model inputs which modifies the relative strength of each player; controlling separately for a player's ***scoring potential*** and ***scoring power*** as defined in Chapter 3. Inputs relating to each player's scoring power are adjusted to compare the expected progression of a frame depending on the ability of both players (§4.4.1). The scoring power (§4.4.2) and scoring potential (§4.4.3) of a single player are then adjusted relative to the other to analyse the effect on the likely outcome of the frame.

Section 4.5 describes potential applications of the model to support a player's decision making at key points during the frame. It first looks at the potential benefit for a player should they be able to increase the effectiveness of their break-off shot and reduce the chances of their opponent potting the first ball (§4.5.1). It then looks in more detail at the simulated outcomes of a frame and how this varies depending on the status of the frame and which player pots the next ball (§4.5.2). This is used to identify recommend strategies for situations where 1) a player faces a choice between a safety shot or a risky pot, and 2) a player is snookered and has to weigh up the ease of the escape against the risk of leaving their opponent certain chance pot (§4.5.3).

Concluding remarks are presented in Section 4.6.

4.1 Introduction

Snooker is a sport which lends itself well to detailed analysis. The two players take turns visiting the table, making it easier to assess the outcome of each shot and the interaction between successive shots. Shot choice plays a key part in the game as players weigh up the potential benefits from an attacking shot against the increased risk of handing their opponent an easy opportunity to score points. Despite this there are few articles focussed on the analytical side of the game.

4.1.1 Snooker

A description of the game of snooker is provided in Section 1.2.1, with existing literature discussed in Section 1.4.1. No previous analysis has been published looking at the progression of a frame, although some consideration has been given to the strategies involved in shot selection.

Percy (1994) demonstrated how Bayesian methods of predictive inference could be applied, albeit acknowledging the difficulties in obtaining enough data relating to each shot to enable a posterior distribution to be generated – a problem which still exists! In a later paper he discussed how dynamic learning could be used to inform stochastic processes for strategy selection and outcome prediction across a variety of sports (Percy, 2015).

There is a limited amount of data available for analysis. An automated scoring system is used at all professional tournaments, as well as some of the leading amateur tournaments. This automatically updates the score displayed at the venue, while the feed (for professional events) is also used to update World Snooker's live scoring service (World Snooker Tour, n.d.), as well as being sent to bookmakers for use on their websites. For commercial reasons the outputs from this are not publicly available.

Scoring apps can also be downloaded by players to help track their own games. The creators of the scoring app *mysnookerstats (MSS)* have carried out analysis of a variety of players to understand how performance varies among different levels of player (Guest, 2010). The MSS Ratings shown in Table 4.1 are based on the proportion of successful pots by a player which are followed by another. In this research this is referred to as a player's ***scoring power***, as it is effectively a measure of a player's break-building capability.

Table 4.1: The playing standard reflected by different levels of performance as captured by mysnookerstats (MSS) ratings.

MSS Rating	Playing Standard
92+	The very best in the world (only Ronnie O'Sullivan at his best!)
87 – 91	The World's Top 16
83 – 87	The remainder of the Professional Tour players
80 – 83	Q-School hopefuls[1]
75 – 80	Strong amateurs
70 – 75	Good league players
60 – 70	League players
40 – 60	Club players
< 40	Beginners / Novices

[1] Q-school refers to the Qualifying school for the Professional Tour, an annual event with the top performers invited to join the Tour at the start of the following season.

Source: From *Why Should I Use MySnookerStats?* by Guest (n.d.) (https://www.mysnookerstats.com/why-mysnookerstats/). Copyright 2007-2021 by MySnookerStats.com

4.1.2 Modelling sporting contests

Monte Carlo analysis is commonly used to simulate the results of sporting contests; typically to model the outcome of a larger tournament with the purpose of evaluating expected results (Koning et al, 2003), different tournament designs (Scarf et al, 2009) or scoring rules (Scarf et al, 2019). There are, however, surprisingly few examples where Monte Carlo simulation has been used to model the constituent parts of a scoring contest.

Freeze (1974) described an early model of baseball matches by simulating the outcome of individual pitches, which he used to assess how much variations to a team's batting order affected the expected outcome of a match. More recent studies into baseball, such as Hirotsu and Bickel's analysis of the sacrificial bunt (2019), have used a Markov model based on the transitions between states from one batter to the next.

Markov models are often used to predict the results of tennis matches (Spanias and Knottenbelt, 2013). These use the simplifying assumption that points won on serve are near enough independent and identically distributed (Klaassen and Magnus, 2001). A Markov model has also been used to determine the outcome of successive balls in ten-pin bowling (vanDerwerken and Kenter, 2018).

Swartz et al (2009) used Monte Carlo simulation to model runs scored in a one-day cricket match; simulating the outcome of each delivery using characteristics of the batsman and bowler, as well as factors such as the number of balls bowled and the number of wickets lost. They posed a range of questions relating to team selection and batting order which could be addressed using the model.

Alternative approaches to modelling a cricket match have been utilised to arrive at optimal solutions to specific problems; including dynamic programming to determine optimal scoring rates (Clarke, 1988) and simulated annealing to determine optimal batting line-ups in Twenty20 cricket (Perera et al, 2016).

Broadie and Ko (2009) used simulation to model the outcomes of golf shots based on parameters relating to both the golfer and course layout. They were particularly interested in isolating the effect of changing one skill parameter (e.g. driving distance) while holding others constant.

In a similar way, the flexibility of Monte Carlo simulation makes it an ideal choice for modelling a frame of snooker and as the starting point for more detailed analysis of the game and alternative shot choices available to players. A range of different outputs can be extracted depending on the scenario chosen and inputs can easily be adjusted to represent different levels of player.

4.2 Model Development

This section describes the model developed and summarise the data analysis carried out to identify the inputs required.

4.2.1 Conceptual Model

A key decision in the development of any simulation model is to determine its scope. It is not necessary to model all potential outcomes of a shot if some rarely happen, such as potting 2 reds on the same shot. The model described in this research is limited to two outcomes for each shot; either a single ball is potted, or no balls are potted.

Most notably, I have chosen not to model any shots which result in a foul (around 2% of observed outcomes). A key consequence of this is that once one player requires snookers, they cannot win the simulated frame. Upon reaching this state, the simulation continues until the current scoring visit is completed, but no subsequent visits are modelled.

The default starting position for a frame is that there are 15 reds on the table, giving a potential 147 points remaining (based on the black being potted after each red). Both players begin with 0 points. The model has been set up so that alternative starting positions can be specified.

Table 4.2 summarises the process the model follows depending on whether a ball is potted or not. This varies slightly depending on whether the player was attempting to pot a red [1], a colour following a red [2], or a colour at the end of the frame once all reds have been potted [3].

The required inputs to the model therefore relate to the probability of potting a ball on each shot, and (for colours played after a red) the choice of colour played.

Table 4.2: The flow of the simulation model for each type of shot played depending on whether a ball is potted or not

	Shot Played	Process if ball is potted	Process if ball is not potted
[1]	Player ***i*** takes a shot at a red	The number of points scored by Player ***i*** increases by 1 The number of reds remaining reduces by 1 The number of points remaining reduces by 1 Player ***i*** takes a shot at a colour **[2]**	If the difference in points scored exceeds the number of points remaining, then the frame ends Otherwise, Opponent takes a shot at a red **[1]**
[2]	Player ***i*** takes a shot at a colour (with value ***c***)	The number of points scored by Player ***i*** increases by ***c*** The number of points remaining reduces by 7 If reds remain on the table, Player ***i*** takes a shot at a red **[1]** Otherwise, Player ***i*** takes a shot at the yellow **[3]**	The number of points remaining reduces by 7 If the difference in points scored exceeds the number of points remaining, then the frame ends Otherwise, if reds remain on the table, Opponent takes a shot at a red **[1]** Otherwise, Opponent takes a shot at the yellow **[3]**
[3]	Player ***i*** takes shot at yellow / green / brown / blue / pink / black (with value ***c***)	The number of points scored by Player ***i*** increases by ***c*** The number of points remaining reduces by ***c*** If there are 0 points remaining, then the frame ends Otherwise, Player ***i*** takes a shot at the next colour, with value ***c + 1*** **[3]**	If the difference in points scored exceeds the number of points remaining, then the frame ends Otherwise, Opponent takes a shot at the colour **[3]**

4.2.2 Data

To determine suitable inputs to the model, post-match video analysis has been used to collect data on 31,298 shots played during all 734 frames contested over the 46 matches in the 2018 World Championship finals and 2019 Masters. A list of the matches analysed is provided in Appendix III. Both tournaments are contested by the 16 highest-ranked professional players, with an additional 16 players qualifying for the World Championship after progressing through three qualifying rounds. The data gathered therefore reflect the play of top professionals. To be consistent with the scope of the model, shots played during visits which started after one player required snookers are not included in the analysis presented in this chapter.

For each shot I recorded the number of reds (and points) remaining on the table. I also noted whether the player was taking the first shot of a new visit or continuing their visit (having previously potted a ball). For new visits I calculated the number of shots since the last pot, while for continuing visits I identified the number of shots (successful pots) already played. The outcome recorded for each shot was whether a ball was potted; and if so, how many points were scored. A sample of the data recorded is provided in Appendix IV.

Although my initial data collection recorded what type of shot was played (e.g. safety shot or attempted pot) I have not used this in my analysis. Nor have I recorded any information about the position of the remaining balls, or the distance between the cue ball and object ball. The data used are therefore comparable to the information which would be available from an automated scoring system.

4.2.3 Model Inputs – Probability of potting a ball

The first set of inputs required are estimates of the probability that a ball will be potted on any shot played in the frame. Chi-square automatic interaction detection (CHAID) analysis has been carried out on the data collected to inform my approach (as used by Méndez-Domínguez et al, 2019 to identify the best predictors of ball possession efficacy when using the goalkeeper as an outfield player in elite futsal), although I have ultimately used my judgement in determining which are the most appropriate dynamics of a snooker frame to capture. A list of all inputs is provided in Appendix VII. The results of chi-squared tests are presented alongside each of the inputs to support the decisions made.

I first split a frame into different phases depending on how far it has progressed. While this would not be the first factor chosen by CHAID or any other decision tree methodology, it offers the most logical starting point for modelling the progression of a frame.

At the start of the frame the reds are packed together so safety shots are played until there is a chance to pot the first red. As the frame develops the reds become more spread out and there is a greater chance of a player potting a ball and establishing a break. Reds which are well-placed tend to be potted first, so the later reds (and the final red in particular) can be harder to gain position on.

The first four phases are therefore represented by number of reds remaining on the table when the red was played (***1st Red***; ***Reds 2-4***; ***Reds 5-12***; ***Reds 13-15***). The colours played immediately following each red are referred to accordingly – the colour following the 1st Red etc.

During the fifth and final phase, the reds have all been potted and only the colours remain on the table. There was relatively little difference between the pot rates for each of the colours, so a collective set of assumptions were developed.

The most significant factor in predicting the outcome of each shot during each phase of the frame was found to be the current status of the visit. If a player has just potted a ball, then it is far more likely that they will be in position to pot the next one than if they have just come to the table following a shot by their opponent.

A player's options when starting a new visit are highly dependent on the previous shot played by their opponent. If a pot is missed, then it is more likely that a scoring opportunity will be left than if they played safe. I account for this indirectly in my model as a player was found to be more likely to pot a ball on a new visit if they were countering a scoring visit from their opponent than if there had been an exchange of shots since the last ball was potted. Aside from this, the probability of potting the next ball was found to be independent of the length of the exchange.

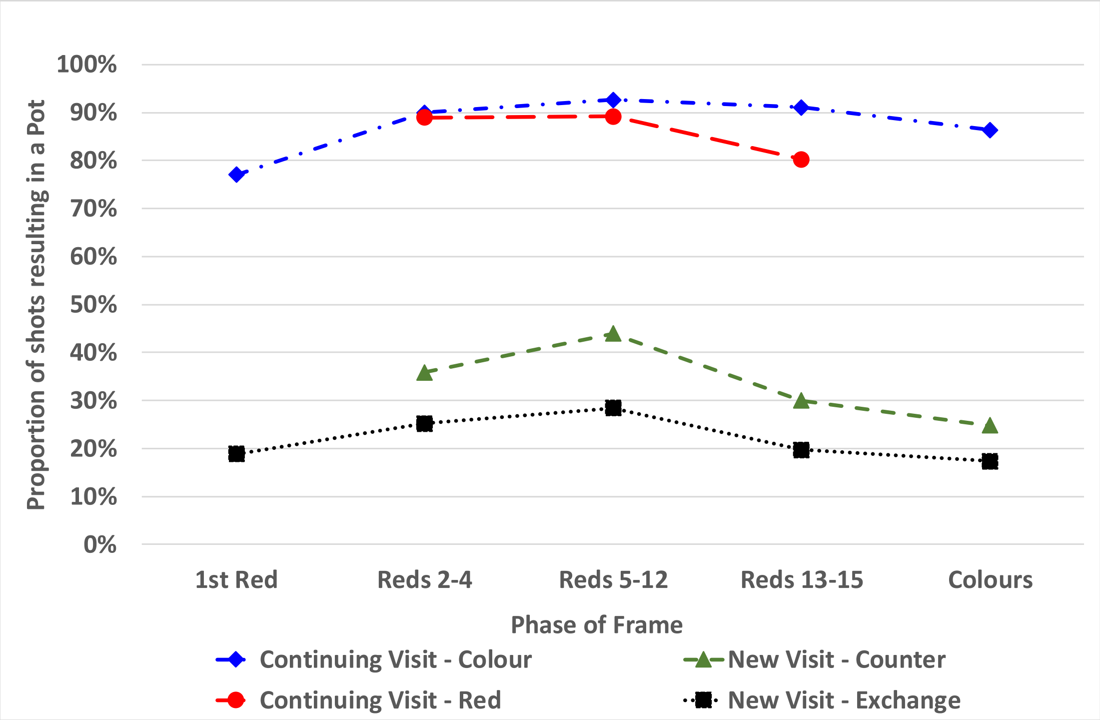


Figure 4.1: Proportion of shots which resulted in a pot depending on the status of the current visit and the phase of the frame

Figure 4.1 shows the proportion of shots played which resulted in a pot over different phases of the frame and how this depended on whether the player was continuing an existing visit or starting a new visit. For continuing visits there were slightly higher probabilities of potting the colours rather than the reds. For new visits there was a higher probability of potting the ball when the opponent had potted a ball on their previous visit (counter) than if they had not (exchange).

Looking at the last colour potted in the visit provided an additional indication of how likely a player was to pot the next red. If a player had just potted the pink or black then they were more likely to pot the next red than if they had potted the one of the baulk colours (yellow, green or brown) – in which case there is often further for the cue ball to travel to retain good position for the next red.

Towards the end of a frame, the more established the scoring visit was, the less likely a player was to miss. The length of the current visit is effectively providing an indication as to how open the frame has been and how well positioned the remaining balls are likely to be.

A colour was less likely to be potted following the first red in a scoring visit than after subsequent reds, likely reflecting the relative difficulty of the pot on the initial red and the resultant impact on a player's ability (and ambition) to retain good position on a colour. This was particularly true nearer the start and end of each frame.

As the frame progressed, the probability of potting a colour tended to increase. The exception was for the colour following the penultimate red, where the player may have to try and develop the final red while potting the colour – although for more established breaks (perhaps where the frame is already won) this appeared to be less of an issue.

4.2.4 Model Inputs – Choice of colour following a red

In modelling the colours played following each red, the model first determines whether a colour was potted and, if it was, only then selects the colour. Although there may be interest in modelling the choice of colour first, that information was typically not collected when no ball was potted – and for safety shots the choice of colour is less relevant.

The value of the colour potted was found to change as an individual visit and the wider frame developed, as shown in Table 4.3.

Table 4.3: Frequency at which each colour was potted after a red at different stages of the frame

	Colour selected (and associated points value)					
Stage of Frame	**Yellow (2)**	**Green (3)**	**Brown (4)**	**Blue (5)**	**Pink (6)**	**Black (7)**
New Visit (1st Red)	10%	13%	8%	30%	12%	27%
New Visit (Reds 2-15)	8%	10%	9%	24%	17%	32%
Continuing Visit (Reds 2-14)						
- Following Baulk Colour	5%	6%	5%	27%	22%	35%
- Following Blue	3%	3%	3%	34%	19%	38%
- Following Pink	2%	2%	1%	19%	46%	30%
- Following Black	2%	2%	2%	17%	13%	65%
Continuing Visit (15th Red)	4%	7%	8%	29%	14%	38%

At the beginning of a new visit (particularly if it was the first red to be potted in the frame), a player was more likely to finish on the blue or one of the baulk colours; perhaps to allow an element of safety in case the red was missed. As the break continued, they were more likely to play for the blue, pink or black; both for their greater points value and for their proximity to the reds. Following the final red, I noted that the blue and baulk colours were again chosen more frequently with the player aiming to finish close to the yellow for their next shot.

An additional dynamic that I identified was a dependency in the choice of successive colours. In particular, if a player potted the black after the previous red then they were more likely to pot the black after the next red as well. Similarly, the pink is often played less frequently at the start of a frame as its spot is closest to the reds, which tend to block the pink's path to the pockets. Once this area of the table is opened up though, the pink can become a more favourable option for retaining position on the reds.

It should be noted that I have not attempted to refine a player's choice of colour in the model depending on the current score (for example, where high-valued colours need to be chosen in order to reach / avoid reaching the situation where snookers are required).

4.2.5 Implementation

The model was developed in R using the shot probabilities listed in Appendix VII and the colour choice distributions presented in Table 4.3.

To validate the model, 10,000 simulations were run of a full frame, starting with 15 reds on the table and the score at 0-0. For subsequent scenarios described in this research, 10,000 simulations were again run, with the starting point dependent on the scenario being studied.

4.3 Verifying and Validating the Model

This section sets out the analysis carried out to check that the model is working as expected and is yielding an accurate representation of a frame of snooker.

The outcomes of different shots played within a frame are analysed in Section 4.3.1 and compared with the input assumptions used. Statistics capturing the progression of the observed and simulated frames, such as the number of scoring visits made, the points scored during these visits and the number of shots played between each scoring visit are then compared in Section 4.3.2.

Section 4.3.3 considers the situation at the start of each scoring visit in respect of the current score and the number of points remaining on the table. The proportion of frames won from similar situations are then compared between the observed and simulated frames.

As a case study for how the simulation model can be used to track the course of a frame, the model is repeatedly run to replicate the situation prior to every shot played in the deciding frame of the 1994 World Championship Final, showing how the chances of each player fluctuated as the frame progressed (Section 4.3.4).

4.3.1 Shot Outcomes

The model uses assumptions about the probability that a player will pot a ball on each shot, so we would expect the overall success rates to be similar. Table 4.4 summarises the outcomes of shots played at different stages of the frame. Pearson's chi-squared test was carried out to assess whether the number of shots resulting in a pot was consistent between the observed and simulated values. None of the differences were significant at the 95% level.

Table 4.4: Modelled and observed outcomes of shots played at different stages of a frame; based on 734 observed and 10,000 simulated frames

			Shots played on new visit				Significance test	
			Observed data		Simulation			
Phase of frame	Object ball	Status of visit	#	% Pots	#	% Pots	chi-sq	p-value
1st Red	Red	New[1]	3,145	23.3	42,881	23.3	0.00	0.99
	Colour	Continuing	731	77.0	10,000	77.0	0.00	1.00
Reds 2-4	Red	New	1,824	28.2	25,736	28.2	0.00	0.99
	Red	Continuing	1,881	89.0	25,645	88.6	0.23	0.63
	Colour	Continuing	2,191	90.0	30,000	90.1	0.03	0.85
Reds 5-12	Red	New	2,853	33.7	38,950	33.4	0.11	0.74
	Red	Continuing	5,380	89.2	73,582	89.7	1.27	0.26
	Colour	Continuing	5,761	92.6	79,006	92.5	0.13	0.72
Reds 13-15	Red	New	1,027	22.0	16,242	21.6	0.11	0.74
	Red	Continuing	1,656	80.3	23,126	80.0	0.07	0.79
	Colour	Continuing	1,554	91.1	22,000	90.7	0.32	0.57
Colours		New	646	18.7	13,336	18.4	0.05	0.83
		Continuing	1,913	86.4	28,853	84.9	3.28	0.07
All		New[1]	9,496	26.9	137,145	26.4	1.16	0.28
		Continuing	21,068	89.0	292,212	88.8	0.51	0.47

[1] Excluding the first shot played in each frame

Table 4.5 summarises the number of times each colour was potted following a red. The simulated data is shown to follow a similar distribution to the observed data, with Pearson's chi-squared test showing no evidence that these are significantly different from one another: X^2 (5) = .52, p = .99.

Table 4.5: Colours potted following a red; based on 734 observed and 10,000 simulated frames

	Observed data		Simulation	
Object Ball	**Pots**	**% Pots**	**Pots**	**% Pots**
Yellow	350	3.8%	4,839	3.8%
Green	429	4.6%	5,935	4.6%
Brown	362	3.9%	4,929	3.9%
Blue	2,138	23.0%	29,677	23.2%
Pink	1,918	20.7%	26,558	20.8%
Black	4,090	44.0%	55,830	43.7%

4.3.2 Frame statistics

A stronger test of the model is whether the frames generated follow a similar pattern to the observed data. Outputs related to the first scoring visit made in each frame are evaluated, looking at how many shots were played before the first ball was potted and how many points were scored during the first scoring visit. Similar analysis is then presented for subsequent visits, along with other features of the observed and simulated frames.

First Scoring Visit

A summary of all statistics produced for the first scoring visit is presented in Table 4.6. Where the measures are continuous in nature, t-tests have been carried out to compare the distributions of the observed and simulated data. When comparing the number of observed and simulated frames which did (or did not) produce a particular outcome, Pearson's chi-squared test for homogeneity is used. The player "breaking-off" (taking the first shot in the frame) is denoted as *Player 1*.

Table 4.6: Summary statistics for the first scoring visit made in each frame

	Means / Proportions		Significance test		
	Observed data	**Simulation**	**t-stat**	**chi-sq**	**p-value**
Shot on which 1st red was potted	5.28	5.29	-0.06	-	0.95
Player 1 makes 1st scoring visit	47.3%	45.4%	-	0.96	0.33
Points scored in 1st scoring visit	30.2	30.8	-0.44	-	0.66
50+ Points in 1st scoring visit	21.4%	23.4%	-	1.50	0.22
100+ Points in 1st scoring visit	6.9%	6.1%	-	0.77	0.38
Total Clearances[1]	3.1%	3.1%	-	0.00	0.99
% Frames 'won' on 1st scoring visit	15.1%	15.9%	-	0.31	0.58

[1] A total clearance is when a player pots all 15 reds with colours and the final 6 colours in the same visit (a total of 36 pots – unless multiple reds are potted in the same shot)

The distributions for the shot on which the first red was potted are shown in Figure 4.2, with a close match between the observed and simulated frames. The observed probabilities of potting a ball on Shot 1 (0.00) and Shot 2 (0.20) are used directly by the model. The simulated data is then based on the assumption that from Shot 3 onwards the probability of potting a red is 0.24. The observed proportion of balls potted on Shot 3 was actually slightly higher than this (0.27), but the difference was not significant and it was not incorporated into the model, resulting in the only notable deviation between the distributions obtained.

Based on the input assumptions used for the model we can calculate mathematically that *Player 1* has a 45.5% probability of potting the first ball in the simulation, putting them at a slight disadvantage at the start of the frame. Section 4.1 considers the effectiveness of the break-off shot and how much this influences the outcome of a frame.

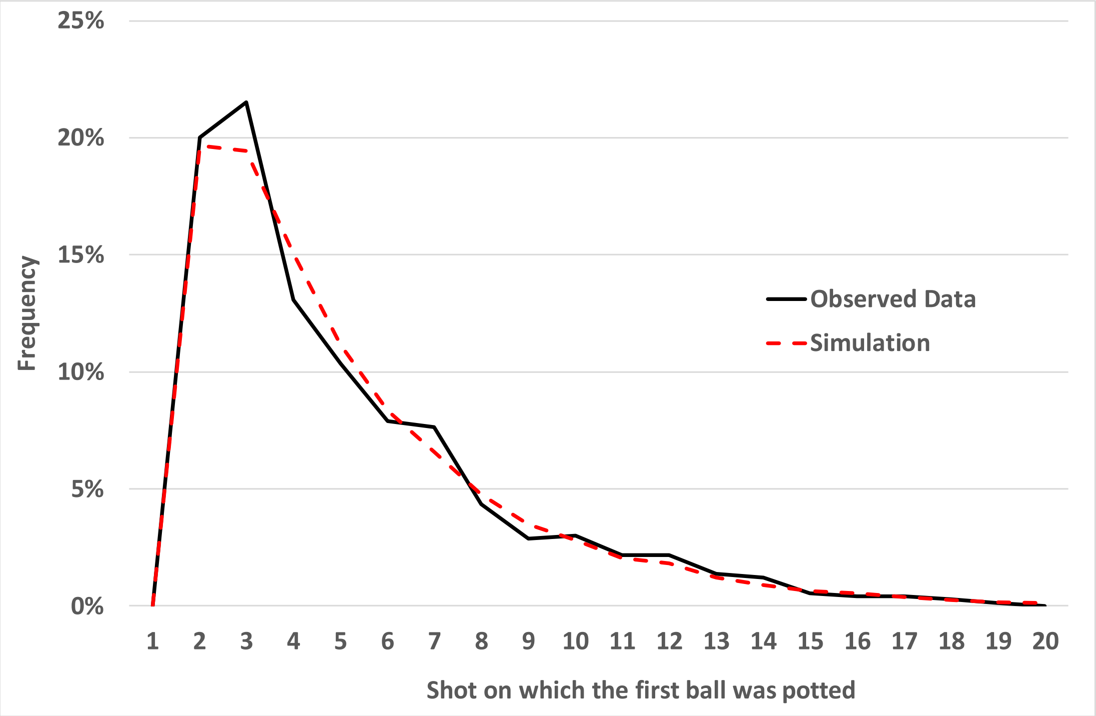


Figure 4.2: Distribution of shots played when the first red is potted in each frame

The number of points scored during the first scoring visit also followed very similar distributions for the observed and simulated data (Figure 4.3). There were a slightly higher proportion of visits in which 50 or more points were scored in the simulation, although a slightly lower proportion in which 100 or more were scored – neither difference being significant. There was also no evidence of a significant difference in the number of frames reaching the "snookers required" stage – i.e. effectively 'won' - by the end of the first scoring visit.

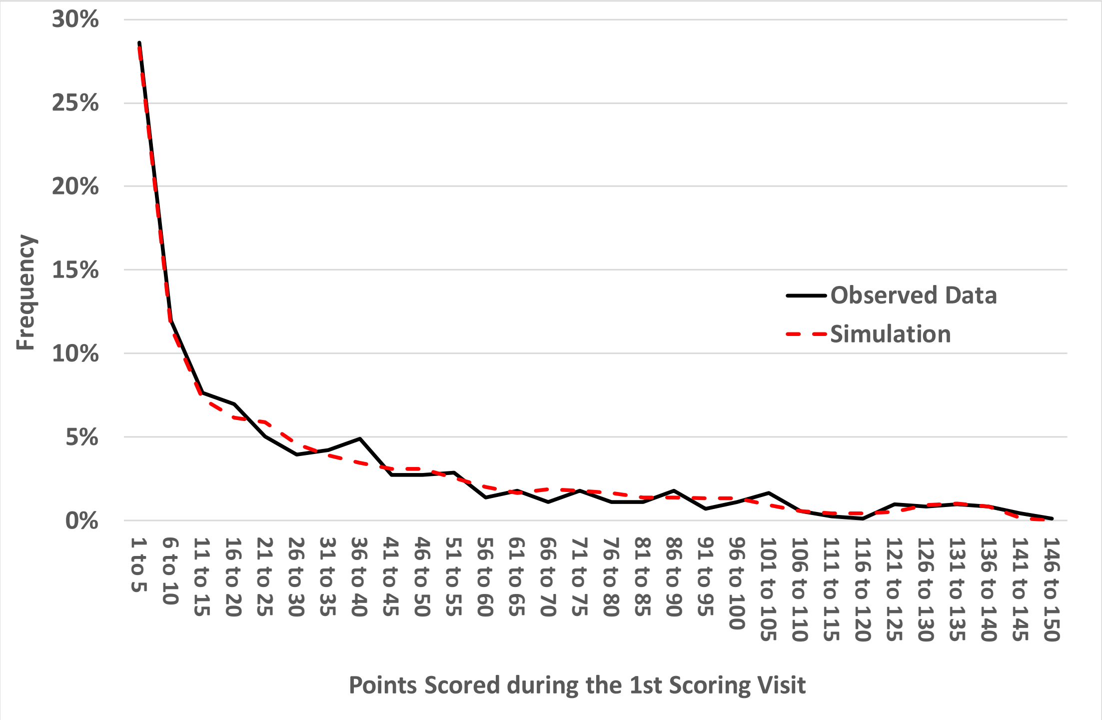


Figure 4.3: Distribution of points scored during the first scoring visit in each frame

Frame Outcomes

Table 4.7 contains statistics relating to subsequent scoring visits made during the frame (before snookers were required) and additional summary statistics relating to the frames recorded.

Table 4.7: Summary frame statistics

	Means / Proportions		**Significance test**		
	Observed data	**Simulation**	**t-stat**	**chi-sq**	**p-value**
Shots between scoring visits	3.47	3.59	-1.41	-	0.16
Points per scoring visit (Visit 2+)	27.0	26.4	0.88	-	0.38
% Frames won by *Player 1*	48.5%	48.9%	-	0.05	0.82
Scoring visits per frame	3.51	3.62	-1.39	-	0.17
Frames with a 50+ break	68.1%	68.2%	-	0.00	0.98
Frames with a 100+ break	14.7%	12.0%	-	4.56	0.03

With *Player 1* less likely to pot the first ball in the model, they went on to win fewer than 50% of the simulated frames. During a match the break-off shot for each frame alternates between the two players, so we would expect the probability of each player in the model winning a match to tend towards 0.5 the more frames that were played.

A similar proportion of observed and simulated frames contained at least one break of 50 points or more; although there was some evidence of a difference in the number of century breaks made, with the simulation model producing a lower proportion than the observed data.

There is some prestige associated with making a century break (i.e. 100 points or more), so it is possible that players were particularly focused on reaching this milestone. If we instead compare the proportion of frames featuring a break of 90 points or higher, the simulated data is again lower (18.5% of frames compared with 19.8% in the observed data) – but not significantly so: X^2 (1) = .75, p = .39. Similarly, there is also insufficient evidence of a difference in the proportion of frames featuring a break of 110 points or more (7.9% of frames compared with 9.8% in the observed data; X^2 (1) = 3.30, p = .07).

The number of shots played between scoring visits (Figure 4.4) followed similar distributions in both the observed and simulated data; as did the number of points scored on each scoring visit (Figure 4.5).

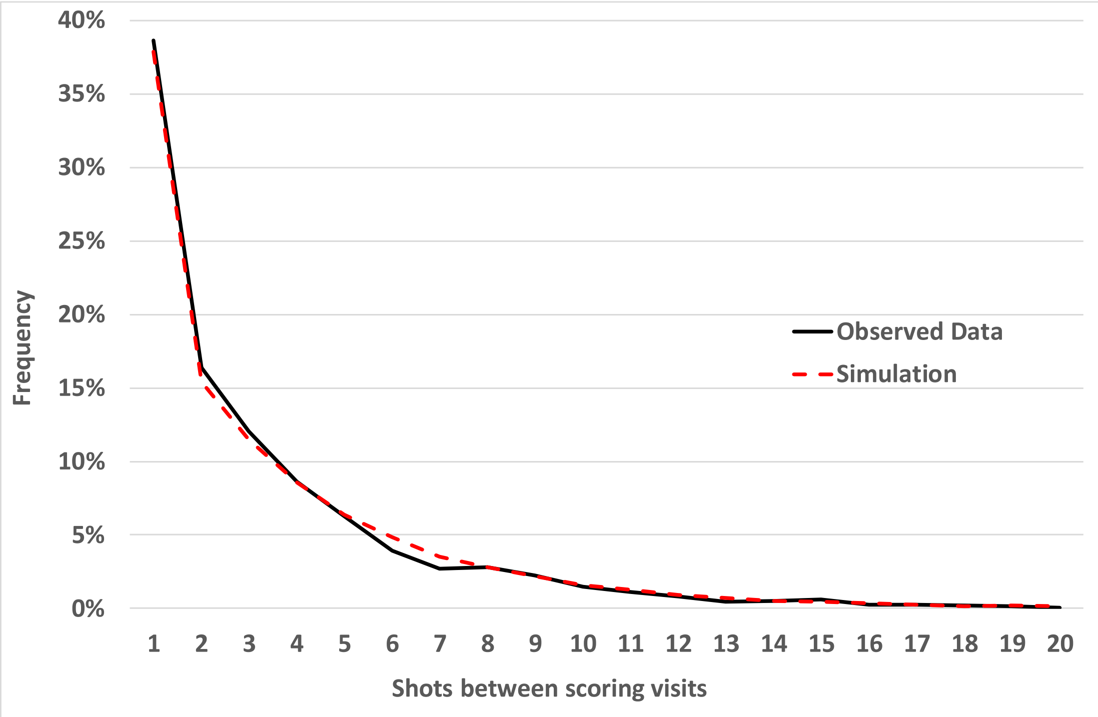


Figure 4.4: Distribution of shots played between scoring visits

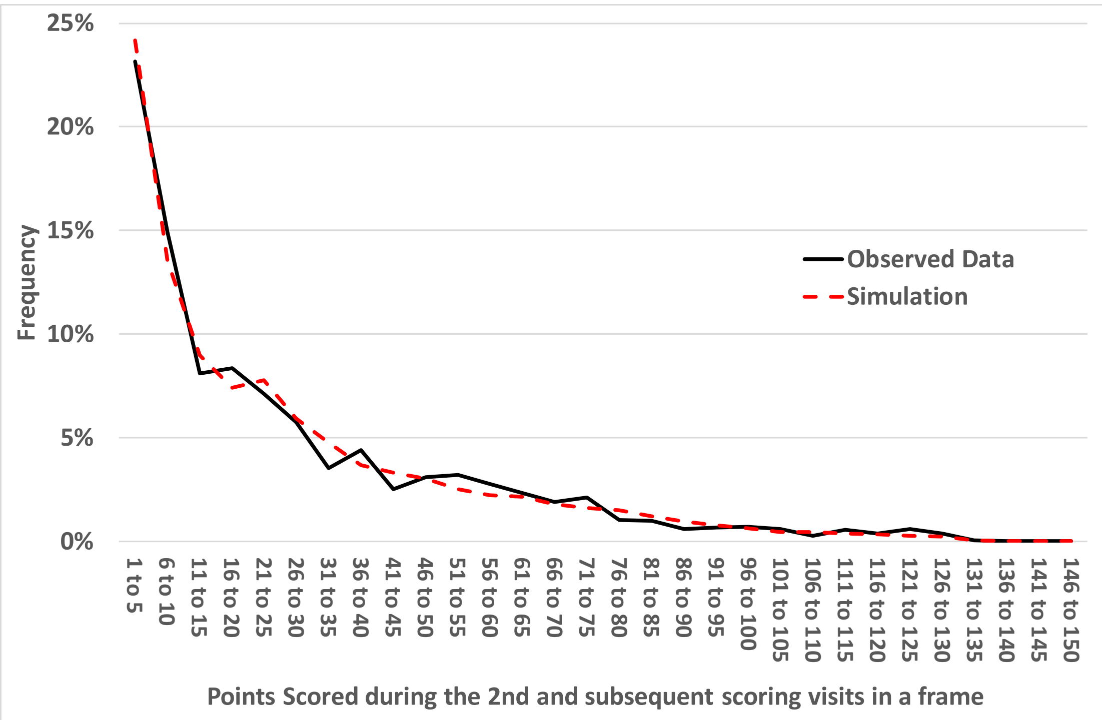


Figure 4.5: Distribution of points scored during the second and subsequent scoring visits in each frame

The distribution of the number of scoring visits played before the frame was won was skewed slightly more to the right in the simulation (although not significantly so). In particular, there were a relatively large number of frames won in exactly 2 or 3 scoring visits during the observed matches (Figure 4.6).

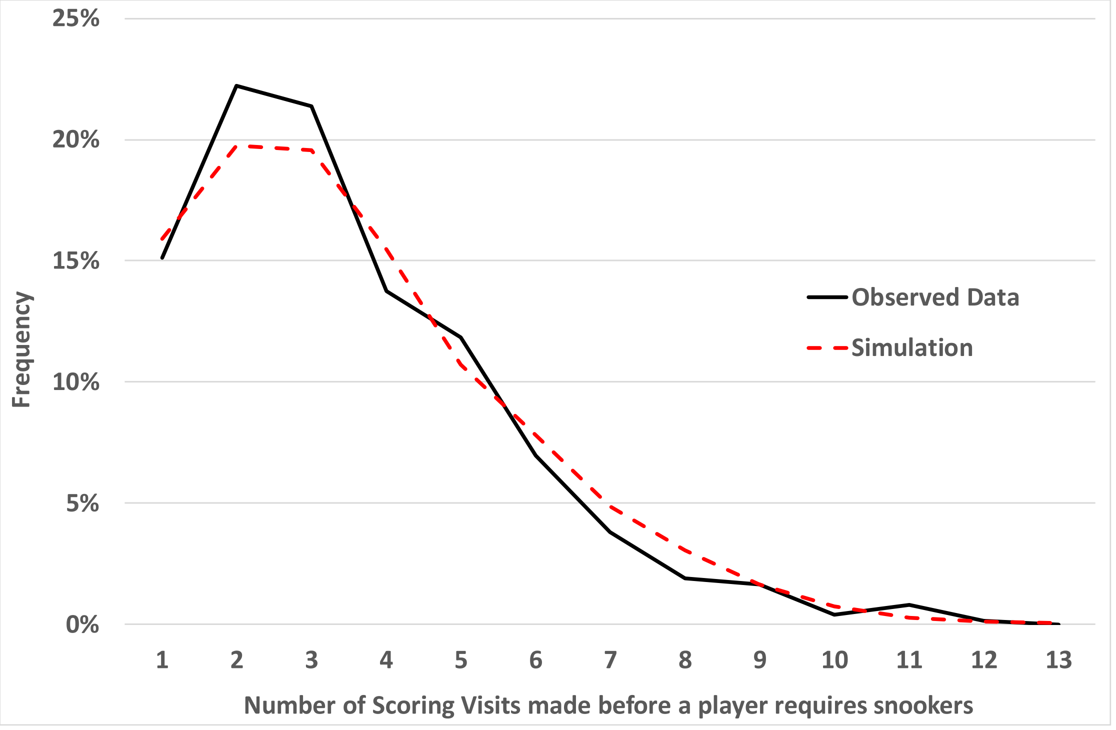


Figure 4.6: Distribution of scoring visits made in each frame before a player required snookers

4.3.3 Win Probabilities

The main output produced by the simulation model is an estimate of the probability that each player will win the frame from a given position. There is insufficient data available from actual matches to compare outcomes from specific situations, but similar instances can be grouped together to provide an indicative comparison of the simulated results against the available observations.

The analysis presented in Table 4.8 considers the position at the start of each scoring visit in terms of the number of points remaining and the current score from the perspective of the player at the table. The proportion of frames which were subsequently won by that player has been calculated for both the observed and simulated data – with figures only presented for a scenario if there were at least 30 such occurrences in the observed data.

Table 4.8: Proportion of observed and simulated frames won from different scenarios under which the next scoring visit started. The figures in brackets show the number of instances analysed in each case.

	Difference in Score						
Points Remaining	**>= 50 behind**	**30-49 behind**	**10-29 behind**	**9 behind – 10 ahead**	**11-30 ahead**	**31-50 ahead**	**> 50 ahead**
Observed Data							
147				60% (727)			
123-139			54% (89)	67% (374)	76% (49)		
91-115		50% (110)	62% (164)	71% (93)	80% (109)	95% (57)	
59-83	37% (65)	40% (82)	47% (55)	70% (64)	89% (55)	98% (54)	90% (40)
7-51			56% (89)	81% (110)	92% (97)	100% (34)	
Simulated Data							
147				63% (10,000)			
123-139			57% (1,189)	65% (5,433)	73% (648)		
91-115		45% (1,457)	57% (2,066)	67% (1,109)	80% (1,443)	89% (774)	
59-83	21% (1,112)	35% (1,102)	54% (815)	72% (799)	83% (744)	93% (871)	99% (601)
7-51			49% (1,559)	74% (1,959)	90% (1,443)	98% (451)	

Grouping the cases together in this way means that the data cannot formally be compared – and a relatively small sample of actual frames are available for analysis – but the simulated results do follow a logical pattern along the same lines as the actual outcomes. The most notable difference is in the relatively high number of actual frames won from 50 points or more behind (24 / 65 = 37%) compared with the simulated data (235 / 1112 = 21%) but further data would be required to assess whether this reflected a limitation in the model or just an anomaly in the frames observed.

The analysis presented in this section shows the value of being "in control" of the table. If one player leads by more than 30 points while reds remain on the table then the model indicates they are very likely to win the frame if they pot the next ball (89% in the case where there are 91-115 points remaining), but their opponent still stands a good chance of winning (45%) if they make the next pot. This is explored in more detail in Section 4.5.2.

4.3.4 Case Study

The simulation model developed has been used to estimate the probability of each player winning based on the situation prior to each shot of the deciding frame in the 1994 World Championship final between Stephen Hendry and Jimmy White.

A high-level summary of the frame is provided in Table 4.9. The pattern of the frame – with both players making two scoring visits – is typical of the frames which informed the inputs to the model, so the level of performance reflected should be reasonably comparable even if overall standards of play are now slightly higher. Figure 4.7 then charts the simulated probability of Stephen Hendry winning the frame at the start of each shot.

Table 4.9: Summary of the final frame from the 1994 World Championship final

#	Summary of scoring visits	Score at end of visit (Hendry - White)
1	White pots the first red (shot #5) but runs out of position after potting the black and plays safe.	0 - 8
2	After White misses a long pot, Hendry has the first real chance of the frame (shot #10). He also loses position though and misses a difficult red (shot #18) to leave a chance for White.	24 - 8
3	White plays a couple of good positional shots to develop the reds and establish a frame-winning opportunity. He misses a straightforward black though (shot #28) to hand the opportunity back to Hendry.	24 - 37
4	Within a couple of shots Hendry is in prime position and completes the clearance to claim frame, match and the Championship.	82 - 37

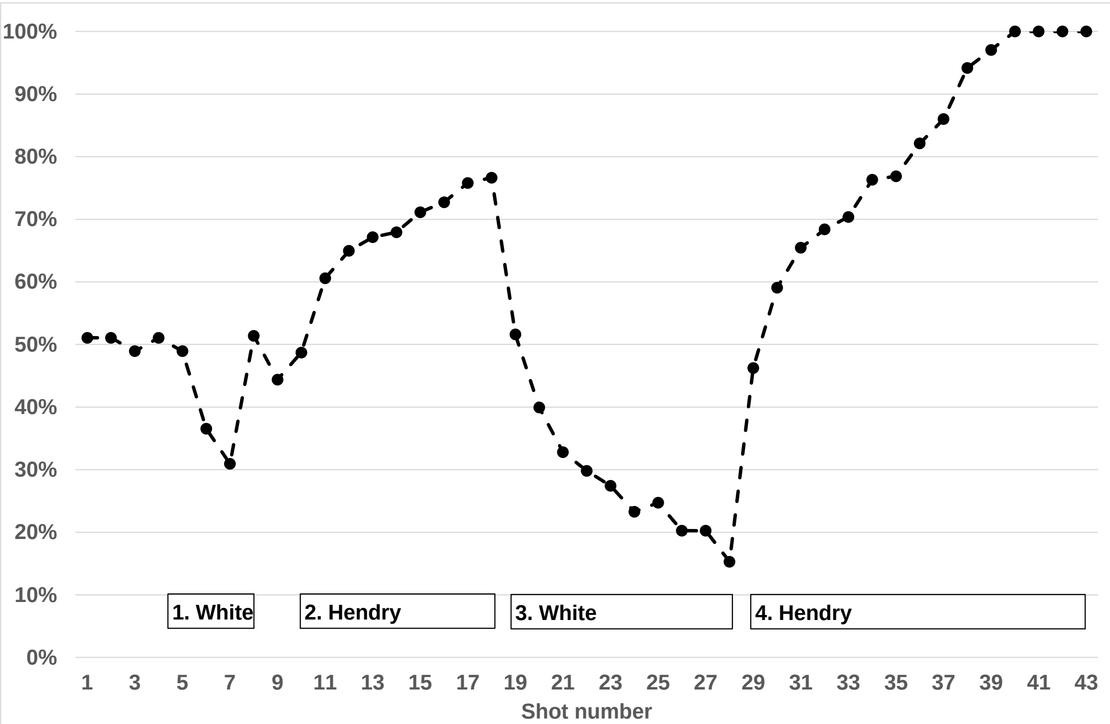


Figure 4.7: Shot by shot simulated probability of Stephen Hendry winning the final frame against Jimmy White during the 1994 World Championship final

White became the clear favourite to win the frame once he had established his second scoring visit. At the point he missed the crucial black, the simulation model gave him an 85% chance of winning the frame. If anything, the opportunity was even more promising than this given the position of the balls.

Similarly, the model was initially cautious about Hendry's chances of winning - on the basis that the last couple of reds can be more difficult to gain position on. In fact, the reds were all out in the open, so the clearance looked to be a formality even before the final red was potted (shot #37).

Nevertheless, simulating the frame in this way provides an informative method for quantifying and depicting how the chances of each player fluctuates over the course of the frame and in assessing the impact of potential turning points.

4.4 Sensitivity Analysis

As well as analysing the outputs from the model under the core set of inputs we are also interested in understanding how sensitive these outputs are to any changes in the inputs. This also helps us to understand how the progression of a frame depends on the ability of the two players.

To assess this, we can view each of the inputs as either representing the probability of potting a ball at the start of a new visit, or as a continuation of an existing visit.

The overall proportion of successful pots at the start of a new visit reflects a player's ***scoring potential***, as described in Chapter 3. In part this reflects a player's potting ability, often where the cue ball is some distance from the object ball. Arguably to a greater extent though, these inputs reflect the opponent's safety prowess / tactical awareness; players with greater capability in this respect (or perhaps just those showing more caution in their shot selection) will be less likely to leave a scoring opportunity for their opponent. A scaling factor is introduced into the model to modify a player's scoring potential by adjusting each of the underlying input parameters by the same proportion.

Collectively, inputs relating to the probability of a player successfully continuing their visit reflect the player's ***scoring power***, another measure introduced in Chapter 3. The greater their skill at manoeuvring the cue ball between shots, the higher their subsequent probability of potting a ball. A second scaling factor is introduced into the model to modify one or both player's scoring power by adjusting each of the underlying input parameters by the same proportion.

Setting both scaling factors at 1.00 reflects the default model described in Section 4.2, broadly representing a Top 16 player.

This section summarises the impact on the progression of a simulated frame from using these scaling factors to modify the input assumptions to reflect different levels of ability for one or both of the players.

4.4.1 Changing the scoring power of both players

The first element of this analysis considers the effect of the break-building prowess of the two players on the progression of the frame. Under each scenario the model was run with a different scaling factor to adjust the scoring power of both players by the same amount, leaving them evenly matched. The scoring potential of the two players was not modified.

The first five values chosen are believed to reflect the range of abilities within the professional ranks. The final four values were chosen to mirror some of the levels identified by *mysnookerstats* (see Table 4.1) – ‘Q-School hopefuls’, ‘Strong amateurs’, ‘Good league players’ and ‘League players’. Additional data would be required to accurately model the ability of different players, but this serves to provide an indication of the effect on frame progression and demonstrate the sensitivity of the inputs used.

Table 4.10 shows the effect of applying different scaling factors. Note that the effect on the scoring power of the players is typically slightly larger than the scaling factors used. This arises from using the length of the current visit as a factor in determining the outcome of shots played towards the end of the frame, so the success of shots played at the start of the frame has an additional influence over the success of subsequent shots played.

Table 4.10: Impact on scoring rates from changing the proportion of balls potted following a pot by both players

Scaling Factor	Scoring Power	Scoring Visits	Points per Scoring Visit	High Breaks		Total Clearances
				50+	100+	
1.04	93.3%	2.49	43.2	88%	31%	13%
1.02	91.0%	3.06	33.8	79%	20%	6%
1.00	**88.8%**	**3.62**	**27.6**	**68%**	**12%**	**3%**
0.98	86.7%	4.19	23.4	58%	8%	1%
0.96	84.6%	4.71	20.3	49%	5%	1%
0.93	81.4%	5.57	16.7	37%	2%	0%
0.89	77.2%	6.68	13.4	24%	1%	0%
0.84	72.2%	8.00	10.8	12%	0%	0%
0.76	64.4%	9.96	8.1	4%	0%	0%

Even a small change to the scoring power of the two players has a pronounced effect on the average points scored on each visit. The nature of the game also changes – at the highest level the model indicates that there would often only be a couple of scoring visits per frame as players score so heavily.

The proportion of frames in which a 50+ or 100+ break is made varies considerably across the professional ranks. The simulated figures are comparable to data extracted from *Cuetracker.net* on century breaks made during the 2020/21 season (Florax, 2021). These show that the 4 highest-ranked players made 100+ breaks in 27% of the frames they won and players in the bottom half of the rankings made 100+ breaks in 6% of the frames they won. [Breaks made in frames won is referenced here as a proxy for the rate expected from two players of equal ability competing against one another.]

4.4.2 Changing the scoring power of one player

The second set of scenarios considers how the chances of a player winning the frame are affected by the relative strength of their scoring power compared with their opponent.

Under each scenario the model was run with a different scaling factor to adjust the scoring power of a single player – the player taking the first shot in the frame (*Player 1*). A scaling factor above 1 therefore indicates that *Player* 1 has a greater scoring power than their opponent. The scoring potential of the two players is not adjusted.

The range of values chosen in this analysis are broadly consistent with levels recorded across the 46 matches analysed. Of the 20 players contesting at least 30 frames, the lowest aggregate success rate recorded was 82.5% (315 / 382) and the highest was 92.2% (1130 / 1226).

Table 4.11: Impact on the performance of Player 1 and the proportion of frames won from changing their probability of potting a ball during a scoring visit

	Player 1 Performance Measures			
Scaling Factor	**Scoring Power**	**Points per Scoring Visit**	**% Scoring Visits made**	**% Frames Won**
1.04	93.1%	41.6	49.9%	**60.1%**
1.02	91.0%	33.5	49.2%	**54.3%**
1.00	**88.8%**	**27.6**	**49.2%**	**48.9%**
0.98	86.7%	23.4	48.8%	**43.6%**
0.96	84.7%	20.4	48.6%	**39.5%**
0.94	82.7%	18.0	48.3%	**34.9%**
0.92	80.7%	16.0	48.5%	**31.4%**
0.90	78.6%	14.4	48.3%	**27.8%**

Table 4.11 shows that the proportion of scoring visits made by *Player 1* is largely unaffected by the scaling factors applied but it is the number of points scored at each visit which drives the changes in the proportion of frames they are modelled to win.

Note that as *Player 1* is modelled as playing the first shot in every frame they have slightly less than a 50% chance of winning a 'fair' simulated frame (i.e. when using a scaling factor of 1.00). In reality, the player taking the break-off shot would alternate, so this would even out over a longer match.

On the basis that players with a scoring power of around 83% and above are of a professional standard (Table 4.1), the results from these simulations are consistent with observed win percentages over the last two seasons (Collingwood et al, 2022). On average, a middle-ranking professional would be expected to win around 40% of frames against a Top 16 player, while a lower-ranked professional might win 33%.

4.4.3 Changing the scoring potential of one player

The third element of this analysis considers how the chances of a player winning the frame are affected by the strength of their tactical play relative to their opponent's.

The measure of scoring potential is not as clear an indicator of an individual player's level of performance as it is partly dependent on the performance of their opponent as well as their style of play (for example, whether they are more inclined to play safe than take on a difficult pot). Nevertheless, if one player is taking more opportunities to score than their opponent then we would expect them to gain an advantage.

Under each scenario, the inputs relating to the probability of potting a ball at the start of a *new* scoring visit (reflecting a player's scoring potential) are adjusted for *Player 1* using a scaling factor. A scaling factor above 1 therefore indicates that *Player 1* has a relatively strong safety game compared with their opponent and is more likely to pot the next ball following any safety exchange. The scoring power of the two players is not adjusted.

The range of values chosen in this analysis are broadly consistent with levels recorded across the 46 matches analysed. Of the 20 players contesting at least 30 frames, the lowest aggregate success rate recorded was 20.3% (47 / 232) and the highest was 37.6% (65 / 173). This latter figure was a bit of an outlier, with the next highest at 31.4% (157 / 500).

Table 4.12 shows that against weaker opponents, *Player 1* is more likely to pot a ball on a new visit so makes a slightly higher proportion of the scoring visits made during the simulated frames. This in turn increases their chances of winning the frame. Against stronger opponents, *Player 1* has fewer opportunities to score and wins a lower proportion of frames as a result.

Differences in the relative scoring potential of the two players would appear to have less of an influence on the frame outcome than their relative scoring power. This is not unexpected given that in the matches analysed 69% of shots played were part of a continuing visit as opposed to the start of a new visit.

Table 4.12: Impact on the performance of Player 1 and the proportion of frames won from changing their opponent's probability of potting a ball on a new visit

	Player 1 Performance Measures			
Scaling Factor	**Scoring Potential**	**Points per Scoring Visit**	**% Scoring Visits made**	**% Frames Won**
1.20	32.6%	27.4	53.4%	**54.4%**
1.15	31.2%	27.4	52.5%	**53.2%**
1.10	29.7%	27.2	51.3%	**51.9%**
1.05	28.4%	27.4	50.2%	**50.3%**
1.00	**27.0%**	**27.6**	**49.2%**	**48.9%**
0.95	25.6%	27.7	47.8%	**46.9%**
0.90	24.3%	27.7	46.7%	**45.3%**
0.85	22.9%	27.7	45.2%	**43.8%**
0.80	21.6%	27.8	43.7%	**42.0%**

4.5 Applications

There are a variety of potential applications for the model. Section 4.5.1 models the effectiveness of the break-off shot to explore what impact this has on the outcome of the frame.

Section 4.5.2 presents estimates of the probability of each player winning from different stages of the frame should they pot the next ball. This is used in Section 4.5.3 to evaluate alternative approaches available to a player. Firstly, where they have to decide whether or not to take on a risky pot. Secondly, where they are snookered (unable to hit the object ball directly) and must weigh up the risks incurred from either playing a straightforward escape which is likely to leave their opponent an easy opportunity to score; or attempting a more difficult escape which may see them concede penalty points.

4.5.1 Effectiveness of the break-off shot

The position of the balls for the first shot of each frame is the same, so in theory a player should be able to reach a high level of consistency with the shot. A top professional would certainly expect the cue ball to finish close to the baulk cushion – ideally in a way in which the baulk colours were blocking the path to most of the reds. They have less control over the reds themselves - although the red they hit will usually finish safe, a red on the other side of the pack will be released and may finish in a position from which it can be potted.

I was surprised to find that a ball was potted on as many as 20% of occasions immediately following the break-off. Subsequently, the player who broke-off (played the first shot in the frame) was observed to pot the first ball less frequently than their opponent. A one-sided binomial test of the observed outcomes (*N* = 734, *K* = 346) indicates that there is not enough evidence to conclude that this is significantly less than random chance (p = .07), but it would be interesting to test this with a larger sample of data. Over 60% of the observed frames were won by the player making the first scoring visit so I was still interested in analysing how much the effectiveness of the break-off shot might affect the outcome of a frame.

To do this I changed the probability of potting a ball on Shot 2 (default = 0.200) to reflect differences in the quality of *Player 1*'s break-off shot. The better the shot, the lower the probability that *Player 2* will pot a red on their turn. The probabilities associated with subsequent shots were not adjusted. The outcomes compared are the proportion of frames in which *Player 1* subsequently pots the first ball, and the proportion of frames they go on to win, shown in Table 4.13.

Table 4.13: The impact on the outcome of a frame from changing the probability that Player 2 pots a red on their first shot

Probability that *Player 2* pots a red on their first shot	**Proportion of frames where *Player 1* pots 1st red**	**Proportion of frames won by *Player 1***
0.00	56.4%	51.7%
0.05	54.2%	51.3%
0.10	51.2%	50.3%
0.15	49.0%	49.9%
0.20	**45.5%**	**48.9%**
0.25	43.2%	48.1%
0.30	40.5%	47.8%

The results show a direct impact on the proportion of frames where *Player 1* pots the first red. Under the default inputs, *Player 2* has a 20% chance of potting a ball on their first shot, making them more likely to pot the first ball in the frame. If *Player 1* could reduce *Player 2*'s chances to around 10% then my modelling indicates that the advantage switches and *Player 1* is now slightly more likely to pot the first ball. The impact on the ultimate outcome of the frames is relatively low, although in elite sport competitors are looking for such marginal gains which may give them a crucial advantage over their opponent.

It should be noted that this analysis is based on *Player 1* increasing the effectiveness of their break-off shot by achieving a higher level of consistency, either though practice and / or a small modification to the shot attempted. A significant change in approach (e.g. rolling up to the reds) may not have the same effect as the shot faced by *Player 2* would also be very different, potentially making it easier for them to restrict their opponent's opportunity to score in return.

More comprehensive data would be required to determine whether some players are consistently able to limit their opponent's chance following the break-off shot; the sample sizes available are not substantial enough. For interest though, of players breaking-off in at least 20 frames during the recorded matches I observed the following extremes:

- Neil Robertson's opponents were only able to pot a ball on the following shot once out of 22 frames (5%) and Robertson made the first pot on 13 occasions (59%) – although despite this, he only went on to win 9 of the 22 frames when he broke-off (41%).
- Conversely, Kyren Wilson's opponents successfully potted a ball on their first shot in 13 out of 44 frames (30%). As a consequence, Wilson potted

the first ball in just 19 of the frames in which he broke-off (43%) - although he did go on to win exactly 50% of them.

4.5.2 Impact of making the next scoring visit on the outcome of a frame

Section 3.3 compared the observed and modelled outcomes of frames based on the difference in the score and the number of reds remaining when each scoring visit started. In this section, that analysis is extended to focus on specific scenarios and simulate the outcomes of frames at each point.

Figure 4.8 shows the probability of winning at different instants during a frame based on the number of balls remaining at the start of the visit (with each estimate based on a separate simulation of 10,000 frames). A range of scenarios have been run based on the player being 30 points ahead; 15 points ahead; scores level; 15 points behind and 30 points behind.

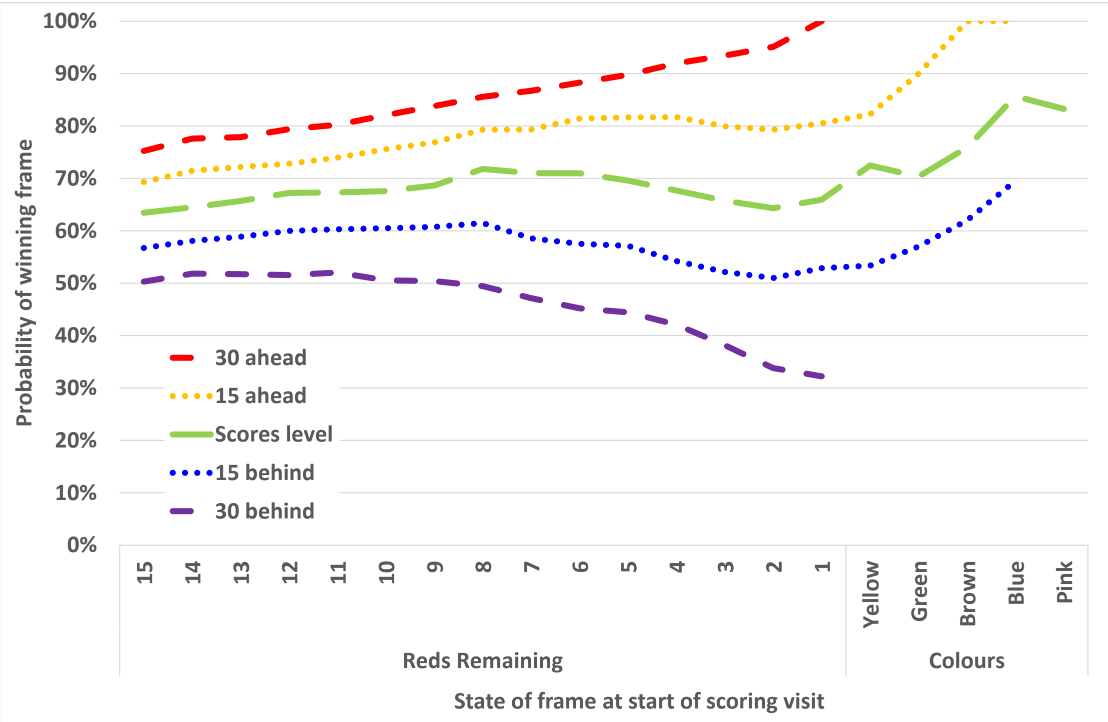


Figure 4.8: Proportion of frames won by the player who makes the next scoring visit at different stages of the frame given the score at the start of the visit

Note that the simulation was not run in cases where snookers are already required (i.e. where a player is 30 points ahead on the colours, or 15 points ahead with only pink and black remaining). Scenarios are included where a player is ahead or behind at the start of the frame, although these would only occur if a player had conceded penalty points before a ball was potted.

There is, unsurprisingly, no overlap between the lines produced; the more points a player leads by at any stage of the frame, the greater their chances of winning. With a player 30 points ahead or behind the lines are relatively smooth; the further the frame has progressed, the more likely it is that the player leading will go on to win the frame. As noted in Section 3.3, even if a player is 30 points behind, they still stand a good chance of winning the frame if they can take the next opportunity to score.

Where the scores are closer, the probability that the player potting the next ball goes on to win the frame is slightly higher while there are a number of reds left on the table presenting a good opportunity to make a decisive break. It then falls slightly for the last few reds, which may not be as well positioned.

If the frame is still close when it reaches the colours, then the next player to pot a ball will be favourite to win the frame. The number of pots required to win becomes a factor in determining how likely it is that each player will win – resulting in the non-monotonicity of the probabilities presented in Figure 4.8. For example, with scores level on the blue the player who pots the blue just requires either the pink or black, while their opponent needs both. With scores level on the pink, the player potting the pink still requires the black. If they miss, then their opponent can also win by potting the black.

4.5.3 Evaluating shot choice

There are a number of variables which we cannot fully model, such as the position of the remaining balls, but we can make some general assumptions to help devise appropriate strategies for different situations faced within a frame of snooker.

Risky Pot or Play Safe?

The first scenario considered is where *Player 1* is at the table at the start of a new visit and has the choice of playing safe or taking on a pot.

If they play safe, their probability of winning the frame can be written as:

$$P(Win \mid Safe) = SV_1 + (1 - S)(1 - V_2) \tag{2}$$

where S is the probability that *Player 1* pots the next ball after the initial safety shot and V_i is the probability that *Player* i wins the frame given that they pot the next ball, which is estimated using the simulation model. Where $V_1 \neq V_2$ this would typically imply that either one player has currently scored more points than the other and / or there is a difference in the relative scoring power of each player. Towards the start of the frame, it could also imply that the safety play of one player is stronger than the other, enabling that player to make a higher proportion of the scoring visits.

If they take on the pot, then with probability ***x*** a ball is potted with a 'fair' chance of continuing the visit (i.e. consistent with the model inputs). Otherwise, *P* represents the probability that *Player 1* pots the next ball after their initial attempt at a pot is missed. Their probability of winning the frame having taken on the pot can therefore be written as:

$$P(Win \mid Pot) = \boldsymbol{x}V_1 + (1-\boldsymbol{x})[PV_1 + \; (1-P)(1-V_2)] \qquad (3)$$

In order to justify taking on the pot, we seek criteria for $\boldsymbol{x}$, which for estimated values of V_i, *S* and *P* would satisfy $P(Win \mid Pot) \geq P(Win \mid Safe)$.

Equation (2) can be rearranged as:

$$P(Win \mid Safe) = S(V_1 + \; V_2 - 1) + (1 - \; V_2)$$

and similarly, equation (3) can be rearranged as:

$$P(Win \mid Pot) = \boldsymbol{x}(V_1 + V_2 - 1) \; + (1-\boldsymbol{x})P(V_1 + \; V_2 - 1) + (1 - \; V_2)$$

Cancelling the term $(1 - \; V_2)$ and dividing throughout by $(V_1 + V_2 - 1)$ this leaves us with the following criterion for choosing the pot over the safety shot:

$$\boldsymbol{x} \; \geq (S-P)(1-P)^{-1} \qquad (4)$$

It is notable that this criterion is only dependent on *S* and *P* and not on V_i. Under this model of a frame, a player's strategy should only be affected by how likely they are to pot the next ball and not by the current score in the frame or the scoring power of their opponent.

In the scenario where two players of equal ability are level on points with 5 reds (67 points) remaining, we can use the simulation outputs displayed in Figure 4.8 to estimate that the next player to pot a ball has a 70% chance of winning the frame (i.e. $V_1 = V_2 = 0.7$).

From equation (2) we see that the probability of *Player 1* winning the frame if they choose to play safe is therefore 0.4*S* + 0.3. If *S* = 0.5 then this intuitively means that $P(Win \mid Safe) = 0.5$.

To resolve equation (3) we consider how the probability of winning the frame given different estimates of ***x*** varies according to the consequences of the shot being missed, represented by *P*:

- *P* = 0 reflects the riskiest shot; if the pot is missed then their opponent is certain to pot the next ball. The probability of *Player 1* winning the frame from this opportunity is 0.4***x*** + 0.3.
- *P* = 0.25 reflects a shot with a 50% chance of leaving the balls safe – from which *Player 1* has a 50% chance of potting the next ball. Their probability of winning the frame is now 0.3***x*** + 0.4.

The horizontal (red) lines in Figure 4.9 show $P(Win \mid Safe)$ for different values of *S*, while the diagonal (black) lines show $P(Win \mid Pot)$ for different values of *P* and ***x***.

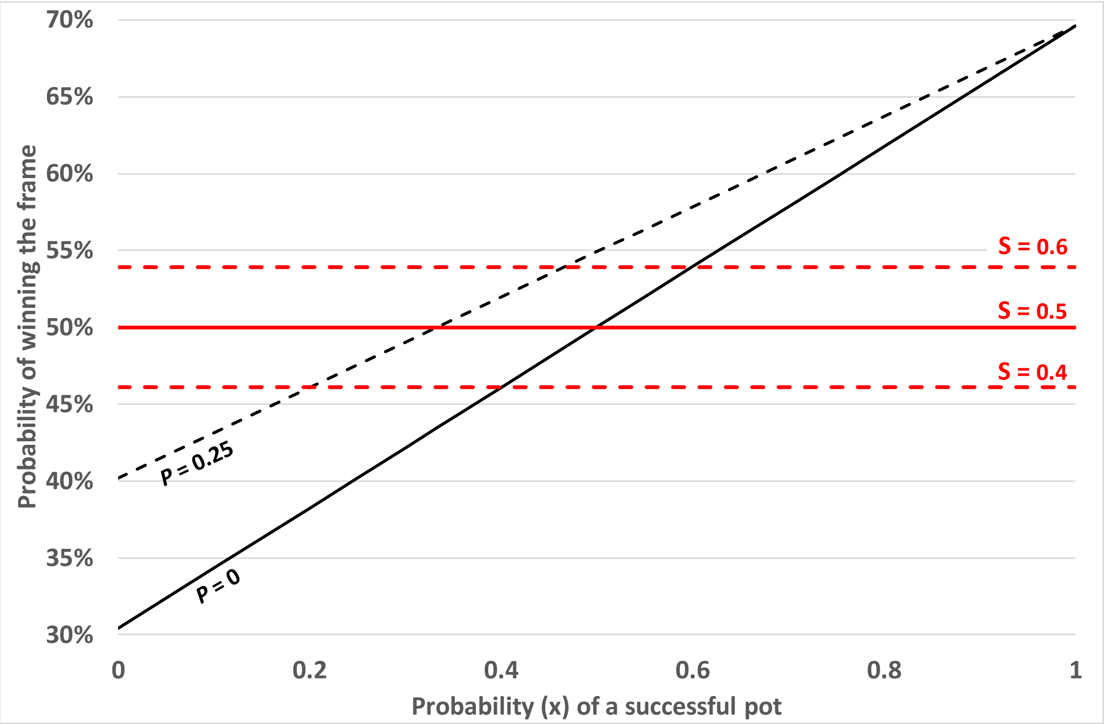


Figure 4.9: Probability of winning the frame with the scores level and 5 reds (67 points) remaining when a) playing safe, or b) taking on a pot with different expectations of success at selected levels of risk

This can be used to identify criteria for ***x*** which would justify taking on the pot for different estimates of *S* and *P*, shown in Table 4.14. These criteria also hold in scenarios where there is a difference in the current score or in the scoring power of the two players.

Table 4.14: The expected success rate for a particular pot which is required in order to justify taking on the pot rather than playing safe, for specified estimates of the probability of potting the next ball after playing safe (*S*) and the probability of potting the next ball if the pot is taken on but missed (*P*).

	P = 0	**P = 0.25**
S = 0.25	***x*** ≥ 0.25	***x*** ≥ 0.00
S = 0.40	***x*** ≥ 0.40	***x*** ≥ 0.20
S = 0.50	***x*** ≥ 0.50	***x*** ≥ 0.33
S = 0.60	***x*** ≥ 0.60	***x*** ≥ 0.47

In a match between two top professionals, we would anticipate that the neutral assumption of *S* = 0.5 would be reasonable in most situations. A general rule that follows from this is that a player should always take on the pot if they would expect to be successful at least 50% of the time – even if they would leave a certain scoring opportunity for their opponent if they missed.

Where the player believes that they can play the pot with a degree of safety (e.g. *P* = 0.25), it would be reasonable for them to be more aggressive and take on a more difficult pot (***x*** ≥ 0.33) as a missed pot would not always be costly.

Where the player thinks they can gain a genuine advantage in the safety exchange (e.g. *S* = 0.6), then they may benefit from waiting for a better opportunity (***x*** ≥ 0.60). Conversely, where the player has no straightforward safety shot (e.g. *S* = 0.25), a difficult pot (***x*** ≥ 0.25) may be the best option.

Snookered

When a player (*Player 1*) finds themselves *snookered* (i.e. unable to hit a red directly) with multiple reds left on the table they could take one of two alternative approaches to the shot.

One approach (sometimes preferred by Ronnie O'Sullivan) would be to maximise their chances of hitting a red by playing at pace towards a group of reds, hoping to get lucky and not leave an easy chance for their opponent (a "hit and hope"). The likely outcome is that they would not give away any penalty points but would hand their opponent a good chance of potting a ball.

The alternative approach (more commonly taken by professional players) is to attempt a more difficult escape which minimises the chances of leaving their opponent an easy pot, perhaps by rolling up to a solitary red near a cushion. The risk from this approach is that they are more likely to miss the red and concede penalty points, with their opponent also having the option to ask them to replay the shot. Successive misses can occur in such a situation, with the player giving away numerous penalty points before finally hitting the red.

Equation (2) from the previous scenario can again be used to estimate the probability of *Player 1* winning the frame if they follow each of the potential approaches. In this scenario, both players are modelled as being of equal ability and the scores are level before the shot is taken.

For the first approach, it is assumed that no penalty points are conceded by *Player 1* (so $V_1 = V_2$) but that their chances of potting the next ball are now very low – the cases where *S* = 0 (i.e. a certain pot for their opponent), and *S* = 0.125 (i.e. a 25% chance of leaving the initial shot safe, with a 50% chance of winning the resultant safety exchange) are considered. Using the simulation outputs displayed in Figure 4.8, if there are currently 10 reds remaining the player who pots the next ball is estimated to have a 68% chance of winning the frame (i.e. $V_1 = V_2 = 0.68$). *Player 1*'s chances of winning the frame therefore stand at 32% (where *S* = 0) or 37% (where *S* = 0.125).

For the alternative approach, it is assumed that once the player hits the red, they do not leave a pot for their opponent and that there is an even chance of each player winning the subsequent safety exchange (i.e. *S* = 0.5). In this case, equation (2) simplifies to $0.5 + V_1 - V_2$. If they hit the red at the first attempt, $V_1 = V_2$ and their chances of winning the frame are 50%. Otherwise, if they concede penalty points before hitting the red, $V_2 > V_1$ and their chances of winning the frame are reduced.

With 10 reds remaining, the analysis in Section 4.2 estimates that their chances of winning would fall to 42% if they concede 15 penalty points before hitting the red, while if they were to concede 30 penalty points their chances fall to 34%. [Missing the reds on a single shot would typically concede 4 penalty points, so the examples given are very roughly comparable to missing the red on 4 occasions, and on 7-8 occasions.]

Figure 4.10 plots the potential outcomes from playing a cautious escape (in red) and from playing a riskier 'hit and hope' (in black) at different stages of the frame. During the first half of the frame (i.e. with at least 5 reds remaining), conceding 15 points in penalty points but not leaving a chance is always preferable to handing the opponent a scoring opportunity. It is rare that a professional will miss more than 3 or 4 times in succession, so the more cautious approach (assuming that the player will succeed in keeping the balls safe) is generally advised. This is also the case where the scores are not level – subject to the player being able to concede points without requiring snookers themselves.

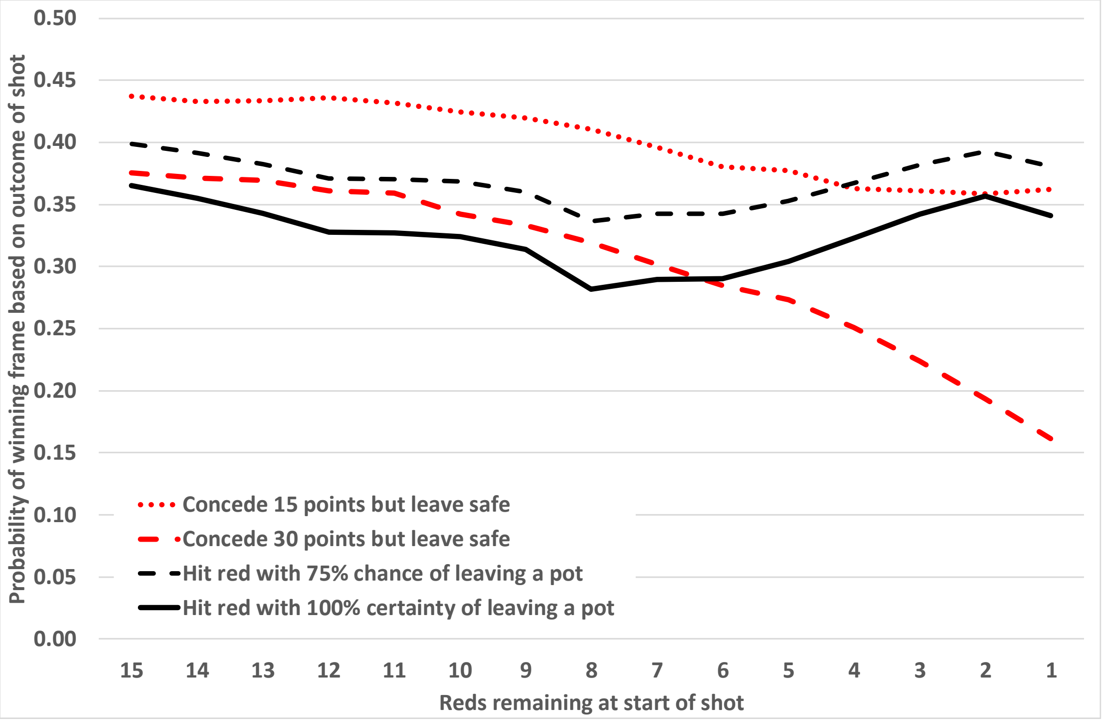


Figure 4.10: Probability of winning the frame depending on the outcome of the shot played when snookered and the number of points conceded in the process

As the frame progresses, it can be concluded that the difficulty of the escape should be considered; conceding 15 penalty points becomes more costly than allowing the opponent to pot the next ball – on the basis that the last couple of reds may be difficult to gain position on, preventing the opponent from making a frame-winning break. A 'hit-and-hope' shot may also stand a greater chance of leaving the balls safe when there are only a couple of reds left on the table.

4.6 Conclusions

This chapter described the development of a simulation model which accurately reflects the progression of a typical frame of snooker played between two top professionals. The analysis is based on over 30,000 shots and indicates how the probability of potting a ball on each shot changes as the frame progresses.

The number of points scored during each scoring visit closely matches the observed data and the scoring power of a player within the simulation can be adjusted to reflect players with slightly greater / weaker break-building prowess. This scaling factor could be used to reflect the scoring power of players at a lower level of the game as well, although further data would be required to validate this.

The safety prowess and tactical awareness of players is harder to measure, although the inputs to the model can be refined to reflect how likely it is that a player will leave an opportunity for their opponent. This has a direct impact on the proportion of scoring visits made by each player (their scoring potential), which potentially gives one player a crucial advantage over another. It would also be important to understand if there is also an effect on scoring power, i.e. whether the quality of the scoring opportunities is affected as well as the quantity. This may potentially affect both players if longer safety exchanges tend to leave more reds close to the cushion / colours away from their spots.

In theory it would be possible to determine suitable scaling factors to represent the abilities of each individual professional snooker player, allowing for more detailed analysis of how likely one player is to defeat another, and which aspect of their game provides them with this edge. A lot of additional data would be required, but this could be generated from the automated scoring system, rather than requiring a significant amount of manual collection. A potential approach to estimating a player's scoring potential and scoring power based on available data is discussed in Section 5.4.4.

Developing a method of simulating the frames played by individual players would allow for more formal analysis of which factors differentiate the top players from the rest and how this translates to them winning more matches.

Examples of how the model could be used to help analyse particular situations which occur during a frame were discussed.

- The effectiveness of the break-off shot has a limited impact on the outcome of a frame, but this could nevertheless give one player an advantage at a crucial stage of match.
- As a general rule, a player is advised to take on a risky pot if they would expect to execute it successfully more often than not – even if missing it would leave a certain chance for their opponent. If there is an element of safety attached to the shot then they could afford to attempt a more difficult pot; whereas if they felt they could gain an advantage in the safety exchange then they would be advised to wait for a better opportunity.
- When snookered with multiple reds still on the table, a player is generally advised to attempt a more cautious escape which minimises the risk of leaving their opponent a chance - even if they concede penalty points in the process.

The model could be developed to explore other situations too, such as whether a player who is ahead would benefit from pushing a colour to a safer position; reducing the risk of their opponent clearing the table to steal the frame.

For this version of the model, I have chosen not to capture any shots which result in a foul. Towards the start of the frame, it would be straightforward to introduce these as a potential outcome from each shot; primarily occurring during a safety exchange. In the latter stages of a frame, particularly where one player requires snookers, additional assumptions would be required to reflect the changing objectives of each player and the types of shot played in these circumstances. While in theory the inclusion of fouls would make the model more realistic, I believe that there are limited benefits from doing so.

A more natural development would be to extend the model to simulate full matches rather than a single frame. A Bayesian updating rule could be introduced to reflect actual performance as the match develops, as developed by Kovalchik and Reid (2019) to model changes in the proportion of points won on serve as a tennis match progressed, and by Song and Shi (2020) to model the effect of in-play adjustments to team parameters on the results of basketball matches.

5 Conclusions and Further Work

This chapter summarises the findings and contributions from the three elements of research presented in this thesis: rating & ranking players (§5.1), measuring performance (§5.2) and modelling the progression of a snooker frame (§5.3). Areas of potential further work are explored in Section 5.4 and concluding remarks presented in Section 5.5.

5.1 Rating & Ranking Players

In any competitive environment we are interested in accurately ranking the participants in some way, thus identifying the strongest competitor and their closest rivals, while providing an assessment of where all other participants fit within the hierarchy. Within sport, rankings are commonly used to determine eligibility and generate seedings for tournaments. At the end of a season, they can also control promotion and relegation from a league or tour.

More than this, we ideally want to quantify the difference in ability between any two competitors, which enables us to translate this into an expectation of the likely outcome of each match. Aside from the potential interest for bettors and bookmakers, the accurate rating of players can also provide valuable insights to schedulers, spectators and players.

Tournament organisers can use the information to anticipate how long a match (or series of matches) may take to play, while different tournament designs can be modelled and compared. The chances of unexpected results can be estimated, providing perspective for those participating in and following the sport. Over a longer timeframe, anticipating the progression of a player's ranking can help them to manage their effort over a season.

There is no perfect way of ranking and rating players, with various methods and options available. Measuring performance over different time periods will yield different rankings, while results can be either be treated equally or weighted towards more recent, or more high-profile matches. A player's performance can be assessed purely by their progression in a tournament or can be scaled according to the level of opposition they faced. The merits of these alternatives need to be considered in determining a suitable rating system.

Chapter 2 set out four different methods for rating and ranking players with the objective of quantifying the relative ability of professional snooker players. The predictive ability of the models was compared using professional snooker results from the 2017/18 and 2018/19 seasons.

The merits of the models developed are discussed in Section 5.1.1, along with suggestions for how each of the models could be developed. The approach taken to comparing the predictive ability of the models is considered in Section 5.1.2, acknowledging the benefits of analysing subsets of matches to better understand the limitations of each model. Potential applications of the models are then discussed in Section 5.1.3.

5.1.1 Evaluation of the Models Developed

Four different models were described in Section 2.2. Logit models were created based firstly, on the difference in world ranking points earned by players over the last two years (§2.2.1), and secondly, the difference in the percentage of frames won (§2.2.2). Two paired comparison models were also created, using formulations devised by Bradley-Terry (§2.2.3) and Elo (§2.2.4).

The relatively large amount of prize money offered during the World Championship has a clear impact on the world ranking of the top players. Otherwise, the ranking of players was relatively similar between each of the models using at least 2 years of data (§2.3.1). The predictive ability of all four types of model chosen was also largely comparable (§2.3.2).

The model based on the world rankings was noticeably weaker at reflecting the ability of less experienced players (§2.4.1); a direct consequence of excluding the previous results of returning professionals and all amateur players. There is justification for using an accumulative system to determine the official rankings, but consequently there is a clear limitation in its use as a predictive model.

The model based on win percentages over the last 2 years produced reliable predictions, on a par with the Bradley-Terry and Elo models. A potential limitation is that it does not account for the strength of opposition faced by players. In particular, the highest-ranked players progress through to the latter stages of competitions more frequently and face a higher level of opponent as a result. This is discussed further in Section 5.4.1.

Win percentages offer a further advantage in that they are relatively straightforward for non-analysts to understand. This research indicates that they can play a meaningful role in communicating the relative performance of players to a wider audience.

The Bradley-Terry model showed better discrimination than the Win Percentage model and for the highest-ranked players (who play the most matches) there was some evidence that it produced better predictions. This did not always appear to be the case for lower-ranked players; it was noticeable that the overall results were no better than those of the Win Percentage model (§2.4.2).

Utilising two years of historical results generally improved the predictive ability of the models. There may be a case for introducing a factor which reflects a player's recent form more heavily, especially where there has been an improvement in a player's performance over the last year (§2.4.3). More generally, there would be an argument for trying to capture the variation in a player's performance. This is discussed further in Section 5.4.2.

It would be interesting to explore whether the Bradley-Terry model would benefit from the inclusion of additional years of data, particularly for those players who compete less frequently. More recent results could be weighted more heavily to mitigate any change in performance levels over time.

One advantage of the Bradley-Terry and Elo models is that data from amateur tournaments could easily be added, with the different level of opposition faced automatically accounted for in the generation of the relative ratings. This would increase the amount of information available for players who have just joined or returned to the professional tour.

The Elo model did appear to benefit from the use of a larger set of historical data, with predictions for less experienced players exhibiting lower levels of bias (§2.4.1). A useful development to the model would be to vary the weight allocated to the most recent results depending on the experience of the player. The ability of higher-ranked players is likely to be more stable, justifying the use of a lower weight when updating their ratings. This is also considered in Section 5.4.2.

For this research, the models have been updated around 10 times a season to mirror the way in which the official World Rankings are used to determine the seedings for each tournament. This means that a player's rating will not include results from the most recent tournament(s) played if an update is not applied until a later date. Taking this work forward, it is anticipated that the models would be updated more frequently so that the ratings for each player at the start of each tournament (or stage of a tournament) would include all results from matches played over the last two years.

5.1.2 Evaluation Methods

A variety of methods have been used in previous academic literature to compare the results of predictive models. The approach used in this research has followed that taken by Kovalchik (2016), with a range of metrics used to assess the levels of accuracy and bias in the predictions made.

Kovalchik derived her calibration ratio from the expected and actual wins for the tennis player with the higher World Ranking but I have preferred to base mine on the higher rated player according to the model being analysed. The calibration of each model is therefore assessed on its own basis rather than also being influenced by the strength of alignment between the model and the World Rankings.

In attempting to replicate her results I identified an error in her application of the Bradley-Terry model, which meant that her predictions were reflecting the probability of a player winning a single game, rather than the whole match. This would explain the relatively poor calibration and log-loss scores achieved by the model in her analysis.

In presenting the results of this analysis in Section 2.4, I identified specific subsets of matches which are relevant to understanding differences in the models and in comparing their performance. This is rarely covered in academic papers but helps to highlight the strengths and limitations of each model and provide insights into how they could be developed. This analysis indicates that the effects of form and the relative strength of opposition previously faced are not as clear-cut as anticipated and that further analysis would be required to improve the reliability of the models.

5.1.3 Potential Applications

The intention of this element of the research was to develop a reliable model for anticipating the outcome of a snooker match between any two professional players, which could be used as the basis for further analysis of the game.

A natural application is to use the models created to simulate the progression of tournaments and potentially even the entirety of a professional season. We can also analyse the progression of a match to test the assumption of independence between the outcomes of individual frames.

Modelling Tournaments

Section 2.5 provided an example of how the expected outcomes of a tournament could be modelled and how alternative tournament designs could be evaluated based on their expected impact on different players.

The change to the qualifying rounds of the World Championship is not the only variation in tournament design which has recently been implemented or proposed.

- The Saudi Arabia Snooker Masters was due to be introduced during the 2020/21 season but was cancelled due to the Covid-19 pandemic. The proposed tournament structure would have seen players entering in different rounds based on their seeding; players seeded 65+ would enter in Round 1, players seeded 33-64 would enter in Round 2, while players seeded 1-32 would receive byes through to Round 3 (World Snooker Tour, 2019).
- Two additional tournaments were played during the 2020/21 season which featured a series of round-robin group stages rather than a traditional knock-out tournament. The Championship League saw participants split into groups of 4, with the winner progressing to the next group stage. The World Snooker Tour (WST) Pro Series saw participants split into groups of 8, with the top two qualifying for the next group stage. The former even allowed for the possibility of drawn matches as players effectively contested a "Best of 4 frames" match, with possible outcomes of 3-0, 3-1 or 2-2.

A notable feature of the round-robin design is that players are guaranteed to play multiple matches (3 in the Championship League, 7 in the WST Pro Series), whereas half of entrants will only play a solitary match in a traditional knock-out tournament.

The variation in the lengths of professional snooker matches was also discussed in Section 1.2.2, with Appendix V showing how the length of the match influences the expected outcome of a match contested by players of different abilities. The effect that this has on the outcomes of tournaments can also be analysed in this way; for example, comparing the chances of each individual winning if matches are the “Best of 7 frames” or the “Best of 19 frames”.

The way that the seeding of players within a tournament influences its progression could also be analysed. The top 16 seeds are usually separated from each other in the draw so that they won’t play each other until the 4th round (the “Last 16”). The British Open (reintroduced in 2021) is one exception to this, with the draw made completely at random.

On the other hand, the UK Championship draw is completely determined by the seedings. The 1st round (the “Last 128”) is designed so that the #1 seed plays the #128 seed, and the #2 seed plays the #127 seed etc. Subsequent rounds follow a similar pattern, so in the 2nd round (the “Last 64”) the #1 seed would potentially play the #64 seed and the #2 seed plays the #63 seed etc.

The efficacy of this was raised during the 2021 UK Championship, after Shaun Murphy (seeded #6) was knocked out in the 1st round by an amateur, Si Jiahui (effectively seeded #123). Murphy controversially argued that “amateur players should not be permitted to play in professional tournaments” (Hafez, 2021), with Jiahui one of 8 amateurs invited to enter the tournament to ensure a full complement of 128 players.

One of the points raised was that Jiahui, who had been a professional for the previous two seasons, was arguably a stronger opponent than some of the professionals seeded above him who had only recently joined the Tour. There is indeed some merit in this argument as he won 47% of his frames during the 2020/21 season, giving him the 75th highest win percentage. This research has shown more generally that the world rankings are limited in their ability to accurately rank amateur players and new professionals – especially early in the season when the 2021 UK Championship was contested. It is, therefore, very likely that some players will end up with easier (or more difficult) 1st round opponents compared with the intentions of the tournament design.

Modelling Seasons

A natural extension from modelling a single tournament would be to model an entire season of professional events. A key area of interest would consider the speed at which players might progress up (and down) the rankings and how likely it is that a player joining the Professional Tour will break into the Top 64 within two years (and secure their place for another year). What results would they need to achieve this and what level of performance would be required?

Match Progression

The models described in Chapter 2 were all based on the outcomes of frames played, with Figure 2.1 showing a strong correlation between the proportion of frames won and the proportion of matches won by each player. The models created assume that the outcome of frames played are independent of one another and provide us with a baseline expectation of how a match is likely to progress. Comparing this against actual results will help to assess how reliable an assumption this is, or whether our expectations should change as the match progresses based on the outcome of previous frames. This is discussed in more detail in Section 5.4.4.

5.2 Measuring Performance

Statistics provide a condensed summary of a match, giving the audience an understanding of how the two players are performing and how this has influenced the current score. Some statistics are ultimately just intended to provide a high-level picture of the match (such as the current score), while others are presented with the purpose of providing a meaningful indicator of the performance of each player.

Valid and reliable performance indicators provide valuable information for players and their coaches, whether for comparing the statistics of different players or evaluating how an individual's level of play has progressed over time. More generally they can help to identify the relative strengths of players and highlight where improvements could be made.

Chapter 3 set out alternative measures of performance which could be used in snooker instead of the existing statistics that are produced. The merits of adopting summary statistics relating to the number of scoring visits made by each player are discussed in Section 5.2.1. Dynamic measures which focus on how play transitions from one shot to the next are described in Section 5.2.2.

A key issue in assessing the performance of individual players is the limited amount of publicly available data. The potential for developing basic statistics using the official scoring system is examined in Section 5.2.3, along with a discussion of the type of analysis which could be possible using technology available in other sports.

5.2.1 Summary Measures and Scoring Visits

My analysis concluded that there is limited value in the existing measures of *Total Points Scored*, *Balls Potted* and *Highest Break* (§3.3.1). Presenting the average points scored per frame would at least allow for some level of comparison across different matches but this still does not provide a meaningful indication of a player's performance. Reporting the number of 100+ and 50+ breaks made during the match would also be more informative than the highest break recorded, albeit making relatively limited use of the available information.

Snooker frames are made up of a series of *visits*, with a player's points largely accrued through a subset of visits in which a ball is potted (*scoring visits*). Information relating to the scoring visits made during a match therefore provides valuable insight into how it has progressed and as a basis for analysing differences in the performance of both players[1]. Two alternative measures were set out in Section 3.4.1:

The ***number of scoring visits*** made by each player provides an indication of the number of scoring opportunities created by each player. To enable a comparison across matches this can be presented as the average number of scoring visits per frame.

The ***average number of points*** accumulated per scoring visit then provides an indication of how productive these visits have been. In conjunction, these summary scoring measures help to provide an indication of where one player may have gained an advantage over the other.

[1] A small proportion of points (3.4% in the 46 matches observed) are accrued through penalty points, but these have not been analysed within this research.

5.2.2 Performance Indicators and Dynamic Measures

The existing measures of player performance are based on their success at executing individual shots – whether attempted pots or safeties. They are subjective measures, which require someone to assess whether a pot or safety shot was being attempted. An additional distinction is made for shots from long distance and those played using the rest, while the success of a safety shot is also based on the observer's judgement. There are no official definitions for each type of shot which make the statistics difficult to replicate, whilst providing scope for potential inconsistencies in the production of the official statistics.

At the highest level, the *Pot Success* rate does to a large extent reflect the relative performance of each player (§3.3.2), although it does not account for the difficulty of the pots attempted or the consequences from making or missing the pots. As a result, it is questionable whether the measure reliably reflects the progression of the match between players of a lower level of ability, or when relatively few frames are played. In these cases, the link between *Pot Success* and match outcomes is less apparent.

An example of this is the quarter-final between Jimmy White and Peter Lines during the 2020 World Seniors Championship. Lines officially recorded an impressive *Pot Success* rate of 90.9% (50/55), compared with 87.6% for White (85/97). His recorded *Safety Success* rate was also higher (87.3% v 83.7%). This wasn't an accurate reflection of the match though - Lines won the first frame but White dominated the remaining frames and won 4-1. Although Lines missed just 5 pots during the match he frequently ran out of position and could not make the most of his opportunities to score.

Analysis of the scoring visits made by each player provides a much more accurate representation of the match. Lines made an average of 10 points from his 15 scoring visits but White comfortably outscored his opponent with an average of 27 points from his 10 scoring visits.

The measures of *Long Pot Success* and *Rest Success* focus on particular types of shot but these occur relatively infrequently and provide limited information about the way a match has progressed. The exact definition of a long pot is unclear and will reflect a wide variety of attempted pots, some of which will be far more difficult than others. It is interesting to compare the relative success which different players have in using the rest, although it would be more informative to develop the measure to incorporate alternative methods of playing shots beyond a player's normal reach.

A measure of a player's safety play is clearly valuable, but the existing measure of *Safety Success* is very subjective and has limited value in explaining who gained the advantage in key safety exchanges and how this translated to the number points scored by each player. Further work is required to scope out potential measures of safety play and this is discussed in Section 5.4.3.

In a sport such as snooker, where the outcome of one shot has a very clear impact on the subsequent one, shots cannot purely be evaluated in isolation. A successful pot is more valuable if it is followed by another, while the success of a safety shot is ultimately determined by which player pots the next ball. The dynamic nature of a frame should be captured within measures of performance.

This research proposes two core measures which evaluate a player's *Scoring Potential* and *Scoring Power*. The merits of these measures were discussed in the context of measuring a player's performance (§3.4.2) and also in modelling the progression of a frame (§4.4).

A player's ***Scoring Potential*** represents the proportion of visits made in which a ball is potted and reflects the overall probability that a player will pot a ball given that they are starting a new visit. This is influenced by both players so is a less effective indicator of an individual's performance, although the difference in performance between the two players was found to be a factor in determining the outcome of a match.

A player's ***Scoring Power*** represents the proportion of successful pots which are followed by another and reflects the overall probability that a player will pot a ball given that they are continuing a visit. There is evidence that this is a key indicator of the ability of a snooker player, and not just at the professional level. The relative performance of the two players during a match was shown to be a factor in determining its outcome.

These measures are just a starting point, but they do provide a sound basis for further analysis into how scoring visits start and progress and where some players gain an advantage over others. They are both objective measures, which enables them to be derived directly from information generated by the scoring system and greatly increases the number of matches which they could be produced for. There would still be merit in producing more subjective measures to enable a more precise assessment of a player's performance, but there is no need for such statistics to be the only ones produced.

5.2.3 Data Expansion

The limited amount of within-match data currently available is the primary reason why snooker lags far behind other sports in the way data is used to analyse player performance. Much of the analysis within this thesis is based on a bespoke data collection from 46 matches using post-match video analysis, a tiny fraction of the 2,000+ professional matches contested each year. This is in sharp contrast to the way that data is collected and used in other sports, as described in Section 1.4.3.

All professional snooker matches use an automated scoring system which records the player taking each shot and the number of points scored. The data from this could be used to produce the statistics described in this thesis, both for all individual matches and at an aggregate level for each professional player. This would offer far greater scope for understanding where matches are won and lost, how performance differs between players and how much an individual player's performance varies over all matches played.

A larger pool of data would also enable additional statistics to be generated:

- It would be interesting to determine the proportion of frames in which a player potted the first ball, and how many times they (or their opponent) potted a ball on the shot following the break-off.
- The average number of points scored during each scoring visit could also be calculated for different stages of the frames – based on the number of reds remaining on the table at the start of the visit.

- A player's propensity to win frames from different situations could be measured, comparing actual outcomes with an expected outcome given the current score in the frame.
- Further analysis could be carried out to develop meaningful measures of the more tactical side of the game. This is discussed in Section 5.4.3.

The use of ball tracking, in a similar way to golf, would generate a massive leap forward in the type of performance analysis that could be carried out. The geometry relating to each pot could be captured (the distances and angles between the cue ball and object ball, and between the object ball and the pocket) and based on historical data an expectation of the pot success could be generated. A player's success in executing different types of shot could then be compared against a particular standard of player. An extension of this approach would seek to estimate the expected points scored during a player's visit based on the position of all balls on the table.

Some of this would take time and money to set up but the technology exists and is used extensively in other sports, so it will hopefully be introduced into snooker in the not-too-distant future. In the meantime, analysis of data from the scoring system would enable a core set of indicators to be developed, providing the groundwork for further exploration.

5.3 Modelling the Progression of a Snooker Frame

Chapter 3 considered different measures of performance and their relative importance in determining the outcomes of matches. To analyse this in more detail we need to explore how a frame progresses depending on the ability of the two players.

Chapter 4 detailed the development of a Monte Carlo simulation model of the progression of a frame. Simulation is a tool used to create a representation of a system to enable further exploration and enhance understanding. In modelling a frame of snooker, we are particularly interested in understanding how the probability of each player winning the frame evolves as the frame progresses.

This is clearly of interest to the betting industry and anyone wishing to place a bet "in play", but also provides context to the general viewer as well. Perhaps more significantly, it provides potential benefits for players and coaches in informing strategies used by players. Should someone play more cautiously to defend a lead, or be more inclined to take a risk if they find themselves behind? How is this affected by the strength or playing style of their opponent?

An evaluation of the model created is summarised in Section 5.3.1, with the ability of the model to estimate expected outcomes at different stages of the frame explored in Section 5.3.2. The potential for using the model to aid decision making is discussed in Section 5.3.3.

5.3.1 Model Evaluation

Chapter 4 extended the analysis behind Chapter 3 to model the progression of a frame, using a series of inputs to reflect the outcome of each shot at different stages of a frame. Analysis was carried out on 31,298 shots played during 46 matches in the 2018 World Championship and 2019 Masters; broadly reflecting the performances of a player ranked in the Top 16.

A subset of the model inputs relates to the probability of a player potting a ball at the start of a new visit, collectively representing a player's ***Scoring Potential***. This was found to be slightly higher immediately after the opponent had finished a scoring visit, typically reflecting occasions where a pot is missed.

The remaining inputs relate to the probability of a player potting a ball when continuing a visit (having potted a ball on the previous shot), collectively representing a player's ***Scoring Power***. The length of the current visit and the previous colour which was potted were identified as having some influence over the way a visit progressed.

The simulation model created was shown to fit the data well, with simulated frames exhibiting similar features to the observed matches with respect to the distributions of the number of scoring visits made, the length of the scoring visits and the number of shots played between scoring visits (§4.3.2). The only superficial difference was that the proportion of simulated frames containing a 100+ break was lower than in the observed matches, possibly due to players adapting their strategy slightly to ensure they reach that milestone.

Scaling factors were introduced into the model to reflect different levels of play by adjusting the core inputs relating to a player's scoring potential and / or scoring power (§4.4). A significant amount of additional data would be required to accurately represent individual players, but this approach yields plausible outcomes for frames contested by players of different levels of ability. Should it become possible to generate reliable estimates of the typical levels of each player's performance then the model appears capable of producing a reasonable simulation of their matches.

5.3.2 Expected Outcomes

It is not possible to formally evaluate the model's estimates of each player winning the frame as it develops. The simulated outcomes did, however, look realistic considering the state of the frame at the start of each new scoring visit (§4.3.3). The one notable difference was in the relatively high number of actual frames won from a deficit of 50+ points, but further data would be needed to assess whether this reflected a limitation in the model or was just an anomalous feature of the observed matches.

The communication of expected outcomes could be a valuable addition to the commentary in televised snooker matches. In a similar context, Jakeman (2021) described the development of an algorithm used in cricket called *WinViz*, which estimates the probability of each team winning as the match develops. Jakeman provided a quote from the former England captain and commentor, Mike Atherton, describing how he compares the model's prediction with his own estimates: "Most of the time I'll be within two or three percentage points, but when I'm massively out of sync, that's the starting point for a conversation."

The model is ultimately reflecting what has typically happened in similar situations and it would be extremely challenging to incorporate details which were specific to the current frame, such as the position of all the balls. There would always be scope for an expert commentator to complement the model's estimate with their own insight. It should be noted that in this context, a simulation model, which can take time to run, may not be the most appropriate tool for producing real-time win probabilities. A deterministic approach, perhaps using logistic regression, may be more appropriate.

An alternative use of the model in this vein would be to compare simulated and actual outcomes for individual players. In a similar way to analysis of expected goals in football, this could help to evaluate a player's ability to win a frame from different situations.

5.3.3 Decision Making

One insight from the analysis of expected outcomes is that in frames between top professionals, a good lead for one player can easily be negated if their opponent takes the next opportunity to score. To demonstrate this, a handful of scenarios were simulated to show the probability of winning at different stages of the frames given the current difference in the score and used as the basis for further exploration of the strategies involved in snooker (§4.5.2).

A further interesting insight from this is that the choice of shot should not be influenced by the current score; just on the choice of pot or safety shot which gives the player the best chance of potting the next ball. This is a slightly simplified analysis as it ignores the positions of the balls and the quality of the opportunity which is likely to present itself. Nevertheless, it does demonstrate the importance of having 'control' of the table and how quickly momentum can change in a frame at the highest level.

The significance of potting the next ball also applies to the case where a player finds themselves snookered (§4.5.3). A difficult escape which may concede multiple penalty points but leaves the balls safe, is generally preferred over a simple escape with a much higher chance of leaving a pot for the opponent. Given the potential risks of conceding either penalty points or the next pot, it is conceding the next pot which a player should treat as the bigger threat.

Section 4.5.1 considered the significance of the break-off shot in determining the outcome of the frame. Some players have started experimenting with a more cautious break-off shot to minimise the risk of leaving a pot, and the data suggests that there is certainly some justification in being concerned about how frequently a ball is potted on the second shot of the frame. Further data would be required to evaluate the actual effectiveness of any alternative shot, but the analysis presented in this research at least provides a framework for understanding how much of an influence the break-off shot has on the progression of the frame and the potential impact of any alternative approach.

The simulations run using the model have been based directly on the inputs generated from the observed data. Speculative adjustments could be made to these inputs to reflect different scenarios encountered during a match, such as how a frame might be affected by one of the baulk colours being located near a cushion and therefore harder to pot. To what extent would this affect the chances of either player winning the frame, and how might this affect their strategies? To extend this further, if a player is ahead in the frame should they opt to put a colour safe rather than taking on a risky pot in the hope of winning the frame during their current visit?

5.4 Further Work

This section identifies aspects of the research which warrant further analysis.

In analysing the rating models, it was noted that accounting for the strength of opposition faced by a player may lead to a more reliable assessment of their rating, although the predictive ability of the pairwise comparison models were not superior to the model based purely on overall win percentages. This is revisited in Section 5.4.1, analysing how much variation there was in the strength of opposition faced by different players. The potential for developing the Win Percentage model to incorporate this as a factor is also discussed.

The analysis described in this research has focussed on a player's typical level of performance. Further work is required to understand how much this varies and whether this has a significant impact on the results we would expect to observe. This is discussed in Section 5.4.2, which considers how natural variation in performance creates a prediction ceiling for any model.

The way in which the number of matches contested by a player affects the level of confidence in our assessment of their rating is also discussed and an alternative version of the Elo model is described. The development of the rating models highlighted current form as a potential factor – particularly where a player's results over the last year had been better than the previous year. Further analysis would be required to determine the best way of measuring form, whether based on a player's results or their underlying performance.

In measuring a player's performance, the tactical element of the game is the most challenging to evaluate given the limited amount of data available. Section

5.4.3 discusses potential measures of the quality of a player's safety shots, the style of their play and effectiveness of their shot selection. A more comprehensive set of data would be required to produce these, although using data from the scoring system would enable analysis to be carried out on the number of shots played between each scoring visit and the number of scoring visits each player made.

The simulation model described in Chapter 4 reflects the typical performance of a player ranked in the Top 16 as observed in two tournaments. Section 5.4.4 considers how assumptions could be developed to enable the input parameters to be scaled in order to reflect the performance of individual players over a longer period of time.

In developing the ratings models an assumption was made that the outcomes of each frame were independent of one another. Section 5.4.5 sets out analysis which could be carried out to look at how a match progresses and whether there is any indication that the outcome of a frame is dependent on how the match has progressed so far.

5.4.1 Strength of Opposition

Accumulative ranking models, such as snooker's world rankings, are based on how far a player progressed in a tournament. It does not matter whether a player had to defeat a succession of top-rated players or was fortunate to draw lower-ranked opponents - the same amount of world ranking points are awarded. Over the course of a season there are likely to be differences in the strength of the opponents faced by each player. Rather than just considering how many matches a player won, accounting for the players they faced may produce a more reliable assessment of their rating relative to other players.

One of the differences between the Win Percentage and Bradley-Terry models is that the latter is based on pairwise comparisons, effectively accounting for the strength of opposition faced by each player. The analysis reported in Section 2.4.2 found some evidence that this could improve the accuracy of match predictions when a Top 16 player faced someone ranked 17-64, but there was very little difference in the overall performance of the models. This section considers this further and discusses how the Win Percentage model could be modified to account for differences in the strength of opposition faced.

Figure 2.2 provided an indication of how the average strength of opponents varied across all professional players during the 2017/18 and 2018/19 seasons. This showed a clear difference between players at different levels of the World Rankings, with players ranked in the Top 16 tending to play stronger than average opponents. This is understandable in that players who regularly contest the latter stages of tournaments (and thus achieve a higher ranking) will have won a relatively high proportion of frames in the process. The same will typically be true of the opponents they face in later rounds.

This analysis also indicated that the average opposition faced by players typically only varied by around 1% one way or the other. To understand what the natural spread of opponents would look like, I drew random samples from the 2-year win percentages achieved by professional players at the end of 2017/18 to represent the opponents faced over 2 years. This yielded a similar spread, which perhaps goes some way to explaining why the strength of opposition does not appear to be a critical factor in establishing a relative rating for players and why the Bradley-Terry model could not significantly improve on the predictions generated by the Win Percentage model.

There may still be some merit in incorporating strength of opposition into a rating model and the Win Percentage model could itself be modified to account for the strength of opposition faced by each player. A Ratings Percentage Index has been used by the National Collegiate Athletic Association (NCAA) in the USA to rank teams across different conferences where it is not possible to organise fixtures between every pair of teams. This calculates a team's Rating Percentage based on their own win percentage (with a weight of 0.25), along with that of their opponents (weight = 0.5) and their opponents' opponents (weight = 0.25) (Pickle and Howard, 1981 as cited in Barrow, 2013).

A simpler method is likely to be sufficient for snooker given the relatively small variation in the strength of opposition faced. One approach to deriving a Weighted Win Percentage (***WWP***) would be to scale a player's Win Percentage (***WP***) based on the weighted average of their opponents' win percentages (***OP***).

$$\boldsymbol{WWP} = \boldsymbol{WP} + \boldsymbol{OP} - 0.5$$

If ***OP*** = 0.5 then no adjustment is made, while if a player’s ***WP*** = 0.5 then we conclude that they are an equivalent standard to their opponents and so ***WWP*** = ***OP***. More generally, if the level of opposition faced is x% higher than the average (i.e. 50%) then ***WWP*** = ***WP*** + x, with a similar downward adjustment if the level of opposition faced is x% lower than average.

Additional historical data are required to calculate the ***OP***. For example, in determining a player’s ***WP*** ahead of the 2019 World Championship, the previous 2 years of results are required – including results from the 2017 World Championship. To also determine their ***OP***, we require the ***WP*** of their opponents ahead of the 2017 World Championship, which is based on matches played from the 2015 World Championship onwards.

Table 5.1 shows that there is little difference in the relative rating of the 10 highest rated players at the end of the 2018/19 season. Looking more broadly at all professionals, their ***WWP*** was typically within 1-2% of their ***WP***. Long-time professional Andrew Higginson (officially ranked 57^{th} at the end of that season), saw the largest reduction in his rating (***WP*** = 52.5%, ***OP*** = 45.9%, ***WWP*** = 48.4%), with Ronnie O’Sullivan seeing the largest increase (***WP*** = 65.7%, ***OP*** = 54.4%, ***WWP*** = 70.1%)

Table 5.1: Players with the highest win percentages at the end of the 2018/19 and how they ranked when weighting their results to account for strength of opposition faced

Win Percentage			Weighted Win Percentage	
Rank	**Rating**	**Player**	**Rank**	**Rating**
1	66%	Ronnie O'Sullivan	1	70%
2	62%	Judd Trump	2	65%
3	62%	Mark Williams	3	64%
4	61%	Neil Robertson	4	62%
5	60%	John Higgins	5	62%
6	60%	Stuart Bingham	6	61%
7	59%	Kyren Wilson	9	60%
8	59%	Mark Selby	7	60%
9	59%	Joe Perry	12	58%
10	58%	Mark Allen	8	60%

As with the Win Percentage models described in Section 2.2.2, a logit model can be created to estimate the probability of each player winning a single frame against each other based on the difference in their Weighted Win Percentage. Using frames played in 2016/17 to fit such a model, a coefficient factor of 3.788 was identified for the difference in ***WWP***. [The coefficient used for the 2-year Win Percentage model was 3.764.]

An assessment of the predicted results from this model for the 2017/18 and 2018/19 seasons shows very little difference compared with the 2-year Win Percentage model (Table 5.2). The most noticeable difference is an improvement in the log-loss score and calibration measure for matches contested by a player ranked in the Top 16 against a player ranked 17-64 – similar to the results achieved by the Bradley-Terry model.

Table 5.2: A comparison of the predictions made by the 2-year Win Percentage model and the Weighted Win Percentage model

	Win Percentage model	Weighted Win Percentage model
All Matches		
Prediction Accuracy	68.8% (3041 / 4421)	68.8% (3040 / 4421)
Log-Loss score	0.592	0.591
Calibration	1.01	1.02
Discrimination	6.61	6.73
Top 16 v 17-64 (651 matches)		
Log-Loss score	0.604	0.598
Calibration	0.92	0.99
Top 16 v 65+ (551 matches)		
Log-Loss score	0.482	0.485
Calibration	1.02	1.04
17-64 v 65+ (1,562 matches)		
Log-Loss score	0.583	0.585
Calibration	1.02	0.99

The Weighted Win Percentage model described above could be refined, although ultimately it would appear that further analysis is required to understand what is driving the differences in the results and whether additional factors need to be accounted for.

It could be that top-ranked players tend to face lower-ranked opponents (ranked 65+) in the early rounds when they are still getting used to the table conditions.

These matches may also take place in front of smaller audiences than they are used to, or outside of the main arena, which could have a bigger impact on their performance compared with an opponent who is more used to these conditions. If they face higher ranked opponents (those ranked 17-64) slightly later in the tournament then they may be closer to their best form.

An alternative hypothesis is that the performance of players ranked 65+ is more variable, and that when they are at the top of their game, they pose more of a threat to the top-rated players than their typical level of play would suggest.

5.4.2 Variations in Performance

This research has focussed on a player's typical level of performance, whether in estimating their overall rating, or in the probability of a ball being potted on their next shot. Understanding the variation in a player's performance warrants further investigation and could inform further development of the models presented in this thesis.

This section considers variation in performance with respect to:

- Understanding the maximum potential for any prediction model (prediction ceiling) given that results are inevitably subject to some random variation which cannot be predicted.
- Considering whether the relative uncertainty in a player's rating could be incorporated into the models.
- Determining whether variation in a player's results should be attributed to natural variation or whether it reflects a short-term change in form or a more substantial change in their ability.
- Analysing the variation within a player's performance to understand the variation in their match results.

Outcome Uncertainty and Prediction Ceilings

One method used to estimate the maximum potential for any model is based on a measurement of the variation in observed results across all competitors (V_{OBS}). This is compared with the expected variation in results if all competitors were of an equal ability; the amount of variance which can be attributed to randomness (V_{RAND}). In a series of win / loss contests between competitors of equal ability, V_{RAND} is calculated as $(0.5 \times 0.5)/\boldsymbol{N}$, where $\boldsymbol{N}$ is the number of games played over a season (Birnbaum, 2011).

The proportion of variation in observed results which is purely due to chance is therefore estimated as: $\frac{V_{RAND}}{V_{OBS}}$. The remainder of the variation can theoretically be attributed to differences in ability or other factors which influence the outcome of a contest (such as home advantage).

The best a predictive model could feasibly achieve is to reflect the variation which is due to skill or other factors. Even this prediction ceiling is unlikely to be reached as it would require a complete understanding of all measurable differences between the competitors.

A similar approach can be applied to snooker, with the added complication that players will contest different numbers of matches depending on how far they progress in each tournament.

Looking separately at the results of professionals in the 2017/18 and 2018/19 seasons, the strongest performance was from Ronnie O'Sullivan in 2017/18, who won 45 of the 52 matches he played (87%). Three players contested at least 10 matches in a season but failed to win any.

The weighted mean performance of all professionals (not counting the results of amateur players) was a win percentage of 51.3%, calculated as $\bar{x} = \frac{\sum w_{is}}{\sum n_{is}}$, where w_{is} represents the number of matches won by player i in season s, and n_{is} represents the number of matches played by player i in season s.

The weighted variance in the proportion of matches won by professionals during 2017/18 and 2018/19 is similarly calculated as $\hat{\sigma}^2 = \frac{\sum n_{is}(x_{is} - \bar{x})^2}{\sum n_{is}}$, where x_{is} represents the win percentage of player i in season s. This gives an estimate of V_{OBS} = 0.02926.

Across all professional players, the average number of matches played per season was 33.1, which gives an estimate for V_{RAND} = 0.00755. To account for the difference in matches played, I also simulated the results from two seasons based on 128 players of equal strength contesting 20 knock-out competitions a season (a very rough approximation to the snooker season). Over 1,000 trials, the average variation in performance gave an alternative estimate of V_{RAND} = 0.00669.

These give estimates for the % variation due to randomness between 23-26%, indicating that a model with a prediction accuracy of 75% would be exceptional (and probably unrealistic). The models tested in this research, with an accuracy of 68-69% would appear to be doing quite well, albeit with some theoretical room for improvement.

Scarf at el (2021) considered the level of outcome uncertainty in different sports and proposed a broader model based on variation in the strength of the competitors, the scoring rate of the sport and any score dependence.

For snooker, score dependence is largely negligible – any (dis)advantage from breaking-off in a frame is minimal (§4.5.1) and there is no clear evidence of any dependence in the outcomes of frames (considered further in Section 5.4.4).

The scoring rate has an influence in the sense that the number of frames contested does impact on outcome uncertainty (demonstrated in Appendix V). We would therefore expect longer matches to have more predictable outcomes. The number of scoring visits required to win a frame is lower for higher-ranked players (§3.4.1), so this would also suggest that outcome uncertainty would be higher for two top-ranked players, compared with two lower-ranked players with a comparable difference in strength.

Incorporating Uncertainty

As noted above, the win percentages for different players are based on differing numbers of matches played. The more frames a player has contested over the last couple of years, the more confidence we have that their win percentage accurately reflects their true ability. Accounting for the level of uncertainty in player ratings could help to enhance the models.

Using a normal approximation, a confidence interval for each player's win percentage ($\hat{p}$) can be calculated as $\hat{p} \pm z\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$.

Over the 2017/18 and 2018/19 seasons, Judd Trump contested the most frames (962). The 95% confidence interval for his win percentage over this period is therefore 0.62 $\pm$ 0.031 [0.59, 0.65]. Simon Lichtenberg contested the fewest frames (123). The 95% confidence interval for his win percentage is therefore much wider: 0.28 $\pm$ 0.079 [0.20, 0.36].

A way in which the Elo model can be adapted is to use a different updating weight for different players based on the number of matches they have played.

The data journalism site *FiveThirtyEight* have applied Elo models to sports and vary the weight used to update a player's rating based on the number of matches they have contested (Morris & Bialik, 2015). The formula they use to allocate a weight to player i is $\frac{K}{(n_i+o)^s}$ where K is a constant multiplier, n_i is the number of matches played by player i, o is referred to as the offset and s determines the shape of the function. The values chosen by Morris and Bialik were $K = 250$, $o = 5$ and $s = 0.4$

I have followed this approach to apply sliding weights to the Elo model described in Section 2.2.4. Setting values of $K = 40$, $o = 5$ and $s = 0.5$ was found to work well; the lower value for K chosen on the basis that we are modelling frames rather than matches - as with the initial Elo model created. An extended exercise could seek to optimise these parameters for a given set of training data.

The effect of applying this formula was to set a weight of around 15 for a player who had only contested 1 or 2 matches over the previous 2 years; reducing to 10 for a player who had contested 11 matches; 5 for a player who had contested 59 matches and 4 for a player who had contested 95 matches (Figure 5.1).

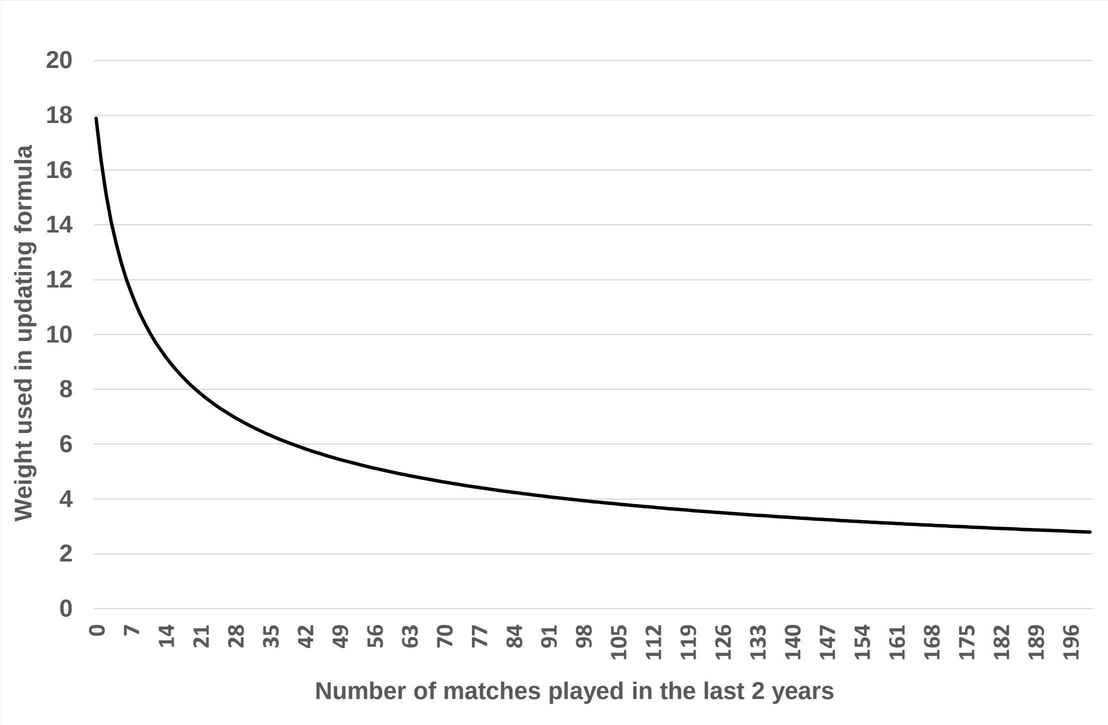


Figure 5.1: Weight used in the updating formula depending on the number of matches played by the individual over the last two years

This revised model performed slightly better against each of the metrics analysed (presented in Table 5.3), indicating that the application of sliding weights would be worthwhile.

Table 5.3: A comparison of the predictions made by the Elo model with fixed weights and an adaptation of the Elo model using sliding weights

	Elo (fixed weight)	**Elo (sliding weights)**
Prediction Accuracy	68.7% (3039 / 4421)	68.9% (3048 / 4421)
Log-Loss score	0.593	0.587
Calibration	1.04	1.03
Discrimination	7.11	7.20

The Elo model could also be adjusted to allow for a ratings deviation in common with the revised model devised by Glickman (1999), although it is perhaps more debatable whether there is greater uncertainty associated with the performance of a snooker player if they play less frequently. Ronnie O'Sullivan famously won the 2013 World Championship despite having only played one professional match in the previous year.

Changes in Form

The section above has considered natural variations in a player's results and how this affects our confidence in the rating of each player. We are also interested in being able to identify changes in the level of a player's performance. This could be tackled in a variety of ways:

A ***measure of a player's current form*** could be developed to understand what impact this has on a player's results. McHale and Forrest (2005) developed a predictive model based on a golfer's finishing position in their last 6 competitions (along with a longer-term indicator of their ability) and found that recent results did have a strong impact on a player's likely finishing position in the following tournament.

This is arguably more straightforward for a sport such as golf, where each player completes at least 2 rounds during each tournament and their score is not directly affected by any other player. In snooker, a player can perform reasonably well but still lose in the first round – or struggle to find their best form but do just enough to defeat a succession of opponents.

Nonetheless, a separate measure of a player's current form over the last *n* matches, or last *t* months could be developed and incorporated into the logit models based on World Rankings or Win Percentages.

An alternative way of developing the World Ranking, Win Percentage and Bradley-Terry models would be to ***introduce a decay factor***, such that recent performances would be weighted more heavily (McHale and Morton, 2011). It is likely that this would support the use of a longer series of historical results, with older results contributing to (but not unduly influencing) a player's rating.

Further analysis would be required to assess the merits of either approach. The former requires a suitable time frame to be identified allowing for the variation in the number of matches contested by different players. The latter essentially assumes that more recent results are better predictors of a player's relative rating, which this analysis has shown is not always the case.

All of the models described in this research are essentially attempting to establish a reliable rating of a player's ability. An alternative way of approaching the effect of form is to consider whether we can identify when a player is performing at a different level of ability than their current rating would suggest. Rather than automatically incorporating a measure of current form in the model, we would only do so in the case that a player's results were consistently better (or worse) than those predicted by our current assessment of their ability. The objective then becomes one of identifying when this has happened and determining how best to adjust our assessment of their ability.

Changes in Performance

Rather than just looking at a player's results, the analysis carried out in this research points towards an alternative way of identifying and capturing changes in performance.

With sufficient data we would be able to analyse how a player’s performance measures vary across different matches. The effects of the match length and the level of opposition faced could also be considered.

The simulation model developed would then provide an indication of how much of an impact such a variation in performance would have on the probability of one player winning a frame against another.

These results could then be used to inform the development of the prediction models. Analysing changes in a player’s performance measures may also be a more effective way of assessing the current form of a player, or whether there has been a sustained improvement in their level of play.

5.4.3 Developing Measures of Safety Play

The measure of a player's *scoring power* described in Chapter 3 reflects a player's break-building capability and is a strong indicator of performance.

It is more challenging to identify a reliable indicator of a player's safety game. The official measure of *safety success* could be suitable if sufficient data were available to reliably measure the performance of all professional players. There would also need to be widely recognised method for classifying a safety shot as being either successful or unsuccessful. The subjective nature of the measure is a significant limitation though, meaning that it is only currently produced for televised matches. Even if the data which is currently collected were publicly available, there would not be enough to evaluate the typical level of performance for most players. Taking each match in isolation, the measure offers only limited value in explaining the outcome.

This research has indicated that there are additional measures which could be developed using data from the automated scoring system based on the number of scoring visits made in a match. With access to sufficient data there is likely to be value in exploring these in more detail.

One potential area of analysis is to assess the type of frame or match being played based on the number of scoring visits and the number of shots between each scoring visit. An open frame would consist of a higher proportion of pots and is likely to have relatively few scoring visits, with players able to score heavily from the chances they get. A more tactical frame would consist of a larger number of shots which don't result in a pot; while we might also expect that fewer points are scored from each scoring visit.

We would expect to see that some players are more likely to be involved in one type of frame over the other depending on whether they are prepared to take on more attacking shots (leading to an open frame) or prefer to be more cautious (leading to a more tactical frame). It would also be interesting to see if some players are more likely to win one type of frame or another – and whether this tallies with their style of play.

An alternative approach could seek to identify key moments in a frame through which an individual's tactical play could be assessed.

- Do some players pot the first ball in a frame more frequently than others?
- If a lengthy safety exchange has developed (i.e. a number of shots have been played since the last pot), are some players more likely to pot the next ball?
- When a player's scoring visit comes to an end, how frequently does their opponent pot a ball on their next shot?
- Do some players have a better record at winning a frame when they require snookers?

More sophisticated measures of an individual's safety play would require additional information to be collected manually – it is unlikely that a modification to the scoring system would be appropriate as any additional data entry would only hinder the scorer's ability to keep score. Capturing some basic information on the type of shot played would, however, allow enable further analysis of a player's shot choice and whether they were more likely to concede chances to their opponent through a missed pot or a poorly executed safety shot (§3.4.3).

5.4.4 Simulating Individual Players

The simulation model described in Chapter 4 is based on the typical performance of a player ranked within the Top 16. Section 4.6 discussed the potential for modelling a frame between players with different levels of performance, scaling the input parameters to reflect the scoring potential and scoring power of individual players.

The data collected for this research only covers two tournaments, so while a player's performance within the observed matches could be measured, this would not necessarily provide a reliable indication of their play over a longer period of time. A significant amount of additional data would be needed to develop suitable scaling factors for all professional players.

Data from the automated scoring system, which covers all professional matches played, would allow this. For each player this would provide shot-by-shot data from around 100-500 frames contested over the course of a season (depending on the extent of a player's participation in tournaments and how far they progressed). Access to the data can be acquired through *SportRadar*, but this service is only available to licensed bookmakers and media companies.

As an alternative, it may be possible to create an informed adjustment by using publicly available data to supplement the data collected in this research. There would be three separate steps involved in this approach.

Step 1 – establishing the relative importance of a player's scoring potential and scoring power on the outcome of a frame.

The post-match video analysis carried out for this research covered 734 frames. Looking at those players contesting at least 25 observed frames, a regression model can be created for the proportion of frames won by a player based on their scoring potential and scoring power.

An appropriate outcome variable would be the proportion of frames won (FW) by each player above the average (y = FW - 50%).

Scoring power for an individual (Pow) can be captured relative to the average scoring power observed (x_1 = Pow - 89%).

In looking at scoring potential, it is more salient to look at the observed scoring potential of a player (Pot) relative to their opponents (Opp), as one is likely to be influenced by the other. The aggregated scoring potential of their opponents during the observed frames can be calculated as the proportion of all new visits made by the player's opponents where a ball was potted. (x_2 = Pot - Opp).

A resultant model produced on this basis gave the following equation for estimating a player's win percentage based on their scoring power and their scoring potential relative to their opponents:

$$FW = (Pow - 0.89) * 1.96 + (Pot - Opp) * 1.02 + 0.5$$

The coefficients estimated for this model are consistent with the simulated findings presented in Chapter 4, which estimate the effects of changes to a player's scoring power (Table 4.11) and scoring potential (Table 4.12) on their probability of winning a frame.

Step 2 – using the rate at which a player makes 50+ breaks to estimate their scoring power

The number of 50+ breaks made by a player over the course of a season is captured by CueTracker.net (Florax, 2019d), with the rate at which they make these breaks providing an alternative indication of their relative scoring power. Based on ranking events contested during the 2018/19 and 2019/20 seasons, players ranked in the Top 16 at the end of the 2019/20 season averaged a rate of 39.3 50+ breaks per 100 frames (0.393 for a single frame), with Ronnie O'Sullivan achieving the highest rate of 49.6 (0.496 for a single frame).

Table 4.10 provides an indication of the relationship between scoring power and 50+ breaks made. Running the simulation model using a scaling factor of 1, the two players had a combined scoring power of 88.8% and collectively made a break of 50+ in 68% of the frames simulated (which we can infer to be 34% for each player). Increasing their scoring power by 2.2% resulted in 50+ breaks made in 79% of simulated frames (inferred to be 39.5% for each player).

An increase in a player's scoring power of 1% therefore roughly increased the rate at which they made 50+ breaks by 2.5 per 100 frames; equivalently if the rate at which 50+ breaks are made is higher by 1 per 100 frames, this is indicative of an increase in a player's scoring power of 0.4%.

If we take the average scoring power of a player in the Top 16 to be 89%, we can estimate a player's scoring power based on the rate at which they made 50 breaks. O'Sullivan therefore had an estimated scoring power of 93.1% over this period [0.89 + (0.496 – 0.393)*0.4]. Luca Brecel recorded the lowest rate of 50+ breaks made (0.319) so had the lowest estimated scoring power of 86.0%.

Step 3 – estimating a player’s scoring potential based on their win percentage and scoring power

The average proportion of frames won by the Top 16 players in these two seasons was 59.0%. The coefficient for scoring power from our regression equation (1.96) can therefore be used to estimate the proportion of frames we would expect each individual player to have won based purely on differences in their scoring power [E(FW) = 0.590 + (Pow – 0.89) * 1.96].

This can be compared with their actual win percentage, with the inference being that the difference is due to the relative difference in their scoring potential compared with a typical opponent (taken to be 27% based on the observed matches). The coefficient in the regression model for scoring potential relative to the opponent was 1.02, so a player’s scoring potential is estimated as: E(Pot) = 0.27 + [FW – E(FW)] / 1.02

Based on our estimate of scoring power for Luca Brecel, he would only have been expected to win 53% of frames. He actually won 56%, suggesting that his scoring potential was around 3% higher than his opponents, giving an estimate of 29.9%. O’Sullivan won a lower percentage of frames than indicated by his scoring power (66% v 67%), giving an estimate for his scoring potential of 25.6%. Estimates for each of the players ranked in the Top 16 at the end of 2019/20 are presented in Table 5.4.

Table 5.4: Estimates of scoring power and scoring potential for each player ranked in the Top 16 at the end of the 2019/20 season, based on the rate at which they made 50+ breaks in 2018/19 and 2019/20 and the proportion of frames they won during this period.

Player (World Ranking at end of 2019/2020)	50+ breaks / 100 frames	Scoring Power	Expected Frame Win %	Actual Frame Win %	Scoring Potential
Ronnie O'Sullivan (1)	46.0	93.1%	67.1%	65.7%	25.6%
Judd Trump (2)	48.8	90.6%	62.3%	62.1%	26.7%
Mark Williams (3)	45.2	89.5%	60.0%	62.1%	28.9%
Neil Robertson (4)	41.1	91.4%	63.8%	60.8%	23.9%
John Higgins (5)	42.0	88.5%	58.0%	60.0%	28.8%
Mark Selby (6)	40.2	89.5%	60.2%	59.0%	25.8%
Mark Allen (7)	32.4	89.3%	59.7%	58.2%	25.4%
Kyren Wilson (8)	39.4	88.9%	58.9%	59.2%	27.2%
Barry Hawkins (9)	35.0	88.2%	57.6%	57.6%	26.9%
Ding Junhui (10)	34.5	88.1%	57.3%	58.2%	27.8%
Jack Lisowski (11)	37.2	89.1%	59.2%	57.8%	25.5%
David Gilbert (12)	38.3	87.2%	55.7%	55.5%	26.7%
Stuart Bingham (13)	37.0	88.6%	58.3%	59.9%	28.5%
Shaun Murphy (14)	37.9	87.9%	56.9%	55.6%	25.7%
Luca Brecel (15)	29.1	86.0%	53.2%	56.3%	29.9%
Stephen Maguire (16)	33.9	87.5%	56.1%	56.7%	27.4%

Additional analysis would be desirable to the test the assumption that all inputs could be scaled – but this would require data from the scoring system. This would explore whether there was:

- Any difference in scoring rates across a frame? (e.g. are some players more adept at developing the last reds?)
- Any interaction between scoring potential and scoring power? (i.e. do more defensive players affect the scoring power of their opponent?)

5.4.5 Progression of a Snooker Match

Chapter 4 presented a model of a snooker frame and from here it is a natural progression to consider how to model an entire match. There is insufficient data to analyse how a player's performance changes during a match and whether it is affected by winning or losing frames. The models created in Chapter 2 can, however, be used to evaluate whether a player's probability of winning the next frame is influenced by what happened during previous frames. In using the models to predict match outcomes an assumption was made that each frame was independent, but how accurate is this and could the models be improved by making an alternative assumption?

To explore this, I have initially looked at how all "Best of 7 frames" matches (1,877) progressed during the 2017/18 and 2018/19 seasons. This was the most common match length, but the analysis could easily be extended to longer matches.

Table 5.5 compares the actual and expected proportion of frames won by the higher-rated player at different stages of the frame – grouped according to whether the higher-rated player was level, ahead or behind at the start of the frame. The expected results are based on the 2-year Win Percentage model described in Section 2.2.2. The calibration measure is presented to provide an indication of potential bias in the expected results, with scores below 1.0 indicating that the higher-rated player won more frames than expected.

There is no significant evidence of bias in any of the results, although there are some notable features which would warrant further analysis.

Table 5.5: A comparison of actual and expected frames won by the higher-rated player during "Best of 7 frames" matches played in the 2017/18 and 2018/19 seasons.

Current Score	# Frames	% Frames won by higher-rated player		Calibration
		Actual	Expected	
0-0	1,875	57%	58%	1.01
1-1	885	58%	58%	1.00
2-2	672	56%	58%	1.03
3-3	501	53%	57%	1.07
All level	2,230	57%	58%	1.02
1-0	1,075	58%	58%	1.00
2-0	627	60%	59%	0.97
3-0	379	67%	59%	0.89
2-1	759	55%	58%	1.05
3-1	544	60%	58%	0.98
3-2	596	58%	58%	1.00
All ahead	3,980	59%	58%	0.99
0-1	800	55%	57%	1.05
0-2	363	52%	57%	1.08
0-3	173	54%	56%	1.05
1-2	564	59%	57%	0.97
1-3	326	55%	56%	1.02
2-3	475	53%	57%	1.07
All behind	2,701	55%	57%	1.04
All frames	10,614	57%	58%	1.01

The higher-rated player won fewer frames when the score was 3-3 (known as 'deciders') than the win percentage model expected (53% v 57%). Their previous results may be less relevant for predicting the winner given that the players were evenly matched throughout the contest so far – although the higher-rated player did still prevail in the majority of these frames.

501 matches went to a deciding frame, of which 52% were won by the player who had won the previous frame (against an expectation of 50%). This suggests a slight advantage from having some momentum going into the decider, although there was insufficient evidence of a significant advantage – the confidence interval for the calibration measure was (0.87, 1.04).

At any stage during the match, the player who had won the previous frame also won the next frame on more occasions (53% of 8,739 frames) than was expected (51%).

Extending this analysis to look at additional years of results, or different lengths of matches may yet provide some evidence of a dependence in the outcomes of consecutive frames, although at this stage the assumption of independence would appear to be justified.

A similar implicit assumption made is that the outcomes of matches played in the same tournament are independent of another. A couple of questions of particular interest are:

1) Does a lower-ranked player fare better than expected if facing a high-ranked player in the first round of a tournament, rather than at a later stage?

2) In tournaments where players enter at different stages (such as the qualifying rounds of the World Championship), does a player who has already won a match have an advantage against an opponent playing their first match?

5.5 Concluding Remarks

In seeking to determine what differentiates the best snooker players from the rest, this research first compared and evaluated four methods of quantifying the relative ability of professional snooker players.

This analysis found little difference between the methods in who they identified as the strongest players. In common with the views of more subjective observers, the ordering varies slightly depending on the time-period analysed and whether the method weights performance in some tournaments more heavily than others (whether due to their profile or recency).

There were larger differences between the methods in their rating of lower-ranked players, largely dependent on the quantity of historical results used – with the official world rankings notably ignoring the past results of players when they were amateurs and / or during previous spells on the professional tour.

The proportion of frames won by players over the previous two years was found to provide a reliable and transparent basis for rating players. Developing an assessment of a player's recent form to complement this may be the most effective way to strengthen the relative ratings produced.

This research has also looked at the merits of various measures used to assess the performance of snooker players, concluding that the existing statistics which are produced offer limited value.

There is greater for potential for developing measures based on the scoring visits made by players during a match, considering whether some players are more adept at creating chances for themselves (or at restricting the chances given to their opponent), and whether some players score more heavily from the scoring visits they make.

Objective measures could be developed using data collected automatically by the scoring system. Further data and analysis would be required to establish effective indicators of a player's safety game, although the measure of *scoring potential* proposed in this research provides some indication of this. The complementary measure of a player's *scoring power* captures the dynamic nature of the game and is shown to be a reliable indicator of a player's break-building prowess – a key element of their performance.

These two measures also form the basis of the final part of this research, the development of a Monte Carlo simulation model of a frame of snooker. This was demonstrated to reliably reflect frames played between top-level professionals, with the potential for providing a framework for understanding how differences in player performance affect the outcome of a frame.

The simulation model also offers potential for evaluating the merits of different types of shot, providing some insight into the relative importance of being the player to gain the next scoring opportunity.

A player is generally advised to take on a pot if they estimate their chances of success to be greater than their chances of potting the next ball should they play safe – even if they would certainly leave a chance for their opponent if they missed. When faced with the risk of conceding a scoring opportunity to their opponent, they are advised to play a more difficult shot to avoid this – even if they may concede penalty points in the process.

Appendices

I – Professional snooker tournaments

A summary of all ranking events played during the 2016/17, 2017/18 and 2018/19 seasons

European Events	Asian Events
Riga Masters	China Championship [3]
Paul Hunter Classic [1]	Shanghai Masters [4]
European Masters	International Championship
German Masters	Indian Open
Snooker Shoot-Out [2]	World Open
Gibraltar Open	China Open

[1] Contested for the last time as a ranking event in 2018/19

[2] Contested as a single frame of snooker played under time constraints and modified rules

[3] Non-ranking invitational event in 2016/17; full ranking event in 2017/18 & 2018/19

[4] Full ranking event in 2016/17 & 2017/18; non-ranking invitational event in 2018/19

Home Nations Series	Coral Series [5]
English Open	World Grand Prix [6]
Northern Ireland Open	Players Championship [7]
Scottish Open	Tour Championship [8]
Welsh Open	

[5] Series sponsored by Coral in 2019 & 2020 and subsequently by Cazoo from 2021.

[6] Top 32 players on the current season’s money list

[7] Top 16 players on the current season’s money list

[8] Top 8 players on the current season’s money list. Introduced in 2018/19.

Triple Crown Series
UK Championship
Masters [9]
World Championship

[9] A non-ranking invitational event contested by the Top 16 in the World Rankings

II – Score sheet from the deciding frame played during the final of the 2020 World Seniors Snooker Championship

Frame sheet: White vs Doherty - Frame 9

Shot	Shot time
22:16:02 - Frame 9 ended. White won 108(85)-2.	
White potted the pink. Break 85	23.4
White potted the blue. Break 79	11.3
White potted the brown. Break 74	10.3
White potted the green. Break 70	7.4
White potted the yellow. Break 67	27.9
White potted the black. Break 65	10.1
White potted a red. Break 58	14.3
White potted the black. Break 57	11.5
White potted a red. Break 50	17.8
White potted the pink. Break 49	15
White potted a red. Break 43	25
White potted the black. Break 42	30.6
White potted a red. Break 35	31.7
White potted the green. Break 34	61.9
White potted a red. Break 31	19.3
White potted the black. Break 30	24.8
White potted a red. Break 23	22.4
White potted the blue. Break 22	49.6
White potted a red. Break 17	30.6
White potted the black. Break 16	17.5
White potted a red. Break 9	14.8
White potted the black. Break 8	18.9
White potted a red. Break 1	43
Doherty played safe.	35.2
White played safe.	24.8
Doherty played safe. Break ends at 1.	19.3
Doherty potted a red. Break 1	29.4
White played safe.	67.3
Doherty played safe.	28
White played safe.	32.8
Doherty played safe. Break ends at 1.	35.5
Doherty potted a red. Break 1	31.3
White played safe. Break ends at 23.	23
White potted the blue. Break 23	20.7
White potted a red. Break 18	57.8
White potted the pink. Break 17	17.2
White potted a red. Break 11	12
White potted the yellow. Break 10	29.3
White potted a red. Break 8	38.3
White potted the pink. Break 7	39.4
White potted a red. Break 1	11.7
Doherty missed a red.	43.8
White missed a red.	41.8
Doherty played safe.	39.4
White played safe.	-
21:54:28 - Frame 9. Jimmy White to break.	

00:21:34 duration
45 shots played
15 reds potted
2 yellows potted
2 greens potted
1 brown potted
3 blues potted
4 pinks potted
6 blacks potted
0 fouls committed

III – Matches recorded using post-match video analysis

2018 World Championship

Match	Round	Winner	Loser	Score
1	Last 32	Joe Perry	Mark Selby	10 – 4
2	Last 32	Mark Allen	Liam Highfield	10 – 5
3	Last 32	Kyren Wilson	Matthew Stevens	10 – 3
4	Last 32	Jamie Jones	Shaun Murphy	10 – 9
5	Last 32	John Higgins	Thepchaiya Un-Nooh	10 – 7
6	Last 32	Jack Lisowski	Stuart Bingham	10 – 7
7	Last 32	Ricky Walden	Luca Brecel	10 – 6
8	Last 32	Judd Trump	Chris Wakelin	10 – 9
9	Last 32	Ding Junhui	Xiao Guodong	10 – 3
10	Last 32	Anthony McGill	Ryan Day	10 – 8
11	Last 32	Lu Haotian	Marco Fu	10 – 5
12	Last 32	Barry Hawkins	Stuart Carrington	10 – 7
13	Last 32	Mark Williams	Jimmy Robertson	10 – 5
14	Last 32	Robert Milkins	Neil Robertson	10 – 5
15	Last 32	Ali Carter	Graeme Dott	10 – 8
16	Last 32	Ronnie O'Sullivan	Stephen Maguire	10 – 7
17	Last 16	Mark Allen	Joe Perry	13 – 8
18	Last 16	Kyren Wilson	Jamie Jones	13 – 5
19	Last 16	John Higgins	Jack Lisowski	13 – 1
20	Last 16	Judd Trump	Ricky Walden	13 – 9
21	Last 16	Ding Junhui	Anthony McGill	13 – 4
22	Last 16	Barry Hawkins	Lu Haotian	13 – 10
23	Last 16	Mark Williams	Robert Milkins	13 – 7
24	Last 16	Ali Carter	Ronnie O'Sullivan	13 – 9
25	Quarter-Final	Kyren Wilson	Mark Allen	13 – 6
26	Quarter-Final	John Higgins	Judd Trump	13 – 12
27	Quarter-Final	Barry Hawkins	Ding Junhui	13 – 5
28	Quarter-Final	Mark Williams	Ali Carter	13 – 8
29	Semi-Final	John Higgins	Kyren Wilson	17 – 13
30	Semi-Final	Mark Williams	Barry Hawkins	17 – 15
31	Final	Mark Williams	John Higgins	18 – 16

2019 Masters

Match	Round	Winner	Loser	Score
32	Last 16	Luca Brecel	Mark Allen	6 – 5
33	Last 16	Ding Junhui	Jack Lisowski	6 – 1
34	Last 16	Ryan Day	John Higgins	6 – 5
35	Last 16	Ronnie O'Sullivan	Stuart Bingham	6 – 2
36	Last 16	Mark Selby	Stephen Maguire	6 – 2
37	Last 16	Judd Trump	Kyren Wilson	6 – 2
38	Last 16	Barry Hawkins	Shaun Murphy	6 – 2
39	Last 16	Neil Robertson	Mark Williams	6 – 3
40	Quarter-Final	Ding Junhui	Luca Brecel	6 – 5
41	Quarter-Final	Ronnie O'Sullivan	Ryan Day	6 – 3
42	Quarter-Final	Judd Trump	Mark Selby	6 – 2
43	Quarter-Final	Neil Robertson	Barry Hawkins	6 – 3
44	Semi-Final	Ronnie O'Sullivan	Ding Junhui	6 – 3
45	Semi-Final	Judd Trump	Neil Robertson	6 – 4
46	Final	Judd Trump	Ronnie O’Sullivan	10 - 4

IV – Sample of data collected using post-match video analysis

Shots played during the 1st frame of the 2018 World Championship Final

							Derived information							
			Raw Data				Statistical Data			Modelling				
T	M #	F #	S #	Player	Shot	Pts	Break	S/Pot	S/Pow	Reds	Length	Lead	Rem	Notes
WC	31	1	1	Williams	S			0		15	0	0	147	
WC	31	1	2	Higgins	S			0		15	0	0	147	
WC	31	1	3	Williams	S			0		15	0	0	147	
WC	31	1	4	Higgins	S			0		15	0	0	147	Left a red near corner pocket
WC	31	1	5	Williams	P	1		1		15	0	0	147	
WC	31	1	6	Williams	P	7			1	14	1	1	146	
WC	31	1	7	Williams	P	1			1	14	2	8	139	
WC	31	1	8	Williams	P	7			1	13	3	9	138	
WC	31	1	9	Williams	P	1			1	13	4	16	131	
WC	31	1	10	Williams	P	6			1	12	5	17	130	
WC	31	1	11	Williams	P	1			1	12	6	23	123	
WC	31	1	12	Williams	P	7			1	11	7	24	122	Used rest
WC	31	1	13	Williams	P	1			1	11	8	31	115	
WC	31	1	14	Williams	P	5	37		1	10	9	32	114	Lost position
WC	31	1	15	Williams	MP				0	10	10	37	107	Wild attempt but left safe

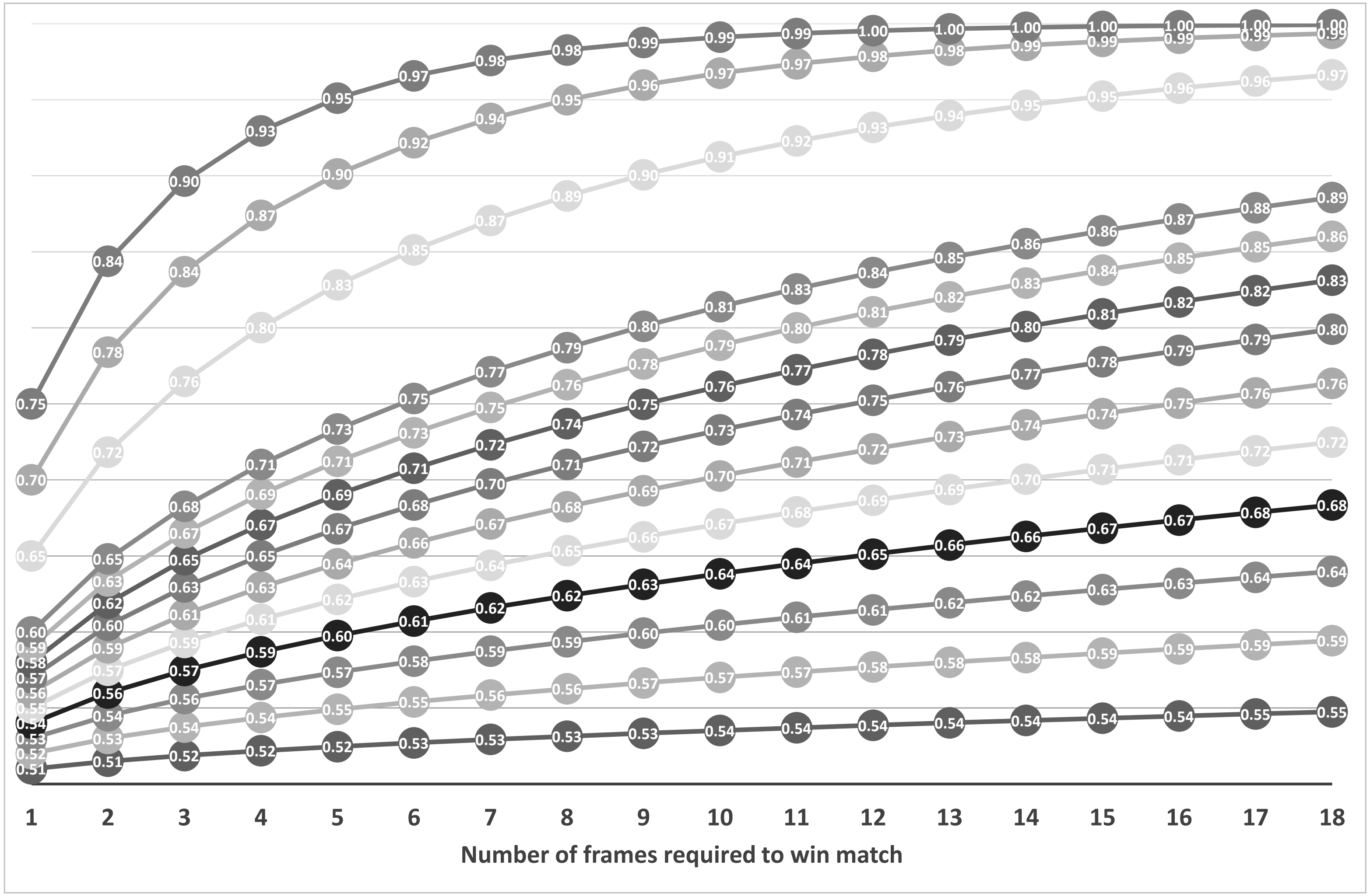
							Derived information							
				Raw Data				Statistical Data			Modelling			
T	M #	F #	S #	Player	Shot	Pts	Break	S/Pot	S/Pow	Reds	Length	Lead	Rem	Notes
WC	31	1	16	Higgins	S			0		10	0	-37	107	
WC	31	1	17	Williams	S			0		10	0	37	107	
WC	31	1	18	Higgins	S			0		10	0	-37	107	
WC	31	1	19	Williams	S			0		10	0	37	107	Left red close to corner pocket
WC	31	1	20	Higgins	P	1		1		10	0	-37	107	Long pot
WC	31	1	21	Higgins	P	5			1	9	1	-36	106	
WC	31	1	22	Higgins	P	1			1	9	2	-31	99	
WC	31	1	23	Higgins	P	7			1	8	3	-30	98	
WC	31	1	24	Higgins	P	1			1	8	4	-23	91	
WC	31	1	25	Higgins	P	7			1	7	5	-22	90	Missed cannon on reds
WC	31	1	26	Higgins	P	1	23		1	7	6	-15	83	Lost position
WC	31	1	27	Higgins	MP				0	6	7	-14	82	Missed difficult cut on black
WC	31	1	28	Williams	P	1		1		6	0	14	75	
WC	31	1	29	Williams	P	5			1	5	1	15	74	
WC	31	1	30	Williams	P	1			1	5	2	20	67	
WC	31	1	31	Williams	P	7			1	4	3	21	66	
WC	31	1	32	Williams	P	1			1	4	4	28	59	
WC	31	1	33	Williams	P	4			1	3	5	29	58	
WC	31	1	34	Williams	P	1			1	3	6	33	51	
WC	31	1	35	Williams	P	5	25		1	2	7	34	50	

							Derived information							
			Raw Data				Statistical Data			Modelling				
T	**M #**	**F #**	**S #**	**Player**	**Shot**	**Pts**	**Break**	**S/Pot**	**S/Pow**	**Reds**	**Length**	**Lead**	**Rem**	**Notes**
WC	31	1	36	Williams	MP				0	2	8	39	43	Missed red along black cushion
WC	31	1	37	Higgins	S			0		2	0	-39	43	Caught green with safety
WC	31	1	38	Williams	P	1		1		2	0	39	43	Long pot
WC	31	1	39	Williams	P	4	5		1	1	1	40	42	Higgins requires snookers
WC	31	1	40	Williams	MP				0	1	2	44	35	Missed long red
WC	31	1	41	Higgins	S			0		1	0	-44	35	
WC	31	1	42	Williams	P	1		1		1	0	44	35	
WC	31	1	43	Williams	P	7	8		1	0	1	45	34	Frame conceded by Higgins

Key:

Raw	**T**	Tournament (WC = World Championship)
	M #	Match number (within tournament)
	F #	Frame number (within match)
	S #	Shot number (within frame)
	Player	Player taking the shot
	Shot	Type of shot played (S = Safety; P = Pot; MP = Missed Pot)
	Pts	Points scored during shot
Derived – Statistical Data	**Break**	Points scored during break (recorded against the last successful pot)
	S/Pot	Included in measure of scoring potential (1 = ball potted at start of visit; 0 = no ball potted)
	S/Pow	Included in measure of scoring power (1 = ball potted on continuation of visit; 0 = no ball potted)
Derived – Modelling Inputs	**Reds**	Reds remaining at the start of the shot
	Length	Number of shots played in the current visit (0 = new visit)
	Lead	Difference in score at the start of the shot from perspective of player taking the shot (negative value = player is behind)
	Rem	Maximum points remaining on the table at the start of the shot
Notes		A sample of the notes made against key events (Long pot, rest used, snookers required, poor safety, lost position etc.)

V – Probability of winning a match (first to win n frames) based on the probability of winning a frame



VI - Transition matrix for a snooker match, populated with data from the 2018 World Championship final based on post-match video analysis

	Mark Williams (MW)				John Higgins (JH)				
	New Visit		Continuation		New Visit		Continuation		Frame End
	Pot[1]	Safety	Pot[1]	Safety	Pot[1]	Safety	Pot[1]	Safety	
MW: Break-off					29% (5)	71% (12)			
MW: New Visit - attempted pot			65% (57)	8% (7)	20% (18)	7% (6)			
MW: New Visit safety		4%[2] (4)			23% (24)	73% (77)			
MW: Continuation - attempted pot			89% (467)	3% (16)	3% (15)	2% (12)			3% (14)
MW: Continuation - safety					4% (1)	78% (18)			17% (4)
JH: Break-off	35% (6)	65% (11)							
JH: New Visit - attempted pot	25% (16)	6% (4)					63% (40)	5% (3)	
JH: New Visit - safety	37% (49)	58% (76)				5%[2] (6)			
JH: Continuation – attempted pot	3% (16)	1% (3)					91% (424)	2% (7)	3% (14)
JH: Continuation - safety	10% (1)	70% (7)							20% (2)

[1] Attempted pot, which may or may not have been successful

[2] Represents shots re-taken following a foul

VII - Input probabilities used in the simulation model

Input	Stage of frame	Observed Shots	Probability of potting a ball	Test against Null hypothesis[1] Null	χ^2	P-value
	1st Red (Start of Frame)					
1	Shot 1 (break-off)	734	**0.000**	0.189	210.93	0.000
2	Shot 2	734	**0.200**	0.233	5.76	0.016
3	Subsequent shots	2,411	**0.243**			
	Reds 2 - 12 (Safety Exchange)					
4	Reds 2 - 4	1,308	**0.252**	0.271	4.05	0.044
5	Reds 5 - 12	1,884	**0.285**			
	Reds 2 - 12 (Countering Scoring Visit)					
6	Reds 2 - 4	517	**0.358**	0.411	9.32	0.002
7	Reds 5 - 12	969	**0.440**			
	Reds 2 - 12 (Continuing Visit)					
8	Last Pot = Baulk Colour	923	**0.823**	0.891	78.10	0.000
9	Last Pot = Blue	1,766	**0.872**			
10	Last Pot = Pink	1,394	**0.930**			
11	Last Pot = Black	3,179	**0.905**			
	Reds 13 & 14 (New Visit)					
12	Safety Exchange	416	**0.207**	0.251	17.02	0.000
13	Countering Scoring Visit	145	**0.379**			
	Reds 13 & 14 (Continuing Visit)					
14	Visit length <= 8 Shots	464	**0.808**	0.842	6.73	0.009
15	Visit length >= 10 Shots	716	**0.865**			
	15th Red (New Visit)					
16	New Visit	466	**0.182**			
	15th Red (Continuing Visit)					
17	Visit length <= 4 Shots	95	**0.589**	0.704	7.44	0.006
18	Visit length >= 6 Shots	381	**0.732**			
	Colour following 1st red of a visit					
19	1st Red of frame	731	**0.770**	0.803	16.12	0.000
20	Reds 2 - 13	1,555	**0.826**			
21	Reds 14 & 15	149	**0.725**			
	Colour following a subsequent red					
22	Reds 2 - 4	1,674	**0.924**	0.941	11.47	0.001
23	Reds 5 - 13 & 15th Red	5,660	**0.946**			
	14th Red					
24	- Visit length < 13 Shots	204	**0.877**	0.915	6.36	0.012
25	- Visit length >= 13 Shots	264	**0.943**			
	Final Colours (New Visit)					
26	Safety Exchange	525	**0.173**	0.187	3.60	0.058
27	Countering Scoring Visit	121	**0.248**			
	Final Colours (Continuing Visit)					
28	Visit length <= 2 shots	233	**0.721**	0.864	63.46	0.000
29	Visit length = 3 - 8 shots	412	**0.828**			
30	Visit length = 9 - 23 shots	741	**0.889**			
31	Visit length > 23 shots	527	**0.920**			

[1] The null hypothesis for each group of inputs is that there is no difference in the probability of potting a ball

References

Abernethy, B., Neal, R. J., & Koning, P. (1994). Visual-perceptual and cognitive differences between expert, intermediate, and novice snooker players. *Applied Cognitive Psychology*, 8, 185–211. https://doi.org/10.1002/acp.2350080302

Agresti, A. (2013). *Categorical Data Analysis*. (3rd ed. – first published in 1990). Hoboken, N.J.: Wiley-Interscience.

Anderson, C., & Sally, D. (2014). *The Numbers Game: Why everything you know about football is wrong*. London: Penguin.

Arastey, G.M. (2018). *Opta Sports: The Leading Sports Data Provider*. Retrieved from https://www.sportperformanceanalysis.com/article/opta-leading-sport-data-provider. Accessed November 17, 2021

Arastey, G.M. (2020). *The increasing presence of data analytics in golf*. Retrieved from https://www.sportperformanceanalysis.com/article/increasing-presence-of-data-analytics-in-golf. Accessed November 17, 2021

Årdalen, H. (2019a). *Finals 2018/19*. Retrieved from http://www.snooker.org/res/index.asp?season=2018&template=7. Accessed November 13, 2021.

Årdalen, H. (2019b). *Historic Seedings 2018/19*. Retrieved from http://www.snooker.org/res/index.asp?template=25&season=2018. Accessed November 13, 2021.

Baker, R.D., & McHale, I.G. (2014). A dynamic paired comparisons model: Who is the greatest tennis player? *European Journal of Operational Research* 236(2), 677-684. https://doi.org/10.1016/j.ejor.2013.12.028

Baker, R.D., & McHale, I.G. (2017). An empirical Bayes model for time-varying paired comparisons ratings: Who is the greatest women's tennis player? *European Journal of Operational Research* 258(1), 328-333. https://doi.org/10.1016/j.ejor.2016.08.043

Barrow, D., Drayer, I., Elliott, P., Gaut, G., & Ostling, B. (2013). Ranking rankings: an empirical comparison of the predictive power of sports ranking methods. *Journal of Quantitative Analysis in Sports*, 9(2), 187-202. https://doi.org/10.1515/jqas-2013-0013

Ben-Naim, E., Vazquez, F., & Redner, S. (2006). Parity and predictability of competitions. *Journal of Quantitative Analysis in Sports*, 2(4), Article 1. https://doi.org/10.2202/1559-0410.1034

Biermann, C (2019). *Football Hackers: The science and art of a data revolution*. Kindle edition. London: Blink Publishing.

Birnbaum, P. (2011). *The Tango method of regression to the mean -- a proof*. Retrieved from: http://blog.philbirnbaum.com/2011/08/tango-method-of-regression-to-mean-kind.html. Accessed November 17, 2021.

Bradley, R.A., & Terry, M.E. (1952). Rank analysis of incomplete block designs 1: The method of paired comparisons. *Biometrika* 39, 324-345. https://doi.org/10.2307/2334029.

Brechot, M., & Flepp, R. (2020). Dealing with randomness in match outcomes: How to rethink performance evaluation in European club football using expected goals. *Journal of Sports Economics*, 21(4), 335-362. https://doi.org/10.1177/1527002519897962

Brier, G.W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1), 1-3. https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

Broadie, M. (2012). Assessing golfer performance on the PGA tour. Interfaces, INFORMS, 42(2), 146-165. https://doi.org/10.1287/inte.1120.0626

Broadie, M., & Ko, S. (2009). A simulation model to analyze the impact of distance and direction on golf scores. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, Austin, TX, USA. 3109-3120, https://doi.org/10.1109/WSC.2009.5429280

Clarke, S.R. (1988). Dynamic programming in one-day cricket – optimal scoring rates. *Journal of the Operational Research Society,* 39(4), 331-337. https://doi.org/10.2307/2582112

Clarke, S.R., Norman, J.M., & Stride, C.B. (2009). Criteria for a tournament: The world professional snooker championship. *Journal of the Operational Research Society*, 60(12), 1670-1673. https://doi.org/10.1057/jors2008.126.

Clayworth, P. (2013, September 5). *Billiards, snooker, pool and darts – Snooker and pool.* Te Ara – the Encyclopedia of New Zealand. Retrieved from http://www.TeAra.govt.nz/en/diagram/38477/table-layouts-snooker. Accessed November 17, 2021.

Chung, D.H.S., Griffiths, I.W., Legg, P.A, Parry, M.L., Morris, A., Chen, M., Griffiths, W., & Thomas, A. (2014). Systematic snooker skills test to analyze player performance. *International Journal of Sports Science & Coaching*, 9(5), 1083-1105. https://doi.org/10.1260/1747-9541.9.5.1083

Collingwood, J. A. P., Wright, M., & Brooks, R. J. (2020). *The analysis and development of performance measures in snooker.* [Unpublished manuscript]. Department of Management Science, Lancaster University.

Collingwood, J. A. P., Wright, M., & Brooks, R. J. (2021). *Simulating the progression of a professional snooker frame.* [Manuscript submitted for publication]. Department of Management Science, Lancaster University.

Collingwood, J. A. P., Wright, M., & Brooks, R. J. (2022). Evaluating the effectiveness of different player rating systems in predicting the results of professional snooker matches. *European Journal of Operational Research*, 296(3), 1025-1035. https://doi.org/10.1016/j.ejor.2021.04.056

Denman, H., Rea, N., & Kokaram, A. (2003). Contest-based analysis for video from snooker broadcasts. *Computer Vision and Image Understanding*, 92(2003), 176-195. https://doi.org/10.1016/j.cviu.2003.06.005

Dixon, M.J., & Coles, S.G. (1997). Modelling association football scores and inefficiencies in the football betting market. *Applied Statistics*, 46(2), 265-280. https://doi.org/10.1111/1467-9876.00065.

Eastaway, R., & Haigh, J. (2011). *The Hidden Mathematics of Sport*. London: Portico.

Eckhardt, R. (1987). Stan Ulam, John von Neumann, and the Monte Carlo method. *Los Alamos Science* (15) 131–137. http://library.lanl.gov/cgi-bin/getfile?15-13.pdf

Elo, A.E. (2008). *The Rating of Chess Players, Past and Present*. (2nd ed. – first published in 1978). Bronx NY 10453: ISHI Press International.

Everton, C. (2014). *Snooker and Billiards*. (2nd ed. – first published in 1991). Marlborough, UK: The Crowood Press Ltd.

Fitzpatrick, A., Stone, J.A., Choppin, S., & Kelley, J. (2019a). A simple new method for identifying performance characteristics associated with success in elite tennis. *International Journal of Sports Science & Coaching*, 14(1), 43-50. https://doi.org/10.1177/1747954118809089

Fitzpatrick, A., Stone, J.A., Choppin, S., & Kelley, J. (2019b). Important performance characteristics in elite clay and grass court tennis match-play. *International Journal of Performance Analysis in Sport*, 19(6), 942-952. https://doi.org/10.1080/24748668.2019.1685804

Florax, R. (n.d.). *Highest Match-Win Percentage – All-time – Professional*. Retrieved from https://cuetracker.net/statistics/matches-and-frames/highest-match-win-percentage/all-time?minimum_played=50. Accessed November 13, 2021.

Florax, R. (2019a). *Tournaments in 2019*. Retrieved from http://cuetracker.net/Tournaments/2019. Accessed November 13, 2021.

Florax, R. (2019b). *2019 World Championship*. Retrieved from https://cuetracker.net/tournaments/world-championship/2019/2918. Accessed November 13, 2021.

Florax, R. (2019c). *Rankings – 2018-2019*. Retrieved from https://cuetracker.net/rankings/2018-2019. Accessed May 23, 2022.

Florax, R. (2019d). *Number of 50 plus breaks – Season 2018-2019 – Professional*. Retrieved from: https://cuetracker.net/statistics/points-and-breaks/number-of-50-plus-breaks/season/2018-2019?status=professional&categories=ranking,minor-ranking. Accessed May 24, 2022.

Florax, R. (2021). *Centuries Made – Season 2020-2021 – Professional*. Retrieved from: https://cuetracker.net/statistics/centuries/most-made/season/2020-2021?status=professional&categories=ranking,minor-ranking,non-ranking,league,invitational,tour-qualifier,6-reds. Accessed May 23, 2022.

Freeze, R.A. (1974). An analysis of baseball batting order by Monte Carlo simulation. *Operations Research,* 22(4), 728-735. http://www.jstor.org/stable/169949

Glickman, M.E. (1999). Parameter estimation in large dynamic paired comparison experiments. *Journal of Applied Statistics*, 48(3), 377-394. https://doi.org/10.1111/1467-9876.00159.

Glickman, M.E. (2013). *Example of the Glicko-2 system*. Retrieved from http://glicko.net/glicko/glicko2.pdf. Accessed November 17, 2021.

Glickman, M.E. (2016). *The Glicko system*. Retrieved from http://glicko.net/glicko/glicko.pdf. Accessed November 17, 2021.

Gneiting, T., & Raftery, A. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association*, 102(477), 359-378. https://www.jstor.org/stable/27639845

Guest, A. (n.d.). *Why Should I Use MySnookerStats?* Retrieved from: https://www.mysnookerstats.com/why-mysnookerstats/. Accessed November 17, 2021.

Guest, A. (2010). *What your positional sense says about you.* Retrieved from: https://www.mysnookerstats.com/what-your-positional-success-says-about-you/. Accessed November 17, 2021.

Hafez, S. (2021). *UK Snooker Championship 2021: Shaun Murphy out, John Higgins and Kyren Wilson through*. Retrieved from: https://www.bbc.co.uk/sport/snooker/59392405. Accessed December 14, 2021.

Haigh, J. (2009). Uses and limitations of mathematics in sport. *IMA Journal of Management Mathematics*, 20, 97-108. https://doi.org/10.1093/imaman/dpm024.

Hirotsu, H., & Bickel, E.J. (2019). Using a Markov decision process to model the value of the sacrifice bunt. *Journal of Quantitative Analysis in Sports*, 15(4), 327-344. https://doi.org/10.1515/jqas-2017-0092

Höferlin, M., Grundy, E., Borgo, R., Weiskopf, D., Chen, M., Griffiths, I.W., & Griffiths, W. (2010). Video visualization for snooker skill training. *Computer Graphics Forum*, 29, 1053–1062. https://doi.org/10.1111/j.1467-8659.2009.01670.x

Hughes, M.D., & Bartlett, R.M. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences*, 20(10), 739-754. https://doi.org/10.1080/026404102320675602

Hughes, M., Caudrelier, T., James N., Redwood-Brown, A., Donnelly, I., Kirkbride, A., & Duschesne, C. (2012). Moneyball and soccer – an analysis of the key performance indicators of elite male soccer players by position. *Journal of Human Sport & Exercise* 7(2) 402-412. https://doi.org/10.4100/jhse.2012.72.06

Jakeman, M. (2021). *Cricket is having its Moneyball moment*. Retrieved from https://www.wired.co.uk/article/cricviz-twenty20-cricket-data. Accessed November 17, 2021.

Klaassen, F.J.G.M., & Magnus, J.R. (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association*, 96(454), 500-509. http://www.jstor.org/stable/2670288

Koning, R.H., Koolhaas, M., Renes, G., & Ridder, G. (2003). A simulation model for football championships. *European Journal of Operational Research*, 148(2), 268-276. https://doi.org/10.1016/S0377-2217(02)00683-5

Kovalchik, S.A. (2016). Searching for the GOAT of tennis win prediction. *Journal of Quantitative Analysis in Sports* 12(3), 127-138. https://doi.org/10.1515/jqas-2015-0059.

Kovalchik, S.A., & Reid, M. (2019). A calibration method with dynamic updates for within-match forecasting of wins in tennis. *International Journal of Forecasting* 35(2), 756-766. https://doi.org/10.1016/j.ijforecast.2017.11.008

Lames, M., & McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport*, 7(1), 62-79. https://doi.org/10.1080/24748668.2007.11868388

Langville, A.N., & Meyer, C.D. (2012). *The Science of Rating and Ranking: Who's #1*. Kindle Edition. Princeton: Princeton University Press.

Lewis, M., 2004, *Moneyball: The Art of Winning an Unfair Game*, Kindle Edition. New York, NY: WW Norton.

McGarry, T., & Schutz, R.W. (1997). Efficacy of traditional sport tournament structures. *Journal of the Operational Research Society*, 48, 65–74. https://doi.org/10.1057/palgrave.jors.2600330

McHale, I., & Forrest, D. (2005). The importance of recent scores in a forecasting model for professional golf tournaments. *IMA Journal of Management Mathematics*, 16, 131-140. https://doi.org/10.1093/imaman/dpi005.

McHale, I., & Morton, A. (2011). A Bradley-Terry type model for forecasting tennis match results, *International Journal of Forecasting*, 27 (2011), 619-630. https://doi.org/10.1016/j.ijforecast.2010.04.004

Méndez-Domínguez, C., Gómez-Ruano, M.A., Rúiz-Pérez L.M., & Travassos, B. (2019) Goals scored and received in 5vs4 GK game strategy are constrained by critical moment and situational variables in elite futsal, *Journal of Sports Sciences*, 37(21), 2443-2451, https://doi.org/10.1080/02640414.2019.1640567

Morris, B., & Bialik, C. (2015). *Serena Williams and the difference between all-time great and greatest of all time*. Retrieved from: https://fivethirtyeight.com/features/serena-williams-and-the-difference-between-all-time-great-and-greatest-of-all-time/. Accessed November 17, 2021.

Norman, J.M. (2015). Is the World Professional Snooker Championship fair? *Journal of the Operational Research Society*, 66(4), 705-706. https://doi.org/10.1057/jors2014.121.

O'Brien, J.D., & Gleeson, J.P. (2021). A complex networks approach to ranking professional snooker players. *Journal of Complex Networks*, 8(6), 1-16. https://doi.org/10.1093/comnet/cnab003

O'Donoghue, P. (2009). Interacting Performances Theory. *International Journal of Performance Analysis in Sport*, 9(1), 26-46. https://doi.org/10.1080/24748668.2009.11868462

O'Donoghue, P. (2015). *An Introduction to Performance Analysis of Sport*. London: Routledge.

Percy, D.F. (1994). Stochastic snooker. *Journal of the Royal Statistical Society, Series D (The Statistician)*, 43(4), 585-594. https://doi.org/10.2307/2348142

Percy, D.F. (2015). Strategy selection and outcome prediction in sport using dynamic learning for stochastic processes. *Journal of the Operational Research Society*, 66(11), 1840-1849. https://doi.org/10.1057/jors.2014.137

Perera, H., Davis, J., & Swartz, T.B. (2016). Optimal lineups in Twenty20 cricket, *Journal of Statistical Computation and Simulation*, 86(14), 2888-2900. https://doi.org/10.1080/00949655.2015.1136629

SABR (n.d.) *How to find raw data*. Retrieved from https://sabr.org/sabermetrics/data. Accessed November 17, 2021.

Scarf, P., Khare, A., & Alotaibi, N. (2021). On skill and chance in sport. *IMA Journal of Management Mathematics*, 33(1), 53-73. https://doi.org/10.1093/imaman/dpab026

Scarf, P., Parma, R., & McHale, I. (2019). On outcome uncertainty and scoring rates in sport: The case of international rugby union. *European Journal of Operational Research*, 273(2), 721-730. https://doi.org/10.1016/j.ejor.2018.08.021

Scarf, P., Yusof, M.M., & Bilbao, M. (2009). A numerical study of designs for sporting contests. *European Journal of Operational Research*, 198(1), 190-198. https://doi.org/10.1016/j.ejor.2008.07.029

Schell, R. (2011). SABR, Baseball Statistics and Computing. *The Baseball Research Journal*, 40(2), 44-50. https://sabr.org/journals/fall-2011-baseball-research-journal/.

Song, K., & Shi, J. (2020). A gamma process based in-play prediction model for National Basketball Association games. *European Journal of Operational Research*, 283(2), 706-713. https://doi.org/10.1016/j.ejor.2019.11.012

Spanias, D., & Knottenbelt, W.J. (2013). Predicting the outcomes of tennis matches using a low-level point model. *IMA Journal of Management Mathematics*, 24(3), 311-320. https://doi.org/10.1093/imaman/dps010

Stats Perform (n.d.). *Opta Event Definitions*. Retrieved from https://www.statsperform.com/opta-event-definitions/. Accessed November 17, 2021.

Stefani, R. (1998). Predicting outcomes. In Bennett, J. (Ed), *Statistics in Sport*. 249-276. London: Arnold Press.

Stefani, R. (2011). The methodology of officially recognized international sports rating systems. *Journal of Quantitative Analysis in Sports*, 7(4), 10. https://doi.org/10.1080/02664769723387

Swartz, T.B., Gill, P.S., & Muthukumarana, S. (2009). Modelling and simulation for one-day cricket. *The Canadian Journal of Statistics*, 37(2), 143-160. https://doi.org/10.1002/cjs.10017

Turner, H., & Firth, D. (2012). Bradley-Terry models in R: The BradleyTerry2 package. *Journal of Statistical Software*, 48(9), 1-12. https://doaj.org/article/a60d27125b574cba83443d54a17bcfbd

VanDerwerken, D., & Kenter, F. (2018). A generative Markov model for bowling scores. *Journal of Quantitative Analysis in Sports*, 14(4), 213-226. https://doi.org/10.1515/jqas-2017-0081

Vaziri, B., Dabadghao, S., Yih, Y., & Morin, T.L. (2018). Properties of sports ranking methods. *Journal of the Operational Research Society*, 69(5), 776-787. https://doi.org/10.1057/s41274-017-0266-8.

Welsh, J.C., Dewhurst, S.A., & Perry, J.L. (2018). Thinking aloud: An exploration of cognitions in professional snooker. *Psychology of Sport & Exercise*, 36, 197-208. https://doi.org/10.1016/j.psychsport.2018.03.003

World Professional Billiards and Snooker Association (2020a). *2020 WSF Open – Results*. Retrieved from https://snookerscores.net/tournament-manager/2020-wsf-open/results. Accessed November 17, 2021.

World Professional Billiards and Snooker Association (2020b). *2020 ROKiT Phones World Seniors Snooker Championship – Results*. Retrieved from https://snookerscores.net/tournament-manager/2020-rokit-phones-world-seniors-snooker-championship/results. Accessed November 13, 2021.

World Snooker Tour (n.d.). *Calendar 2021/2022*. Retrieved from: http://livescores.worldsnookerdata.com. Accessed November 13, 2021.

World Snooker Tour (2019). *Saudi Arabia to host World Snooker Tour event*. Retrieved from: https://wst.tv/saudi-arabia-to-host-world-snooker-tour-event/. Accessed November 17, 2021

WPBSA (2019). *Official rules of the games of snooker and English billiards*. Bristol: The World Professional Snooker & Billiards Association Limited.

Wright, M.B. (2009). 50 years of OR in sport. *Journal of the Operational Research Society*, 60(1), 161-168. https://doi.org/10.1057/jors.2008.170