Essays on Economics of Sports

Disertační práce

Jiří Lahvička

Školitel: doc. Ing. Zdeněk Chytil, CSc.

Praha, ČERVENEC 2014

Prohlášení

Prohlašuji, že jsem disertační práci s názvem *"Essays on Economics of Sports"* vypracoval samostatně a pouze na základě uvedených pramenů a literatury.

V Praze dne 20.07.2014 ……………………………………………

Jiří Lahvička

Statement

I hereby confirm that the dissertation *"Essays on Economics of Sports"* is my own work and that I have referenced all the sources used.

Prague, 2014.07.20 ……………………………………………

Jiří Lahvička

Acknowledgements

The author would like to thank his former dissertation supervisor Peter Bolcha, his current supervisor Zdeněk Chytil, Marek Hudík, Marek Vokoun, Miroslav Svoboda and anonymous journal referees for their helpful comments.

Abstract

This dissertation consists of five articles about economics of sports. The first three articles investigate various types of outcome uncertainty and how they relate to match attendance demand, while the remaining two articles test the efficiency of sports betting markets. The first article presents a new method of calculating match importance. Unlike the previous approaches in the literature, it does not require ex-post information and can be used for any type of season outcome. The second article shows that the additional playoff stage in the Czech ice hockey "Extraliga" lowers the probability of the strongest team becoming a champion and thus increases seasonal uncertainty. The third article demonstrates that the inconsistent findings in the literature about the link between match uncertainty and attendance could be explained by wrongly specified regressions, proposes a new approach to analyzing the effect of match uncertainty and shows that attendance demand is maximized if teams of the same quality play against each other. The fourth article examines the favorite-longshot bias in the context of betting on tennis matches. It shows that the favorite-longshot bias pattern is consistent with bookmakers protecting themselves against both better informed insiders and the general public exploiting new information. The fifth article investigates the supposedly profitable strategy of betting on soccer draws using the Fibonacci sequence. The strategy is tested both in a simulated market and on a real data set and found to lose money.

Keywords: uncertainty of outcome; match importance; match uncertainty; seasonal uncertainty; attendance demand; soccer; ice hockey; tennis; tournament design; Monte Carlo; sports betting; market efficiency; favorite-longshot bias; Fibonacci betting strategy

JEL classification: C53; D12; L83; G14

Contents

Foreword

The earliest articles dealing specifically with economics of sports date back to 1950s and 1960s (Rottenberg 1956; Neale 1964). Since then, economics of sports has developed into a thriving field with its own specialized journals, conferences, methods, and topics. The existing literature can be divided into two approaches; first, the application of standard economic/econometric tools to problems specific to sports industry, such as attendance demand and competition design; second, utilizing rich and detailed sports data to analyze more general issues, such as market efficiency or labor market discrimination. This dissertation consists of five articles straddling both of these approaches; the first three articles investigate various types of outcome uncertainty and how they relate to match attendance demand, while the remaining two articles test the efficiency of sports betting markets.

One of the biggest topics specific to economics of sports is the uncertainty of outcome hypothesis. This hypothesis was first formulated by Rottenberg (1956) and states that a tighter competition with a more uncertain outcome will attract more spectators. The subsequent literature has dealt with both properly defining the uncertainty of outcome and analyzing its impact on match attendance demand (together with other explanatory variables). The sports economic literature distinguishes three different types of uncertainty of outcome (Szymanski 2003, García and Rodríguez 2009); match uncertainty (how uncertain the result of one specific match is), seasonal uncertainty (how uncertain the competition winner and other similar outcomes are), and championship uncertainty (whether there is a long-run domination by one team); however, there are many alternative measures of each type of uncertainty and the impact on attendance demand is far from clear.

The first article called “Using Monte Carlo simulation to calculate match importance: The case of English Premier League” proposes a new method of calculating match importance (a typical variable used in attendance demand models to represent seasonal uncertainty). The previous approaches to defining this variable can be classified into the following five groups: first, using a dummy variable equal to one for all matches in the last several rounds of the season (Paul 2003); second, setting dummy variables equal to one if a team is mathematically certain to reach a given season outcome, such as winning the competition or being relegated (Baimbridge et al.1996; García and Rodríguez 2002; Feddersen et al. 2012); third, determining the dummy variable value through more complex rules based the current team positions, numbers of points, and the number of remaining matches (Baimbridge et al. 1996; Goddard and Asimakopoulos 2004; Simmons and Forrest 2006; Benz et al. 2009); fourth, defining the match importance as an interval variable calculated from the number of points that were eventually necessary to reach a given season outcome (such as winning the competition) and the number of remaining matches (Jennett 1984; Borland and Lye 1992; Dobson and Goddard 1992); fifth, defining the match importance for a specific outcome as the difference between the probability of reaching the outcome if the team wins the match and the probability of reaching the outcome if the team loses the match (Schilling 1994), where the probabilities are estimated by a Monte Carlo simulation (Scarf and Shi 2008; Goossens et al. 2012).

All these methods have apparent limitations; the approaches based on dummy variables are very crude, arbitrary, and (in case of mathematical certainty) overly conservative. The fourth approach pioneered by Jennett (1984) correctly treats match importance as an interval variable, but does not work as well for other types of season outcomes besides the championship and cannot be used for predictions since it requires ex-post information. The fifth approach first proposed by Schilling (1994) does not take into account how likely the team is to actually win or lose the match whose importance is being calculated or how likely it is to reach the seasonal outcome before the match is played. In addition, it can be used only for matches with two possible results (it ignores the probability of a draw). The problem with using a suboptimal method is that any impact of match importance on match attendance demand is underestimated.

The method proposed in the first article in this dissertation builds upon the approach of Schilling (1994). Match importance is defined as strength of relationship between the match result and a given season outcome (e.g. being relegated or not). To arrive at probabilities of various match result – season outcome combinations, probabilities of all remaining match results until the end of the season are estimated based on past results of all teams and then used to repeatedly simulate the rest of the season (the Monte Carlo method). Using actual results of 12 seasons of soccer matches (2000/01-2011/12) in the English Premier League and betting odds, it is shown that the proposed method of calculating match importance relies on realistic match and season outcome predictions. Unlike Schilling's approach, the method allows for more than two match result types. It produces results similar to those of Jennett (1984), but does not require any information unknowable before the match and can be easily adapted to any type of season outcome (not only championship, but also promotion, relegation and so on). The proposed method can also be used to calibrate other, less complex approaches (e.g. various versions of mathematical certainty).

The second article in this dissertation, "The impact of playoffs on seasonal uncertainty in the Czech ice hockey Extraliga," analyzes how a specific tournament design choice impacts seasonal uncertainty. The hypothesis that higher seasonal uncertainty increases attendance has a substantial empirical support (e.g. Szymanski 2003; Pawlowski and Budzinski 2013) and also seems to be accepted by competition organizers. The organizers have two ways of increasing the uncertainty of outcome; first, increase competitive balance (make the team strengths more equal) by redistributing resources through mechanisms such as TV and gate revenue sharing, payroll caps, or giving weaker teams earlier draft picks; second, increase seasonal uncertainty by modifying the tournament design. A major design choice in team sports is between using only a round-robin tournament (e.g. English Premier League) and combining the round-robin tournament with an additional playoff stage (e.g. US Major League Soccer).

The impact off the additional playoff stage on seasonal uncertainty has come into focus only recently. Fort and Quirk (1995) and Szemberg et al. (2012) noted that the regular season winner is far from certain to also win the playoff stage. However, the regular season winner is not necessarily the strongest team in the competition. Longley and Lacey (2012) used team payrolls as a proxy for actual team strengths and showed that the payrolls of NHL teams better predict results in the regular season than in the playoffs, indicating that the additional playoff stage increases seasonal uncertainty. However, this approach is limited by a small dataset and cannot be used to find out how exactly the additional playoff stage impacts the championship chances of specific teams or to analyze various alternative tournament designs.

The second article in this dissertation applies the Monte Carlo simulation method to analyze the impact of the additional playoff stage on seasonal uncertainty in the Czech ice hockey Extraliga.

The Monte Carlo simulation method has been previously used to compare various tournament designs by Scarf et al. (2009) and Scarf and Yusof (2011), but not to compare chances of various teams with or without the additional playoff stage. The Extraliga is a particularly good candidate for this analysis because the regular season uses a balanced schedule not favoring any specific team (each team plays exactly two home and two away games against every other team); this means that the additional playoff stage is not necessary to determine the champion and is simply a design choice.

In the Monte Carlo simulation method, six different sets of realistic team strengths are derived from the actual results of six Extraliga seasons. These strengths are then employed to repeatedly simulate all individual matches in a specific tournament design (the regular season only vs. the regular season followed by the playoffs) and the completed simulations are then used to estimate probabilities of each team winning the regular season and the playoff stage. The simulation results show that although the additional playoff stage heavily favors teams that placed better in the regular season and consists of quite a lot of games, it lowers the average probability of the strongest team becoming a champion from 48 to 39 percent and thus increases seasonal uncertainty. This is similar to the result of Longley and Lacey (2012) for the NHL, but the Monte Carlo simulation approach enables a deeper analysis; for example, it shows that the more dominant the strongest team is, the more their probability of winning the competition is decreased by the additional playoff stage; that the third-strongest to the sixth-strongest teams profit most from adding the playoffs; that obtaining the best seed (compared to the worst seed) roughly triples the championship probability; or that it does not generally make sense for a team to deliberately lose some regular season matches to avoid a specific team in the first round of the playoffs.

The higher seasonal uncertainty makes the Extraliga competition more attractive – the supporters of the strongest team cannot be so sure about the final outcome and the fans of weaker teams have a stronger hope of celebrating the championship title. The fact that securing a higher seed significantly increases championship chances makes the regular season finish interesting for fans of almost all teams. The higher seasonal uncertainty is also likely to translate into a more even distribution of all types of revenues and thus a higher competitive balance. In a positive feedback loop, this should further increase seasonal uncertainty.

The third article in this dissertation called “Does match uncertainty increase attendance? A non-regression approach” investigates whether more balanced sports matches attract higher attendances. While the uncertainty of outcome hypothesis is strongly supported for seasonal uncertainty, the empirical evidence for the impact of match uncertainty on attendance has been mixed. So far, the link between match uncertainty and attendance has been examined by regressing individual match attendance (or its logarithm) on variables representing qualities of both teams, other variables influencing attendance (ticket price, team rivalry, distance between teams, weather…), and a variable measuring how the match is balanced. These studies have investigated different sports, used different ways of measuring team quality (team ranks or points/goals per game) and match uncertainty (difference in team ranks or points per game; absolute value of betting spread; quadratic specification of home win probability derived from betting odds), and arrived at different results; some studies found that higher match uncertainty increases attendance, some found the opposite, some found that attendance increases with home win probability (and possibly starts decreasing if home win probability is higher than 0.6-0.7), some found no significant effect (Borland and McDonald 2003; Buraimo and Simmons 2008; Buraimo and Simmons 2009; Benz et al. 2009; Coates and Humphreys 2011; Pawlowski and Anders 2012).

The third article in this dissertation makes two contributions. First, three simple simulated data sets with no impact of match uncertainty on attendance are used to show that many commonly used regression specifications produce different (and wrong) results about the link between match uncertainty and attendance. This could explain the inconsistent findings in the literature, especially if the actual impact of match uncertainty is weak or nonexistent. Second, a new approach to analyzing the effect of match uncertainty on attendance is proposed. Using data about nine seasons of the English Championship, the article shows that in a pair of matches where both home teams are slight favorites, a switch of the corresponding away teams would decrease the total attendance. On the other hand, if both home teams are underdogs or strong favorites, switching the away teams would increase the total attendance. However, the magnitude of such attendance changes is quite small (several percent). These results are consistent with the uncertainty of outcome hypothesis and suggest that attendance demand is a bell-shaped function of match balance that is maximized if teams of the same quality play against each other. One possible explanation of such a shape could be that there are two groups of potential spectators with different preferences; fans in the first group (seasonal ticket holders, hardcore fans) do not care about match uncertainty and attend all matches if they have free time and no better opportunities, while fans in the second group (occasional spectators) choose to attend only the most interesting matches with one criterion being a proper match balance.

The above results can be directly applied to tournament design; to increase the total attendance of a competition while keeping the number of home and away matches of each team constant, a higher proportion of matches should be played between evenly matched teams. This could be achieved by splitting teams into groups based on team quality instead of on region or by making the tournament design more similar to the Swiss system commonly used in chess. However, the potential attendance increase would likely be small.

The last two articles in this dissertation deal with the topic of sports betting market efficiency. For analyzing efficiency, sports betting markets are preferable over financial markets, since each asset (bet) has a clear value at a specific point in time (after the match) (Thaler and Ziemba 1988). Two main efficiency concepts used in economics of sports are strong efficiency (each bet has the same expected value) and weak efficiency (there is no systematically profitable betting strategy). The literature has mostly concentrated on explaining various biases violating strong efficiency and on trying (mostly unsuccessfully) to find a profitable betting strategy (Thaler and Ziemba 1988; Sauer 1998).

The fourth article in this dissertation called “What causes the favorite-longshot bias? Further evidence from tennis” tries to distinguish between competing explanations for the observation that bets on favorites usually have a higher expected value (lose less money) than bets on longshots. The literature offers three types of explanations for the so-called favorite-longshot bias (Snowberg and Wolfers 2010; Makropoulou and Markellos 2011; Rossi 2011). The first explanation claims that bettors are local risk-lovers and bookmakers take advantage by lowering the odds on longshots. According to the second explanation, bettors overestimate winning probabilities of longshots and bookmakers again take advantage of this psychological bias. The third explanation is based on information asymmetry; bookmakers could potentially lose a lot of money if they underestimate longshots and this mispricing is exploited by either better informed insiders or by the general public reacting faster than bookmakers to new information. Therefore, bookmakers offer lower odds on longshots to protect themselves against this type of loss.

The fourth article in this dissertation uses a data set of almost 45,000 professional single tennis matches to show that the favorite-longshot bias is stronger in later-round matches and in matches in high-profile tournaments, i.e. in matches that are likely to attract high betting volumes; on the other hand, the favorite-longshot bias is also more pronounced in matches between lower-ranked players, which are likely to exhibit low betting volumes. This pattern cannot be explained solely by people being local risk-lovers or overestimating chances of longshots; if all bettors had the same preferences or biases, the type of match should not matter at all. Even if the risk-loving preferences (or the corresponding bias) were exhibited only by occasional bettors, thus causing the stronger favorite-longshot bias in matches that are likely to attract high betting volumes, it would not explain why the bias is also more pronounced in matches between lower-ranked players.

The most plausible explanation of the results seems to be a combination of two information asymmetry approaches: Matches between lower-ranked players are harder to predict, since public information is limited and private information about players' motivation or health problems could play a large role; therefore, it makes sense for the bookmaker to set lower odds on the longshot to minimize possible losses. On the other hand, private information should not play such a big role in later tournament rounds and high-profile tournaments, but in such matches the bookmaker faces a different kind of risk; the general public could react faster than the bookmaker to newly available information. Combined with a high volume of bets, this could mean a considerable loss, so the bookmaker again protects itself by setting lower odds on the longshot.

The last article in this dissertation called "The Fibonacci strategy revisited: Can you really make money by betting on soccer draws?" tests the strategy of betting on soccer draws using the Fibonacci sequence. The strategy was found profitable by Archontakis and Osborne (2007) and this was later confirmed on a bigger data set by Demir et al. (2012). Since the Fibonacci betting strategy is very simple to use, its profitability would mean that sports betting markets are not even weakly efficient (there are some other authors who claim to have found profitable strategies

– e.g. Kuypers 2000; Goddard and Asimakopoulos 2004; Vlastakis et al. 2009 – but these strategies usually rely on hard-to-implement models and identify only a small number of profitable betting opportunities).

The Fibonacci betting strategy is designed for betting on soccer results. It is based on the Fibonacci sequence (1, 1, 2, 3, 5, 8, 13…), where the first two numbers equal one and each successive number is the sum of the two previous numbers. The strategy works as follows: bet $1 (the first number in the sequence) on a draw, if losing, bet $1 (the second number) on a draw in the next match, if losing again, bet $2 (the third number) on a draw in the next match, and so on until a draw actually occurs; after that, start the whole sequence from beginning. Archontakis and Osborne (2007) proved that each sequence of bets ending in a draw is profitable if draw odds are always at least 2.618 (usually true). The authors also tested the Fibonacci strategy on 32 games in 2002 FIFA World Cup and found that it would have generated a profit. The strategy was also tested by Demir et al. (2012) on a sample of 32 seasons of top European soccer competitions and found profitable in all 32 cases. The authors also found the strategy to be profitable in a simple simulated strongly efficient market using 1,000 simulations.

The fifth article in this dissertation first investigates the behavior of the proposed strategy in a simulated strongly efficient market and shows that it actually is not and cannot be profitable in such a market. However, the strategy could still be profitable in a real market under the following two conditions: first, some bets on draws have positive expected values; second, the amounts bet on such matches are high enough to more than compensate for expected losses from the other bets. This could happen if bookmakers underestimated the probability of a draw after a long string of non-drawn matches. Therefore, various versions of the Fibonacci betting strategy are tested on a data set of almost 60,000 European soccer matches and also found to be losing money. The previous positive results in the literature were likely caused by a very low number of trials.

1 Using Monte Carlo simulation to calculate match importance: The case of English Premier League[1]

This article presents a new method of calculating match importance. Match importance is defined as strength of relationship between the match result and a given season outcome. Probabilities of all necessary match result – season outcome combinations are estimated by Monte Carlo simulation. Using actual results of 12 seasons of English Premier League and betting odds, it is shown that both match result and season outcome predictions are realistic. The method provides results that are close to Jennett's approach; however, it does not require ex-post information and can be used for any type of season outcome.

1.1 Introduction

In national team sports competitions, clubs typically strive to win the championship, get promoted to a higher league, qualify for other competitions, or avoid relegation. Because the competitions usually use a round-robin tournament system (sometimes combined with playoffs), not all matches have the same impact on the final outcome of the season – there are highly important last-round matches where a single result can decide which team becomes the new league champion, as well as unimportant matches between clubs without a realistic chance of being either promoted or relegated. It has already been established in the sport economics literature that important matches attract more spectators and that match importance influences

[1] A shorter version of this article (without the appendices) was published online ahead of print on May 22, 2013, in Journal of Sports Economics, doi: 10.1177/1527002513490172.

team performance;[2] however, there is little agreement on how to actually calculate the importance of a specific match. The problem with using a suboptimal method is that any impact of match importance is underestimated.

The goal of this article is to present a new method of calculating match importance. The starting premise is that the stronger the relationship between a match result and the season outcome (e.g. being relegated or not), the more important the match is. To arrive at probabilities of various match result – season outcome combinations, probabilities of all remaining match results until the end of the season are estimated based on past results of all teams and then used to repeatedly simulate the rest of the season (the Monte Carlo method). Predictions of both individual match results and season outcomes are then verified using actual results of 12 seasons of soccer matches (2000/01-2011/12) in the English Premier League and betting odds. Finally, the proposed method is used to evaluate other common approaches to calculating match importance.

1.2 Literature

There are two distinct components of match importance: first, how likely a team is to achieve a certain outcome, such as championship, promotion, or relegation (this is usually called seasonal uncertainty); second, how much a specific match can influence the probability of this outcome. The literature offers various approaches addressing one or both of these components differing in complexity and utilized information.

Probably the simplest possible method of including match importance as an explanatory variable in a regression model is to use a dummy variable that equals one for all matches in the last X rounds of the season (assuming that late matches tend to be more important). This method can be found, for example, in Paul (2003), who used a dummy variable for all NHL matches played in March and April.

[2] The literature about the link between match importance and attendance is summarized in García and Rodríguez (2009); the relationship between match importance and team performance was found by Goddard and Asimakopoulos (2004) and Feddersen et al. (2012).

The second alternative is to use the concept of mathematical certainty/impossibility – for example, when a team leads a competition by seven points, there are two rounds remaining and a win is worth three points, the team is mathematically certain to win the competition (there is not even a theoretical possibility of another outcome). For example, Baimbridge et al. (1996) defined dummy variables for the certain championship and certain relegation in the English Premier League, García and Rodríguez (2002) used a similar approach for Spanish soccer, and Feddersen et al. (2012) for multiple European soccer leagues.

The third approach is to use a more complex rule based on the current team positions, numbers of points, and the number of remaining matches; Baimbridge et al. (1996) and Simmons and Forrest (2006) used dummies for English soccer teams both being in the promotion zone or in the relegation zone; Goddard and Asimakopoulos (2004) asked if a team could be promoted or relegated if all other competing teams got one point in each of their remaining matches; Benz et al. (2009) employed a dummy variable equal to one if a German Bundesliga soccer team was no more than two points behind the current leader and there were at most six rounds until the end of the season.

The fourth approach, which treats match importance as an interval (rather than binary) variable, was introduced by Jennett (1984) for Scottish soccer and later used by others (Borland and Lye 1992; Dobson and Goddard 1992). Jennett's approach applied to the uncertainty of winning the championship works in this way: first, take the number of points that were eventually necessary to win the championship (of course, this ex-post information is not actually available before the end of the season, but it could be argued that it is possible to estimate it) – let's say it is 65. If it is still theoretically possible for a team to reach 65 points, set match importance for this team to 1/(number of matches necessary to reach 65 points), otherwise set match importance to zero. At the beginning of the season, all teams are able to reach 65 points, but they would need at least 22 matches (assuming 3 points for a win), so match importance equals 1/22. As the season progresses, importance for a specific team either increases towards 1 or drops to zero (when it is no longer possible to win). The match in which the eventual winner reaches 65 points must have the importance equal to one (this can happen in the last round or sooner).

The last approach, proposed by Schilling (1994), defines match importance for a specific outcome as the difference between the probability of reaching the outcome if the team wins the match and the probability of reaching the outcome if the team loses the match. To actually estimate these probabilities, Scarf and Shi (2008) and Goossens et al. (2012) use a Monte Carlo simulation approach similar to the first two steps of the method presented in this article. However, neither article offers any verification of season outcome predictions or comparison with other methods.

While this short overview of methods of calculating match importance is by no means exhaustive, other methods are usually quite similar or just combine elements of the approaches described above. More thorough discussion can be found in García and Rodríguez (2009).

All the methods described above have apparent limitations. Using a dummy variable for the last X rounds of the season is very crude (there are many last-round matches that do not decide anything) and does not distinguish between the importance for the home and away team. Using mathematical certainty/impossibility is too conservative; a team is expected not to win the championship much sooner than it is mathematically eliminated. When using the current team positions, numbers of points, and the number of remaining rounds, the chosen rule is necessarily arbitrary – as argued, for example, in Cairns (1987) and Peel and Thomas (1992) – and unlikely to work well in all possible cases. Jennett's method correctly treats match importance as a variable with more than just two possible values, but cannot be used for predictions (uses ex-post information) and cannot be easily adapted for other outcomes other than the championship.[3] Schilling's approach is the only one that takes into consideration strengths of the remaining opponents, but does not take into account how likely the team is to actually win or lose the match whose importance is being calculated or how likely it is to reach the outcome before the match is played. In addition, as noted by Goossens et al. (2012), it can be used only for matches with two possible results (it ignores the probability of a draw). All of the methods above (as implemented

[3] For example, if Jennett's method is used for European qualification, the key qualification matches for the eventual league winner will happen when its qualification is not really in doubt anymore. This problem would be compounded when using Jennett's method for relegation; however, Jennett (1984) proposed to modify the fraction denominator to the number of matches remaining until the end of the season, which partially solves the problem for this specific criterion.

in the literature) also disregard final table ranking criteria. In the rest of this article, it is shown that the proposed method of calculating match importance solves all these problems.

1.3 Data

To show how the presented method works and to verify that it gives realistic results, data about English Premier League are used; however, the method can be easily adapted to any other similar competition. The dataset consists of days/times and final results of all 4,560 matches played in 12 seasons (2000/01-2011/12).

In each season of Premier League, there are 20 teams playing two matches (one home and one away) against each other, so each team plays 38 matches per season. Winning a match is worth three points, drawing a match one point, and losing a match zero points. English Premier League does not use the head-on matches criterion to rank teams with the same number of points, so the final table ranking criteria are total points, total goal difference, and total goals scored (in this order). The first team becomes the champion and the last three teams are relegated to a lower competition – therefore, winning championship and not being relegated are two primary goals that enter into match importance calculations. Other possible goals could be qualifying for a European competition (usually first five teams) or just placing as well as possible.

To verify predictions of individual match results, the latest available betting odds of a big British bookmaker William Hill are used – these odds have been obtained for all but 16 matches played since December 26th, 2005; altogether for 2,477 matches. The odds have been converted to implied probabilities of a home win, draw, and away win.[4] All data including betting odds have been exported from a sports database Trefík.[5]

1.4 Method

As stated above, the importance of a specific match for a specific team can be decomposed into two components – first, the probability that the team reaches a particular season outcome, such as championship or not being relegated; second, how this probability depends on the match result. More formally, it is necessary to calculate the probabilities of various outcomes conditional on the specific match result. The match importance can then be expressed as a measure of association between the match result and the season outcome.

Multiple relevant outcomes imply multiple types of match importance; a particular match can have a small influence on the probability that a team wins the competition, a large influence on the final rank, and a zero influence on the probability that the team is relegated. These various types of match importance are likely to be valued differently by teams and potential match spectators.

The proposed method consists of three steps: first, calculate the result probabilities of all the remaining matches of the season; second, use these probabilities to estimate the probabilities of match result – season outcome combinations; third, calculate the strength of the association between the match result and various season outcomes.

[4] For example, the odds for the Liverpool – Chelsea match played on May 8th, 2012, were 2.10 (home win) – 3.30 (draw) – 3.50 (away win). This traditional form of betting odds indicates what multiple of the original sum the bettor gets if the result actually happens. To convert it to probabilities, the numbers are first inverted: 0.476 – 0.303 – 0.286. The new numbers add up to more than one – in this case 1.065 – to keep the betting agency profitable, so it is necessary to divide them by their sum to get the final probabilities 0.447 – 0.284 – 0.269. This common method is described, for example, in Benz et al. (2009).

[5] The database is available at www.trefik.com. betting The data were exported on June 7th, 2012 and also selectively cross-checked against the website Soccerway (www.soccerway.com).

1.4.1 Step 1: Estimating probabilities of match results

In the first step, it is necessary to estimate probabilities of results of all matches until the end of the season. The selected method should be simple to implement, computationally fast, able to predict exact scores (necessary to calculate final league tables in step 2), and at the same time provide estimates that are "good enough" (i.e. further improvements should have no significant impact on match importance values calculated in step 3).

There are two basic approaches to predicting match results – using betting odds or using the past results of both teams. Although betting odds exhibit small but systematic biases,[6] they can reflect all available information not necessarily included in the past results, such as injuries and suspensions of key players (Peel and Thomas 1992), so should be the superior option. However, the availability of betting odds (especially on exact match scores) is limited and they are usually not available at all for matches further in the future.[7] Therefore, the only feasible choice for simulating the whole season is to use the past results.

To calculate result probabilities of a specific match, this article uses a simplified version of a method commonly employed in the sports betting literature and described, for example, in Maher (1982) or Dixon and Coles (1997). For a given match, the home team's average score in the last 19 home matches (one rolling season, so approximately the last 12 months) and the away team's average score in the last 19 away matches are calculated.[8] The number of goals scored by the

[6] For example, Cain et al. (2000) analyzed betting odds quoted by William Hill for UK soccer matches and found that bets on favorites (as opposed to longshots) had a higher expected value.

[7] Even most commercial databases, such as trefik.com (used in this article), txodds.com, or archived BetFair odds, go back to early 2000s at most and errors and missing data are not unusual. Betting odds on exact match scores are available only for selected competitions (due to lower interest of bettors). An inspection of bookmakers' websites reveals that betting odds on typical league matches are available only about two weeks and/or one round into the future. It could be argued that for research on historical data, future betting odds are not necessary (since everything is already in the past), but this would negate a key advantage of the proposed method – not relying on ex post information that could not be available before the analyzed match to anyone.

[8] To make sure there are always at least 19 previous home/away matches available, the first season in the dataset is used only to provide match history and the match importance values are computed from the second season onwards. However, this does not help in case of freshly promoted teams who have no match history at all. Match histories of such teams are initialized with 19 home matches with the score 1.266:1.378 and 19 away matches with the score

home team and the number of goals scored by the away team in the analyzed match are assumed to be two independent[9] Poisson-distributed variables with the following expected values:

$$\lambda_{\text{home}} = \frac{\text{average goals scored by home team} + \text{average goals conceded by away team}}{2}$$

$$\lambda_{\text{away}} = \frac{\text{average goals conceded by home team} + \text{average goals scored by away team}}{2}$$

Based on the expected values, it is possible to construct corresponding probability distributions of goals scored by each team, compute joint probabilities of all possible match scores, and subsequently also the probabilities of a home win, draw, and away win. For example, to estimate probabilities for Manchester City – Manchester United match (round 36, season 2011/12), the average scores of Manchester City in their previous 19 home matches and of Manchester United in their previous 19 away matches are calculated. These scores are 2.895:0.526 and 1.947:0.789 respectively, so Manchester City are expected on average to score (2.895 + 0.789)/2 = 1.842 goals and Manchester United are expected to score (0.526 + 1.947)/2 = 1.237 goals. Using Poisson distributions, the probability of Manchester City scoring 1 goal is 0.292 and the probability of Manchester United scoring 0 goals is 0.290, so the probability of a 1:0 final score (the actual match result) is simply the probability that Manchester City score 1 goal times the probability that Manchester United score 0 goals and equals 0.292 * 0.290 = 0.085. The probability of Manchester City winning the match can be estimated by aggregating probabilities of all winning final scores as 0.516 (similarly, the draw probability is 0.228 and the Manchester United win probability is 0.256).

0.923:1.856 (these scores are based on the average scores in their home/away matches of all freshly promoted teams in the dataset; the exact values are not that important, since towards the end of the season, when most important matches take place, most of these artificial scores are already replaced by real results). A similar solution could be used for freshly relegated teams in lower competitions or if the extra season is not available.

[9] In the Premier League dataset, the independence assumption seems close to the reality; the correlation between goals scored by home and away teams is close to zero (-0.050 for the whole dataset and +0.048 for the dataset restricted to matches where no team scored more than one goal).

Assuming that team strengths do not change much over the course of the season, it is also possible to use the match histories of all teams available before the analyzed match to estimate probabilities for all the other matches remaining until the end of the season (necessary for the next step).[10]

1.4.2 Step 2: Estimating probabilities of final ranks

In the second step, the actual match results up to the analyzed match are combined with estimated probabilities for all the matches remaining until the end of the season to find out the probabilities of final team ranks conditional on the analyzed match result.[11] For the round 36 Manchester City – Manchester United match mentioned above, it means combining actual results of rounds 1-35 with estimated probabilities for all matches played in rounds 36-38 and calculating probabilities that Manchester City or Manchester United ultimately finish first, second etc. given any specific result of their mutual match. In most cases, the final rank probabilities cannot be found analytically – the number of all possible scenarios is simply too high.[12] However, it is possible to

[10] There are several possible modifications of the algorithm used in this section, such as changing the length of the period used to calculate average scores, giving more weight to recent results, or using a bivariate Poisson distribution; these and other changes are discussed in Dixon and Coles (1997) or Goddard (2005). The period of one year used in this article was chosen based on Goddard (2005), who showed that English soccer results at most 12 months old are several times more important than results 12-24 months old. However, modifying the length of this period has a negligible impact on match importance values calculated in the third step – for example, the Spearman rank correlation between match importance values using a one-year period vs. two-year period is 0.98-0.99.

[11] In this article, matches played on the same day as the match whose importance is being calculated are considered to take place in the future (so their results are not known). This lag can be modified depending on the intended use of the model – for example, when studying sports attendance demand, people could decide to attend a match several days in advance, so even match results several days in the past might be considered unknown.

[12] For the Manchester City – Manchester United match, it would be necessary to go through all possible combinations of results of 25 matches (3 remaining rounds x 10 matches per round, but 5 round 36 matches were already played in the previous two days). Even when taking into account only the win/draw/loss and not the exact score, this gives $3^{25} \approx 8.5 * 10^{11}$ scenarios. The number of scenarios gets intractable for any match not in the last several rounds of the season.

choose a random sample of all scenarios (the Monte Carlo method) to estimate these probabilities with any desired level of precision.[13]

For any analyzed match, the Monte Carlo simulation starts with a league table based on all matches played so far (in the Manchester City – Manchester united example, it is the league table just before the match). In each simulation, the estimated probabilities derived in the first step are used to randomly choose results (exact scores) of all the remaining matches (including the match whose importance is being calculated). For example, the 1:0 score in the Manchester City – Manchester United match has an estimated probability of 0.085, so it has exactly this probability to be chosen.[14] After that, the final league table is calculated (using any ranking criteria applicable). The important outputs of each simulation are the analyzed match result and the final ranks of both teams. The simulation is run as many times as necessary to reach the desired precision.

After all the simulations are complete, the analyzed match results are categorized into a win, a draw, and a loss,[15] and a contingency table is constructed for each team where one variable is the match result and the other variable is the final team rank.

[13] A similar Monte-Carlo-based approach is already used by, among others, the website SportsClubStats.com to predict season outcomes. However, this website does not provide much documentation and no verification of accuracy of their predictions (as of January 6th, 2013).

[14] This can be done by generating two random numbers (goals scored by home and away teams) for each match – the first number from a Poisson distribution with expected value λ_{home} and the second number from a Poisson distribution with expected value λ_{away}. Of course, these numbers (and thus match results) are likely to be different in each simulation. Another option would be to precalculate probabilities of all possible scores in each match and randomly choose among them.

[15] This step transforms the result into an ordinal variable and simplifies the following computation. Another alternative would be a score difference.

		Final rank		
		1st	2nd	Sum
Match result	Win	3,304,376	1,854,511	5,158,887
	Draw	427,562	1,850,343	2,277,905
	Loss	36,049	2,527,159	2,563,208
	Sum	3,767,987	6,232,013	10,000,000

Table 1: Example contingency table for Manchester City

Table 1 shows the contingency table for Manchester City before the Manchester City – Manchester United match. The table is based on 10,000,000 simulations and is restricted to only non-zero values (before the match, Manchester City was in the second place, trailing three points behind Manchester United, but having a better score, and there was no chance of Manchester City finishing worse than second). From the contingency table, it is possible to calculate relative frequencies of various match results, final ranks, and their combinations and use them to estimate the true probabilities that could (theoretically) be found out analytically by going through all possible scenarios.[16] For example, Manchester City had 51.6% probability of winning the match, but just 37.7% probability of winning the league. The probability of winning the league was heavily dependent on the match result – it was 3,304,376/5,158,887 ≈ 64.1% in case of winning the match, but just 36,049/2,563,208 ≈ 1.4% in case of losing (after winning the match, Manchester City later became a new champion). This match should be identified as highly important for Manchester City in terms of winning the championship; however, it had no impact on the probability of relegation, which was zero in any case.

It could be argued that the information in the contingency table might be derived from betting odds on the specific match and on the season outcome. However, even when these betting odds are available, they provide only match and season outcome probabilities (row and column sums in the contingency table – for example, Manchester City defeating Manchester United, or Manchester City winning the championship), not how the season outcome probabilities depend

[16] The standard error of a probability estimate equals $\sqrt{[\pi(1-\pi)/n]}$, where π is the actual probability and n is the number of simulations, so by increasing the number of simulations the estimate precision increases as well. The maximum practical value of n is limited by available computing power and time. Using an optimized MATLAB/Octave implementation, 1,000,000 simulations for each match in one season take one to two hours running in a single thread on a desktop Intel i5/i7 Sandy/Ivy Bridge CPU.

on the match result (individual cells in the contingency table – for example, the probability Manchester City wins the championship if it defeats Manchester United), so there would be not enough information to calculate match importance in the next step.[17]

1.4.3 Step 3: Calculating result-outcome association

In the third step, the contingency table is used to calculate the importance of a particular match from the point of view of a specific team using a specific desired season outcome. As defined above, the importance is expressed as a measure of association between the match result and the final season outcome (i.e. how strongly is the season outcome influenced by the match result).

If the analyzed outcome is anything else than the final rank, the contingency table must first be transformed by aggregating appropriate columns – for example, if the analyzed outcome is relegation, final ranks 1 to 17 are aggregated into the first column (not being relegated) and final ranks 18 to 20 are aggregated into the second column (being relegated). Consequently, the transformed contingency table has 3 rows (win/draw/loss) and 2 columns (no relegation/relegation). Similarly, to analyze match impact on winning the championship, final rank 1 becomes the first column and final ranks 2 to 20 are aggregated into the second column. To get a positive association value, rows and columns should be ordered so that the best result – best outcome combination is in the top left corner.

To make the results comparable to other methods, the chosen measure of association should be on the scale from zero (no importance) to one (maximum possible importance). There are two obvious extreme cases – first, the outcome probability is already zero or one and therefore does not depend on the match result at all; second, the outcome hangs exactly in the balance (e.g. the probability of being relegated is 1/2) and is solely determined by the match result. The chosen measure of association should be equal to zero in the first case and one in the second case. The measure should also take into account that both variables (result and outcome) are ordinal.

[17] Betting odds on season outcomes are not universally available even for major competitions – going through websites of major bookmakers William Hill, BetFair and Bwin on January 6th, 2013, betting odds on the following outcomes were not offered by any of them: relegation in French Ligue 1 (soccer), qualifying for a European competition in the German Bundesliga (soccer), or qualifying for the play-off stage in KHL (ice hockey).

Kendall-Stuart tau-c is a measure of association between two ordinal variables that fulfills all these conditions and is therefore used throughout this article.[18]

		Relegation	
		No	Yes
Match result	Win	0.5	0
	Draw	0.4	0
	Loss	0.1	0
tau-c = 0			

		Relegation	
		No	Yes
Match result	Win	0.5	0
	Draw	0.4	0
	Loss	0	0.1
tau-c = 0.36			

		Relegation	
		No	Yes
Match result	Win	0.5	0
	Draw	0	0.4
	Loss	0	0.1
tau-c = 1			

Table 2: Examples of tau-c values

Table 2 shows tau-c values for three example contingency tables for a relegation outcome (the numbers already represent probabilities of each result-outcome combination). In the left panel, the outcome (no relegation) is already determined before the match and the match importance correspondingly equals the lowest possible value of zero. In the middle panel, the outcome is completely determined by the match result, however, the before-the-match probability of relegation is just 0.1. Therefore, the match importance indicated by the tau-c value of 0.36 is quite high, but far from the maximum possible.[19] Finally, the right panel represents the ideal case of maximum match importance; the outcome is completely determined by the match result, the before-the-match probability of relegation is exactly 0.5, and tau-c reaches its maximum possible value of one. For the Manchester City – Manchester United example (Table 1), the tau-c value is

[18] Tau-c is computed in the following way: first, create all possible pairs of observations (if the contingency table contains n simulations, there are n(n-1)/2 possible pairs); second, divide the pairs of observations into three groups: C (concordant pairs, where both the match result and the season outcome are better in one of the observations and worse in the other), D (discordant pairs, where the match result is better and the season outcome is worse in one of the observations and the other way round in the other), and others (either the match result or the outcome is the same in both observations); the higher the number of concordant pairs and the lower the number of discordant pairs, the higher the match importance. Tau-c is the difference between the numbers of concordant and discordant pairs divided by the highest possible number of concordant (or discordant) pairs given the table size and is given by the following equation, where n is the number of observations and m is the smaller table dimension: tau-c = $(C-D)/[n^2(m-1)/2m]$.

[19] This example illustrates one of the problems with Schilling's method of calculating match importance as a difference between outcome probabilities in case of a win and in case of a loss. This difference is the same (equal to 1) in both the middle and right panels of Table 2, while the situation before the match is clearly very different.

0.585, correctly indicating a very high importance of this specific match for Manchester City's championship chances.[20] More example contingency tables with their corresponding match importance values can be found in Appendix A: Further example contingency tables.

1.5 Verification

Since the match importance value is based on individual match and season outcome predictions, this section verifies that both types of predictions are realistic. These verifications, as well as the rest of this article, are based on 10,000,000 simulations for each of 4,180 matches of 11 seasons of English Premier League.[21] The number of simulations is high enough so that any potential biases in estimated probabilities would be due to imperfect assumptions and not to the Monte Carlo method itself.[22]

1.5.1 Verifying match predictions

To show that the presented method is based on realistic individual match predictions, this subsection analyzes probability estimates of a home win, draw, and away win based on match histories of both teams available just before each match.

The first test looks at all 4,180 matches and simply compares average estimated probabilities with actual relative frequencies of each type of result. The average estimated probability of a home win is 0.467, which is almost identical to the actual relative frequency of 0.468 (the home team

[20] Match importance values obtained from the Monte Carlo simulation are just estimates of their true values, since they are based on estimated probabilities of various result-outcome combinations. Computing standard errors of tau-c estimates is not simple; however, the worst-case match importance standard error for a 3x2 contingency table can be roughly approximated by $1/\sqrt{n}$, where n is the number of simulation runs. This estimate is based on a simple Monte Carlo simulation for a 3x2 contingency table with all outcome-result probabilities equal to 1/6. The simulated standard error was about 10% higher than $1/\sqrt{n}$ for n between 1,000 and 10,000,000 and decreased slowly when deviating from the original contingency table probabilities.

[21] The first season in the dataset (2000/01) provides match history to estimate match result probabilities.

[22] The worst-case standard error of various outcome probability estimates is $1.58 * 10^{-4}$ and the worst-case standard error of match importance estimates can be approximated by $3.16 * 10^{-4}$.

won in 1,958 out of 4,180 matches). Similarly, the average estimated draw probability is 0.250, which is very close to the actual relative frequency of 0.259, and the average estimated away win probability of 0.283 is very close to the actual relative frequency of 0.273.[23] Therefore, it can be concluded that the estimates seem to be unbiased, i.e. no type of result is predicted more often than it should be.

The second test compares the Poisson-based probability estimates based on the up-date-histories (Up-to-date Poisson) with two benchmarks – home win, draw, and away win probabilities implied by the latest available betting odds (Up-to-date betting odds) and a naïve algorithm always predicting 0.468 probability of a home win, 0.259 probability of a draw, and 0.273 probability of an away win (Naïve; the probabilities equal actual relative frequencies in the dataset). Despite its name, the naïve algorithm is still more realistic than typical assumptions used by other methods of calculating match importance, such as (modified) mathematical certainty. To analyze the quality of predictions of matches further in the future (which are necessary to arrive at season outcome predictions), the test also includes Poisson-based probability estimates using only historical matches that are at least 180 days old (Lagged Poisson).

To compare predictive power of these four methods, this article uses a pseudo-likelihood statistic employed by Rue and Salvesen (2000) and later by Goddard (2005). The statistic is equal to the geometric mean of predicted probabilities of actually observed results and can range from 0 to 1, where a higher value corresponds to a higher predictive power.[24] The test is restricted to the subset of 2,477 matches with available betting odds.

[23] None of the differences is statistically significant at $\alpha = 0.05$.

[24] The statistic is equal to 0 if at least one actually observed result is predicted impossible and equal to 1 if every observed result is predicted with certainty. However, if sports results are assumed to be stochastic (not predictable with certainty), the maximum value is much lower than 1 (depending on the proportion of one-sided matches).

		Pseudo-likelihood			
Season	Number of matches	Up-to-date betting odds	Up-to-date Poisson	Lagged Poisson	Naïve
2005/06	213	0.3952	0.3823	0.3773	0.3584
2006/07	378	0.3786	0.3672	0.3680	0.3490
2007/08	380	0.3990	0.3733	0.3757	0.3454
2008/09	370	0.3851	0.3749	0.3765	0.3445
2009/10	379	0.3906	0.3847	0.3763	0.3548
2010/11	378	0.3643	0.3589	0.3626	0.3475
2011/12	379	0.3760	0.3635	0.3620	0.3421
All	2477	0.3832	0.3713	0.3707	0.3481

Table 3: Comparison of match predictions

Table 3 summarizes the test results for individual seasons and for all matches together. There are two key conclusions – first, even the very simple Poisson-based algorithm clearly outperforms the naïve predictions and actually provides estimates not substantially worse than betting odds;[25] second, predictions of matches half a year in the future have practically the same quality as predictions based on the latest available information. The second conclusion validates the assumption that team strengths do not change much over the course of the season.

1.5.2 Verifying season predictions

Even though the Monte Carlo simulation relies on realistic individual match predictions, any imperfections in these predictions could potentially accumulate (or cancel) when estimating probabilities of season outcomes. Therefore, this subsection compares the season outcome probabilities with the outcomes that actually happened.

Again, data about 11 seasons and 4,180 matches are used. Before each match, there are two sets (one for each team) of estimated probabilities of 20 possible final ranks (altogether 4,180 * 2 * 20

[25] It should not be surprising that betting odds (when available) are still a little better, since they include much more information; even a significantly more complex algorithm in Rue and Salvesen (2000) provided predictions only on par with betting odds for one season of English Premier League. Similarly, Goddard (2005) found only very small differences in prediction quality among algorithms of varying complexity based on historical results.

= 167,200 individual rank probabilities). Adding together the individual probabilities for ranks 1-10 in each set provides the probability that the team finishes in the top half of the table; there are 4,180 * 2 = 8,360 such predictions. These probabilities are very suitable for testing since they fully cover all values from 0 (impossible outcome) to 1 (certain outcome) and exactly one half of these predictions come true.

First, it is interesting to look at outcomes predicted to be practically impossible or certain. Out of 8,360 estimates, there are 345 outcomes predicted to be practically impossible (not occurring in any of 10,000,000 simulations) and none of these outcomes happened in reality. Similarly, there are 704 outcomes predicted to be practically certain (occurring in all of 10,000,000 simulations), all of them later actually happening.[26] Analyzing championship, relegation, or individual rank probabilities leads to analogical results. Therefore, it can be concluded that if the Monte Carlo simulation estimates an outcome to be practically impossible or certain, it must be at least very unlikely or very likely to happen. Of course, the same could be said about the mathematical certainty method; however, the Monte Carlo simulation identifies impossible or certain outcomes much earlier in the season. For example, if a team is ultimately not relegated, it is mathematically certain not to be relegated on average 6 rounds before the end of the season. Based on the Monte Carlo simulation, the relegation is practically impossible ($p = 0$) on average 10 rounds before the end of the season and very unlikely ($p < 0.001$) on average 18 rounds before the end of the season.

Another important property to check is the unbiasedness of predictions, i.e. that outcomes with an estimated probability of X% actually happen in about X% of all cases. For this test, the original 8,360 estimated probabilities of a team finishing in the top half of the table are ordered from the lowest to the highest. The already analyzed and correct 345 impossible and 704 certain predictions are removed and the remaining 7,311 estimates are split into 6 equal-sized bins with the first bin containing the lowest 1/6 of the estimated probabilities. In each bin, the average estimated probability should be close to the relative frequency of actual outcomes.

[26] The lowest estimated probability for the outcome that actually happened was 2.1%. The highest estimated probability for the outcome that actually did not happen was 97.3%.

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6
Number of predictions	1,219	1,218	1,219	1,218	1,219	1,218
Average estimated probability	0.0146	0.1300	0.3052	0.5560	0.8590	0.9951
Actual relative frequency	0.0066	0.1552	0.2847	0.5246	0.8819	1.0000
Difference (estimated – actual)	+0.0080	-0.0252	+0.0206	+0.0313	-0.0229	-0.0049

Table 4: Binned predicted vs. actual frequencies of outcome

Table 4 shows that the estimated probabilities for each bin are within two to three percentage points of the actual relative frequencies, so there is no substantial bias associated with a particular probability range.[27] The same comparison, but with 50 bins instead of 6, can be found in Appendix B: Further verifying season predictions.

1.6 Comparison with other methods

As shown in the previous section, the presented method provides match importance values that are derived from realistic predictions of both match results and season outcomes. Unlike other methods, it also takes into account factors such as strengths of the remaining opponents and final table ranking criteria. This section looks more closely at the computed match importance values using championship and relegation criteria and compares them against numbers provided by other common approaches.

Again, data about 11 seasons and 4,180 matches are used. For each match, there are four associated match importance values – one championship and one relegation importance value for each team. Altogether, there are 8,360 championship importance values and 8,360 relegation importance values.

[27] No difference between the estimated probability and the actual relative frequency is statistically significant at $\alpha = 0.05$. When calculating standard errors, it is necessary to take into account that there are only 220 independent random trials (20 teams achieving some final rank in each of 11 seasons), so each bin can contain predictions related to at most this number of trials.

	Average	Standard deviation	Skewness	Min	Percentiles			Max
					90^{th}	99^{th}	99.9^{th}	
Championship	0.0164	0.0452	4.8001	0.0000	0.0693	0.1851	0.4742	0.8253
Relegation	0.0539	0.0747	2.4199	0.0000	0.1509	0.3008	0.6353	0.8761

Table 5: Descriptive statistics of championship and relegation importance, N = 8,360

Table 5 provides basic descriptive statistics. Both championship and relegation importance distributions are extremely skewed towards low values; there are simply not that many important matches. On average, relegation importance is higher, since there are more relegation than championship spots and typically more teams in contention for avoiding relegation than for winning championship. Maximum observed values are close to one in both cases, indicating matches deciding the outcome hanging in the balance.

To simplify the following analysis, all match importance values are classified into the five following groups: zero importance (importance $\leq 10^{-6}$; this value indicates that the season outcome probability was extremely close to zero or one), very low importance (10^{-6} < importance ≤ 0.01), low importance (0.01 < importance ≤ 0.1), medium importance (0.1 < importance ≤ 0.2), and high importance (importance > 0.2).

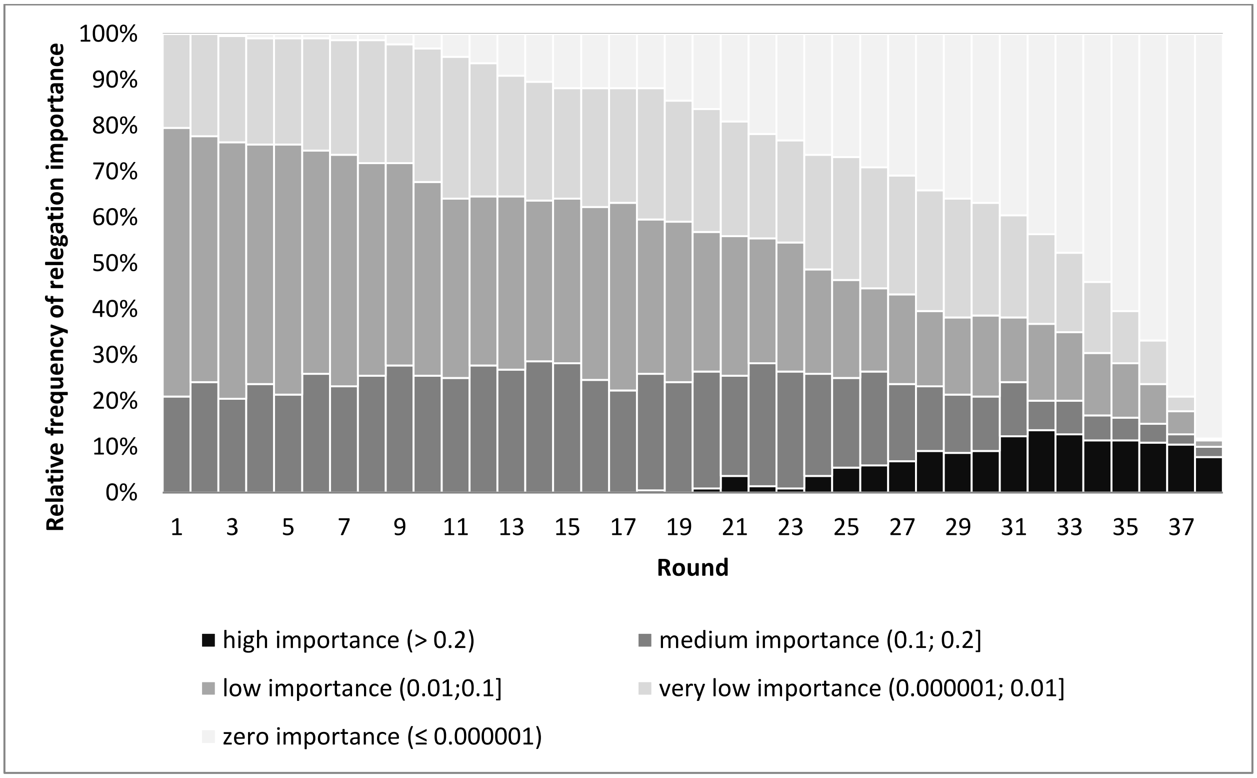

Figure 1: Relative frequencies of relegation importance values in different competition rounds

Figure 1, which shows relative frequencies of grouped relegation importance values in different competition rounds, immediately demonstrates problems of the simplest match importance method of assuming that all matches in the last X rounds are important; first, there is no obvious cutoff round where important matches start to appear; second, most matches even in the last several rounds are simply not that important for any team.[28] Therefore, just using a dummy variable for all matches in the last X rounds severely underestimates any impact of match importance on any other variable.

Another match importance method (mathematical certainty) assumes that a match is important for a given team if a specific outcome is not yet mathematically certain or impossible. A modified approach (used, for example, in Goddard and Asimakopoulos 2004) assumes that in the worst-case scenario, a team will get only X points in each remaining match with all the other teams getting full three points (and placing better if having the same number of points), while in the

[28] Even when including the European qualification criterion, only about one half of all matches in the last four rounds have at least medium importance on at least one season outcome criterion for at least one team.

best-case scenario, a team will get full three points in each remaining match with all the other teams getting X points (and placing worse if having the same number of points). A team is sure to be relegated if it finishes no better than 18th in the best-case scenario; similarly, it cannot be relegated if it finishes no worse than 17th in the worst-case scenario. For X = 0, this method is exactly the same as mathematical certainty, while higher values of X eliminate low-importance matches more aggressively.

		Number of matches	Percentage of matches classified as important (modified mathematical certainty)					
			X=0	X=0.5	X=1	X=1.5	X=2	X=2.5
Relegation importance (Monte Carlo)	zero	1,945	47.7	39.3	30.6	19.6	8.7	1.0
	very low	1,898	100	98.6	96.4	88.6	73.4	39.6
	low	2,583	100	99.9	99.7	99.7	95.7	82.7
	medium	1,612	100	99.8	99.8	99.8	99.2	94.8
	high	322	100	99.4	99.4	98.8	97.5	82.6
	Sum	8,360						

Table 6: Classification of relegation importance values by Monte Carlo vs. modified certainty

Table 6 shows how both the Monte Carlo method and the modified mathematical certainty method with different values of X classify 8,360 relegation importance values. It is evident that the traditional mathematical certainty method (X = 0) is not nearly aggressive enough; 48% of zero-importance and full 100% of very-low-importance matches are still classified as important.

The best setting for the English Premier League dataset seems to be X = 2,[29] which is higher than commonly used in the literature; for example, Goddard and Asimakopoulos (2004) used X = 1. For X = 2, just several percent of medium- and high-importance matches are misclassified as not important, while only 9% of zero importance matches are misclassified as important. This could be considered a good result given the computational simplicity of the modified mathematical certainty method; however, importance is still treated as a binary variable, so a lot of information is necessarily lost.

Similarly to the Monte Carlo method, Jennett's method treats match importance as a variable with many possible values between zero and one. To better compare both methods, match

[29] This is also true for championship importance values.

importance values generated by Jennett's method are also classified into the five following groups matching the distribution of the Monte Carlo values as closely as possible: zero importance (importance = 0), very low importance (0 < importance ≤ 1/25), low importance (1/25 < importance ≤ 1/18), medium importance (1/18 < importance ≤ 1/5), and high importance (importance > 1/5).

		Championship importance (Jennett)					
		zero	very low	low	medium	high	Sum
Championship importance (Monte Carlo)	zero	3,361	85	25	0	12	3,483
	very low	830	1770	695	64	3	3,362
	low	78	375	300	193	10	956
	medium	25	74	140	243	13	495
	high	6	0	3	26	29	64
	Sum	4,300	2,304	1,163	526	67	8,360

Table 7: Classification of championship importance values by Monte Carlo vs. Jennett

Table 7 shows that both methods classify the championship importance values into the corresponding categories quite similarly – 52 of 67 values classified as highly important by Jennett's method have at least low importance assigned by the Monte Carlo method, 58 of 64 values classified as highly important by the Monte Carlo method are considered to have at least low importance by Jennett's method, and only 3.6% of value pairs are more than one category apart. The Spearman rank correlation between uncategorized values of championship importance is 0.781.

A manual inspection of matches where these two methods disagree reveals that there are three reasons Jennett's method occasionally fails: first, it uses ex-post information unknowable before the match (this leads to mistakenly assigning zero importance to ex-ante important matches);[30]

[30] A typical example is Sunderland – Manchester United match played on May 2nd, 2010, in the second-to-last round of the 2009/10 season. Before this match, Manchester United was in the second place trailing one point behind Chelsea. The championship importance of this match for Manchester United computed by the Monte Carlo algorithm is as high as 0.475 (quite reasonable given the situation as known at that point in time and considering that Chelsea had to play a difficult away match at Liverpool in the same round). However, Jennett's method sets the

second, it does not take into account who the opponent is (and therefore does not assign higher importance to matches against teams competing for the same outcome); third, it ignores any table ranking criteria besides total points.

Based on the data, Jennett's method is obviously the best alternative to the more computationally complex Monte Carlo method for championship importance if the season is already over and an occasional misclassification is not a big problem. However, Jennett's approach does not work as well for other criteria, such as relegation or qualifying for European competitions,[31] and cannot be used for prediction.

1.7 Conclusion

As shown throughout the article, the proposed method of calculating match importance relies on realistic match and season outcome predictions, allows for multiple match result types (unlike Schilling's approach), can be used to derive match importance values expressed as continuous variables for any team in any match given any desired outcome (such as championship, promotion, or avoiding relegation), and does not need any information unknowable before the match.

The presented method can be also used as a benchmark for other, less complex approaches. Using the Premier League dataset, Jennett's method of estimating championship importance is found to produce results that are quite close to the Monte Carlo simulation. However, Jennett's method requires ex-post information and does not work as well for other criteria besides championship. If ex-post information is not available and a simple dummy variable for match importance for each team is considered sufficient, the modified mathematical certainty approach provides a

championship importance to zero using ex-post information that Chelsea won the last two matches and Manchester United could not have caught up. Obviously, this could not have been known in advance.

[31] The Spearman correlation is 0.584 for relegation and 0.534 for qualifying for Europe.

reasonable approximation when using the best-case/worst-case scenarios of obtaining one more/less point than the other teams in each remaining match.

1.8 Appendix A: Further example contingency tables

This appendix provides two more example contingency tables for two different values of relegation importance (similar to Table 1 in the main article illustrating championship importance).

		Newcastle relegated?		
		No	Yes	Sum
Newcastle's result	Win	2,526,228	89,756	2,615,984
	Draw	1,473,503	1,089,800	2,563,303
	Loss	0	4,820,713	4,820,713
	Sum	3,999,731	6,000,269	10,000,000

Relegation importance for Newcastle: tau-c = 0.876

Table 8: Contingency table for Newcastle before Aston Villa – Newcastle match (2008/09, round 38)

Table 8 shows the contingency table for Newcastle before their last-round away match against Aston Villa in 2008/09 season. Before that match, Newcastle had a 60% probability of being relegated, but this probability was 100% in case of losing the match, 42.5% in case of drawing the match, and just 3.4% in case of winning the match. The final season outcome hung in the balance and was almost completely determined by the match result, so the relegation importance of this match for Newcastle is 0.876; the highest value in the data set.

		Portsmouth relegated?		
		No	Yes	Sum
Portsmouth's result	Win	3,810,914	153,241	3,964,155
	Draw	2,235,547	322,909	2,558,456
	Loss	2,781,657	695,732	3,477,389
	Sum	8,828,118	1,171,882	10,000,000

Relegation importance for Portsmouth: tau-c = 0.151

Table 9: Contingency table for Portsmouth before Portsmouth – Bolton match (2008/09, round 32)

Table 9 presents a much less dramatic situation before Portsmouth's home match against Bolton in the 32nd round of the same season. Portsmouth had a relatively low probability of 11.7% of being relegated, but winning the next match against their neighbors in the league table would reduce this probability to 3.9%. On the other hand, losing the match would almost double their probability of relegation to 20%. The match was clearly important for Portsmouth, but far from critical, so the corresponding relegation importance value is 0.151. This is much lower than in the previous example, but still higher than 90% of all relegation importance values in the data set.

1.9 Appendix B: Further verifying season predictions

This appendix expands on the analysis of the unbiasedness of season outcome predictions from section Verifying season predictions. Again, the original 8,360 estimated probabilities of a team finishing in the top half of the table are ordered from the lowest to the highest and all the impossible and certain predictions are removed. This time, the remaining 7,311 estimates are split into 50 equal-sized bins instead of just 6. The results are summarized in Figure 2, where each point represents one bin and the diagonal line represents no bias. Again, there is no systematic and substantial deviation from the diagonal line, so the season outcome predictions do not seem to be substantially biased.

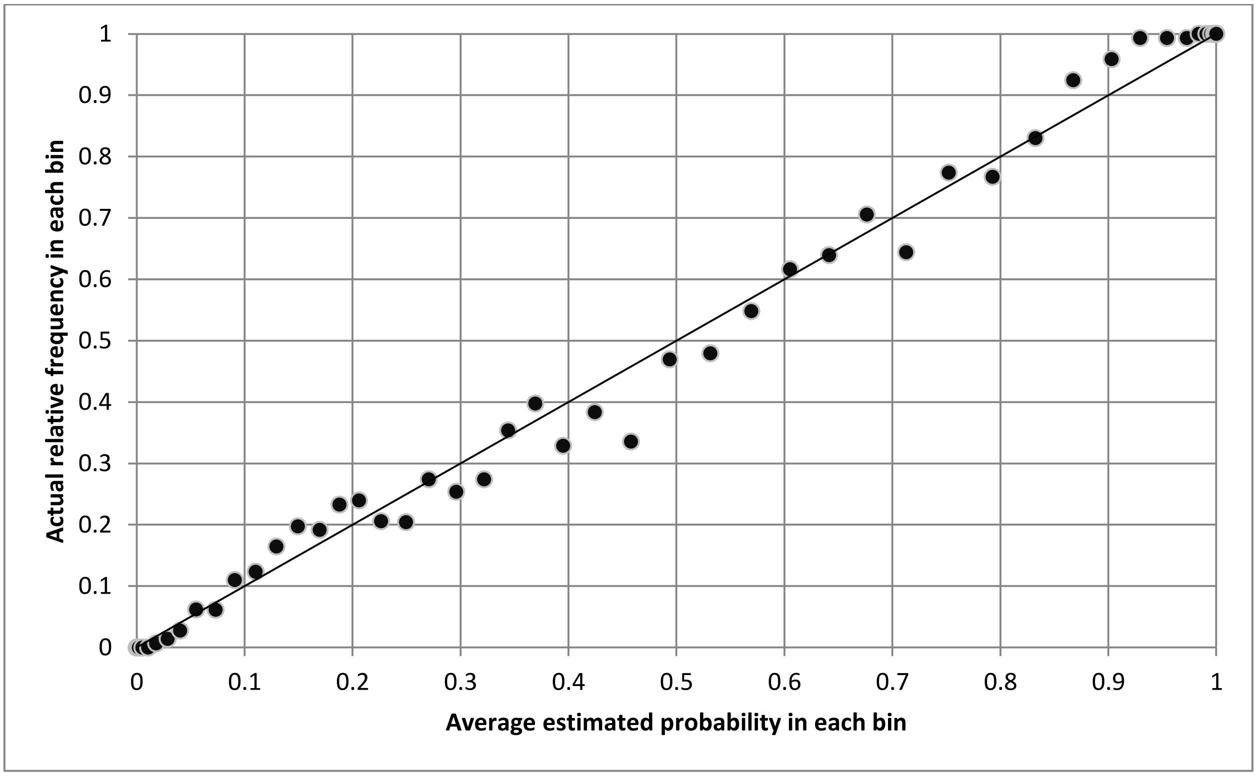

Figure 2: Predicted vs. actual frequencies of outcome, 50 bins

2 The impact of playoffs on seasonal uncertainty in the Czech ice hockey Extraliga[32]

In the top Czech ice hockey competition "Extraliga", 14 geographically close teams compete during a regular season in a pure round-robin tournament. However, the eventual champion is determined in the additional playoff stage and the regular season just decides which teams qualify for the playoffs and how these teams are seeded. This article uses a Monte Carlo simulation to show that although the additional playoff stage heavily favors higher-seeded teams and consists of a lot of games, it lowers the average probability of the strongest team becoming a champion from 48 to 39 percent and thus increases seasonal uncertainty.

2.1 Introduction

One of the most important results of sports economics is the observation that a tighter competition with a more uncertain outcome will attract more spectators. This so-called uncertainty of outcome hypothesis was first formulated by Rottenberg (1956), who noted that a baseball team winning too many games would attract fewer spectators, and later expanded on by Neale (1964), who claimed that a sports league will attract higher attendances if league standings are close and change often.

The sports economics literature distinguishes three different types of uncertainty of outcome (Szymanski 2003, García and Rodríguez 2009); match uncertainty (how uncertain the result of

[32] A shorter version of this article (without appendices) was published online ahead of print on October 23, 2013, in Journal of Sports Economics, doi: 10.1177/1527002513509109.

one specific match is), seasonal uncertainty (how uncertain the competition winner and other similar outcomes are), and championship uncertainty (whether there is a long-run domination by one team). Instead of seasonal or championship uncertainty, some authors use the term competitive balance; however, Scarf et al. (2009) make a useful distinction between competitive balance, which is defined as relative strengths of competing teams, and uncertainty of outcome, which also depends on tournament design.

The hypothesis that higher seasonal uncertainty increases attendance has a substantial empirical support (see Szymanski (2001) for English Premier League, Humphreys (2002) for the American MLB, Pawlowski (2013) for German Bundesliga, Pawlowski and Budzinski (2013) for three major European soccer leagues, or Szymanski (2003) for an overview of multiple studies) and also seems to be accepted by many competition organizers. The organizers have two obvious ways of increasing uncertainty of outcome; first, increase competitive balance by redistributing resources using mechanisms such as TV and gate revenue sharing, payroll caps, or giving weaker teams earlier draft picks;[33] second, increase the seasonal uncertainty by modifying the tournament design. A major choice in team sports is between using only a round-robin tournament (e.g. English Premier League) and combining the round-robin tournament with an additional playoff stage (e.g. US Major League Soccer).

The impact off the additional playoff stage on seasonal uncertainty has come into focus only recently. Fort and Quirk (1995) noted that if playoffs are added to the regular season, the regular season winner is far from certain to also win the playoff stage. Szemberg et al. (2012) analyzed 36 seasons of eight top ice hockey competitions and confirmed that the regular season winner won the playoff stage in just 43percent of the cases. Longley and Lacey (2012) used the NHL and NBA results to estimate probabilities that various seeds win the conference playoffs. However, these approaches do not analyze the probability that the strongest team actually wins the regular season and therefore cannot be used to tell whether the additional playoff stage increases or decreases seasonal uncertainty. Longley and Lacey (2012) also used a clever alternative method and showed that team payroll in the NHL (but not the NBA) better predicts results in the regular

[33] Sports economists have extensively analyzed these mechanisms and are generally quite skeptical about their efficiency and true goals (Vrooman 1995, Szymanski 2001, Szymanski 2003, Szymanski and Késenne 2004).

season than in the playoffs, indicating that the additional playoff stage increases seasonal uncertainty. However, this approach is limited by a small dataset and cannot be used to find out how exactly the additional playoff stage impacts the championship chances of specific teams or to analyze various alternative tournament designs.

Another approach that has been used to compare various tournament designs is the Monte Carlo simulation method. In this method, real competition results are used to estimate team strengths, these strengths (assumed to be constant over the season) are then employed to repeatedly simulate all individual matches in a specific tournament design, and the completed simulations are used to estimate probabilities of various outcomes. Scarf et al. (2009) used the Monte Carlo simulation to compare various designs for the UEFA Champions League. The round-robin design (which would be extremely impractical in reality due to a large number of matches) maximized the probability of the strongest team winning the tournament, while the unseeded 2 leg knock-out design maximized the uncertainty of outcome. A similar approach was used in Scarf and Yusof (2011) to show that seeding favors stronger teams and thus reduces uncertainty of outcome in FIFA World Cup finals.

This article is the first to use the Monte Carlo simulation method to analyze the impact of the additional playoff stage on seasonal uncertainty. The competition chosen for this analysis is the top Czech ice-hockey competition “Extraliga”. In the Extraliga, there are currently 14 geographically close teams that compete during the regular season in a one-group round-robin tournament. However, the eventual competition champion is determined in the additional playoff stage and the regular season is just used to decide which teams qualify for the playoffs and how these teams are seeded. Since the regular season design is balanced and does not favor any team (unlike in the NHL or KHL), the additional playoff stage is not really necessary to determine the competition winner; in fact, before the 1985/86 season, the regular season winner was declared the champion and the season ended without any playoffs.

There are four main reasons why ice hockey in general and the Extraliga in particular are good choices for this analysis. First, ice hockey teams typically rotate 15 to 20 players in each game, so injuries and form fluctuations of individual players change team strengths less than in sports such as basketball and soccer, making the simulation results more reliable. Second, the one-group competition design of the Extraliga simplifies the analysis and means that the playoff stage must be employed by the organizers with the intent to manipulate seasonal uncertainty and not simply to determine the champion (the major reason for using playoffs is unlikely to be to directly increase attendance and revenues, because the total attendance per week during the regular season is actually slightly higher than during the postseason).[34] Third, there are only few player transfers after the season starts, again supporting the assumption that team strengths during the season do not change much.[35] Fourth, the threat of relegation and the absence of draft mean that there are very limited incentives for teams to deliberately lose games.[36]

The Extraliga tournament design raises two related questions. First, does the strongest team have a higher probability of winning the regular season, or the additional playoff stage? Second, how does adding the playoff stage impact the probabilities of all the other teams that they become a new champion? To analyze these questions, this article employs a Monte Carlo simulation based on six different sets of realistic team strengths derived from the actual results of six Extraliga seasons (2006/07-2011/12). The simulation results show that although the additional playoff stage heavily favors teams that placed better in the regular season and consists of quite a lot of games, it lowers the probability that the strongest team becomes the champion (especially if this

[34] During the 2011/12 regular season, the total attendance was 75,881 spectators per week; during the postseason, the total attendance was just 63,006 spectators per week. This is not surprising; while the average attendance per game is higher during the postseason, many teams get eliminated early and have zero attendance. Although the lower postseason attendance is compensated by higher ticket prices, the playoff stage does not seem to be a much better revenue generator than simply making the regular season longer.

[35] Transfers are allowed up to January 31st, i.e. about 8 rounds before the end of the regular season. At this point, the final ranking is still very unpredictable, so teams do not have much reason to, for example, sell or loan top players; even the teams that are unlikely to qualify for playoffs still need to be ready for playout games where they have to fight against relegation. Indeed, the transfer activity during the season is mostly limited to lower-quality players often going to lower-level competitions to even get to play; for example, only 4 out of 50 top players (based on the sum of goals and assists) changed their club after the start of 2011/12 season and3 out of these 4 transferred well before the transfer deadline (source: prestupy.onlajny.cz, accessed on September 6th, 2013). This pattern is different only during the NHL lockouts, which bring many high-quality players for a part of the regular season, but no lockouts happened during the six seasons in the dataset.

[36] See Taylor and Trogdon (2002) for evidence of such strategic behavior (losing to obtain better draft selections) in the NBA.

team is very dominant) and raises this probability for weaker teams (especially if they are significantly weaker than the strongest team, but still above average). This is also true for some obvious modifications of the playoffs. Therefore, the addition of the playoffs to the regular Extraliga season increases seasonal uncertainty.

The rest of the article is organized as follows: Section 2 overviews the Extraliga tournament design and compares it to other competitions; Section 3 describes the individual game model, how team strengths are estimated from actual results, and how the whole season is simulated; Section 4 presents the simulation results; and Section 5 concludes.

2.2 Extraliga overview

The Czech ice hockey Extraliga was established in season 1993/94 (after Czechoslovakia split into the Czech Republic and the Slovak Republic) and is currently the most popular team sports competition in the Czech Republic.[37] Although the specific rules changed several times, the basic tournament design has stayed the same. First, all teams compete in a round-robin tournament that decides which teams qualify for the playoff stage and how they are seeded; second, the playoffs are used to determine the champion and all other final rankings. The same two-part tournament design was also regularly used in former Czechoslovakia since the 1985/86 season and experimented with in the 1970s. Under the current design, the Extraliga enjoys a very high degree of both seasonal and championship uncertainty – in the six seasons (2006/07-2011/12), there were six different regular season winners and five different playoffs winners with the same team winning both in only one season. This section describes the competition rules that were first

[37] In the 2011/12 season, the total Extraliga attendance was about 2.2 million spectators, while the regular season alone attracted almost 1.8 million spectators. In the same season, the top soccer competition "Gambrinus liga" attracted only a bit over 1.1 million spectators. The average regular season game attendance was 4,824 for the Extraliga and 4,710 for the Gambrinus liga. Sources: hokej.cz, fotbal.idnes.cz (both accessed on February 23rd, 2013).

implemented in season 2006/07, were in place during all six seasons analyzed in this article (2006/07-2011/12), and are still practically the same as of September 2013.[38]

The Extraliga consists of 14 teams. In the regular season, which typically runs from September to March, each team plays two home and two away games against all the other teams (4 x 13 = 52 games in total). Each ice hockey game consists of three 20-minute thirds (so-called regulation time). The team scoring more goals is the winner and receives 3 points, while the losing team gets 0 points. A draw is not possible – if a game is undecided in the regulation time, it goes into extra time, which lasts either 5 minutes or until a goal is scored. If the game is not decided in the extra time, a penalty shootout determines which team is considered to have scored the decisive goal. The extra time/penalty shootout winner receives 2 points, while the losing team gets 1 point. In the final regular season league table, teams are ranked according to the following criteria (in that order): total points, points from head-on games against teams with the same number of total points, score difference in these head-on games, total score difference, and total number of goals scored. Since the 2009/10 season, the regular season winner has actually received a minor trophy (the President's Cup); before, there was no trophy at all. After the regular season, all teams play at least several additional games – the top 10 teams qualify for the playoffs, while the bottom 4 teams proceed to the play-out (this additional round-robin stage, whose results are added to the regular season points, determines which team has to defend its Extraliga spot against a lower competition winner).

The playoff stage, which usually takes place in March and April, consists of four rounds – the preliminary round, the quarterfinals, the semifinals, and the final. In the preliminary round, teams that finished 7th – 10th in the regular season compete for two spots in the quarterfinals, where they are joined by the top 6 teams. In each round, teams are seeded according to their regular season final rank and paired so that the highest surviving seed plays against the lowest surviving seed, the second-highest seed plays against the second-lowest seed, and so on. Each pair of teams plays a best-of-five (preliminary round) or best-of-seven (all the other rounds) series of games, so the first team to defeat their opponent three (preliminary round) or four times (all the other rounds)

[38] The rules were compiled from the following websites: cslh.cz (Czech Ice Hockey Association), hokej.cz, and avlh.sweb.cz (Archive of Ice Hockey Results); all websites were accessed on February 15th, 2013.

proceeds to the next round. In each series of games, the higher-ranked team plays the first, second, fifth, and seventh game on its home ice. If a game is tied, the extra time lasts 10 minutes instead of 5. In the fifth (preliminary round), or seventh game (all the other rounds), penalty shootouts are not possible and any extra time lasts until a goal is scored.

The Extraliga playoff stage is quite similar to the system used in the top two ice hockey club competitions in the world – the NHL (USA and Canada) and KHL (Russia and other countries). However, the regular season in these two competitions is different; the participating teams are split into groups according to their geographical location and play more games against geographically close teams. Since teams are not grouped according to their strengths, groups are not designed to be balanced and the regular season winner is not clear (though it can be determined based on the overall record). Therefore, the playoff stage used in the NHL and KHL is somewhat of a necessity due to large distances between teams, but seems completely optional in a small country such as the Czech Republic. It is also interesting that the additional playoff stage is practically nonexistent in European soccer competitions, but used in the Major League Soccer (USA and Canada). On the other hand, other top European ice hockey competitions (e.g. in Sweden, Finland, Germany or the Slovak Republic) are organized very similarly to the Czech Extraliga and do use playoffs. It can be concluded that for small countries, adding a playoff stage to a pure round-robin tournament is simply a design choice, not a necessity. The question is – what is the impact of this design choice on seasonal uncertainty in general and on chances of the strongest team in particular?

At first sight, the Extraliga playoff stage should be quite good at identifying the strongest team. First, each pair of teams plays up to seven games to determine which team moves to the next round; this is much more than one or two matches typically used in soccer. Second, teams are reseeded for each round; this type of seeding was shown to help the strongest teams the most by Scarf and Yusof (2011). Third, higher-seeded teams play any decisive game in the series on their home ice. However, the regular-season round-robin tournament consists of a high number of games as well, so it also seems to be suitable for determining which team is the strongest. Clearly, a more detailed analysis is needed to decide which type of tournament design favors which teams.

2.3 Model

To analyze the impact of the additional playoff stage on seasonal uncertainty, this article uses a three-step Monte Carlo simulation method similar to Scarf et al. (2009). First, actual results of six different regular Extraliga seasons are used to estimate six sets of team strengths. Second, these six sets of team strengths are used to simulate 1,000,000 times each of the six corresponding seasons including the playoff stage down to the level of an individual game score (the actual season results are then used to verify that these simulations are realistic). Third, the resulting dataset consisting of 6,000,000 completed simulations is used to investigate the impact of the playoffs on seasonal uncertainty.

To generate an individual game score between any two teams, this article employs one of several methods introduced by Maher (1982), but modified for ice hockey. The unmodified method assumes that each team's strength can be described by four parameters – attack strength in home games *(HomeAttack)*, attack strength in away games *(AwayAttack)*, defense strength in home games *(HomeDefense)*, and defense strength in away games *(AwayDefense)*. For attack strengths, a higher number is better, while for defense strengths, a lower number is better. If a team i plays at home against team j, the score is composed of two random numbers drawn from two independent Poisson distributions with expected values of $HomeAttack_i$ * $AwayDefense_j$ (goals scored by the home team) and $HomeDefense_i$ * $AwayAttack_j$ (goals scored by the away team).

To be able to simulate any possible game in a season, it is necessary to somehow set 56 parameter values (14 teams x 4 strength parameters per team). It would be possible to randomly generate one or more sets of these parameters, but they would not necessarily correspond to a typical team strength distribution in a season. A better solution is to estimate the parameters based on actual results (Maher 1982, Scarf and Yusof 2011). In this article, the actual results of six Extraliga seasons (2006/07-2011/12) are used to estimate six realistic sets of parameters. For each season, this is done by setting the total expected numbers of regulation-time goals scored and conceded by each team in its home and away games equal to the corresponding actual values in a given season[39] and solving the resulting system of equations (for details, see Appendix A: Estimating team strengths).

The simple model described above does not take into account two factors specific to ice hockey – first, a team trailing by one goal towards the end of the game usually plays much more aggressively and eventually replaces their goaltender with another attacking player, thus dramatically increasing chances of both teams to score; second, a tied game does not end, but goes into extra time (possibly followed by a penalty shootout).

To take into account the option of pulling the goaltender, the model is modified in the following way: First, the home team i scores a random number of goals drawn from a Poisson distribution with an expected value of $7/8 * HomeAttack_i * AwayDefense_j$ and the away team scores a random number of goals drawn from a Poisson distribution with an expected value of $7/8 * HomeDefense_i * AwayAttack_j$ (this represents the score several minutes before the end of the game). If neither team is trailing by one goal, the regulation time score stays unchanged. If the home team i trails by one goal, it scores an additional Poisson-distributed number of goals with an expected value of $3/10 * HomeAttack_i * AwayDefense_j$ and the away team j scores an additional Poisson-distributed number of goals with an expected value of $5/10 * HomeDefense_i * AwayAttack_j$. Similarly, if the home team i leads by one goal, it scores an additional Poisson distributed number of goals with an expected value of $5/10 * HomeAttack_i * AwayDefense_j$ and the away team j scores an additional Poisson-distributed number of goals with an expected value of $3/10 * HomeDefense_i * AwayAttack_j$. These expected values for last-minute goals are quite high and strongly favor the leading team, but they reflect two observations about ice hockey games made by Thomas (2007); first, the average number of goals per minute sharply increases in the last two minutes; second, if a goal is scored when one goaltender is pulled, it is about twice as likely to be scored by the leading team. This modification is also calibrated so that it does not change the expected number of goals scored by each team compared to the unmodified model – this means that the estimated strength parameters are still valid.

[39] All actual season data were gathered from the websites hokej.cz and avlh.sweb.cz (Archive of Ice Hockey Results); both websites were accessed on January 18th, 2013.

To model the extra time, it is simply assumed that if a game is tied after regulation time, an extra time/penalty shootout winning goal will be scored by the home team *i* with the probability of $(HomeAttack_i * AwayDefense_j)/(HomeAttack_i * AwayDefense_j + HomeDefense_i * AwayAttack_j)$ and by the away team *j* with the probability of $(HomeDefense_i * AwayAttack_j)/(HomeAttack_i * AwayDefense_j + HomeDefense_i * AwayAttack_j)$.

In the next step, the estimated strength parameters of all teams in a given season and the individual game model are used to estimate probabilities of a given team winning the regular season or the playoffs. The Monte Carlo approach is the only feasible option due to the high number of games and the complicated tournament design. First, results of all regular season games are randomly generated and points are assigned. Second, these results are used to put together the final table (using all applicable ranking criteria) and decide which teams qualify for the playoff stage. Third, the teams are seeded and paired and all corresponding playoff games are played until there is a competition champion. This process is repeated 1,000,000 times for each set of strength parameters. In the end, there are 6,000,000 completed simulations corresponding to six actual seasons. An *ex ante* probability of any scenario in any season is then approximated by the relative frequency of this scenario in corresponding simulations. Because the number of simulations is very high, the estimated probabilities are very close to the exact probabilities that could (in theory) be obtained by solving the model analytically.[40]

Although there are some possible improvements to predicting individual game results (Maher 1982, Dixon and Coles 1997, Rue and Salvesen 2000, Goddard 2005), the model as a whole is already quite realistic. This can be shown by comparing the aggregate simulation statistics against the corresponding actual results. Specifically, there are no significant differences between the simulated and actual total number of regular season goals (including extra time); the simulated and actual relative frequencies of home/away regulation/extra time wins; the simulated and actual minimum and maximum points in a given regular season; and the simulated and actual relative frequencies of playoff series results (all descriptive statistics and statistical tests are provided in Appendix B: Model verification).

[40] For example, the probability that a given team in a given season wins the regular season or the playoffs is estimated with a standard error less than 0.0005.

Of course, this basic model also has some limitations. First, it assumes that team strengths are constant over the whole season and do not fluctuate much (for example, because of injuries of key players). This assumption was confirmed for English Premier League by Koopman and Lit (2012), who found very small changes of attack and defense strengths over a course of a season, and by the first article in this dissertation called "Using Monte Carlo simulation to calculate match importance: The case of English Premier League," which found that predictions of English Premier League matches based on up-to-date historical results are only negligibly better than predictions based on results that were at least 180 days old (i.e. mostly previous season results). Also, as already said in the Introduction, this assumption should be even more valid for ice hockey than for soccer due to a higher number of players rotated in each game. Nevertheless, the impact of relaxing the constant-strengths assumption is investigated at the end of the next section.

The second assumption is that teams always play as well as they can and do not lose games on purpose. As said in the Introduction, the incentives for underperforming are much weaker in the Extraliga than in other competitions due to the absence of draft and the threat of relegation, but there are still two potentially viable strategies analyzed at the end of the next section; first, the strategy mentioned in Szemberg et al. (2012) of expending less effort during the regular season to have more energy for the playoff stage; second, the strategy of deliberately losing some games at the end of the regular season to avoid a specific opponent in the first round of the playoffs.

2.4 Results

To investigate the impact of the additional playoff stage on seasonal uncertainty, the dataset of completed simulations of six seasons is used to estimate two probabilities for each team in each season – the probability of winning the regular season and the probability of winning the playoffs. Since there are 14 teams per season, there are 6 * 14 = 84 pairs of probabilities. Based on these probabilities, it is possible to determine the strongest team in each season – it is simply the team with the highest probability of winning the regular season (and also the playoffs).[41] The

[41] It does not really matter if the teams are ordered by the probability of winning the regular season or by the probability of winning the playoffs; the ordering is identical for top 4 and bottom 4 teams in every season and very

84 pairs of probabilities (one pair for each team in each season) are represented by points in Figure 3. The strongest team in each season is marked by a bigger and darker point. On the diagonal line, the probability of winning the regular season equals the probability of winning the playoffs.

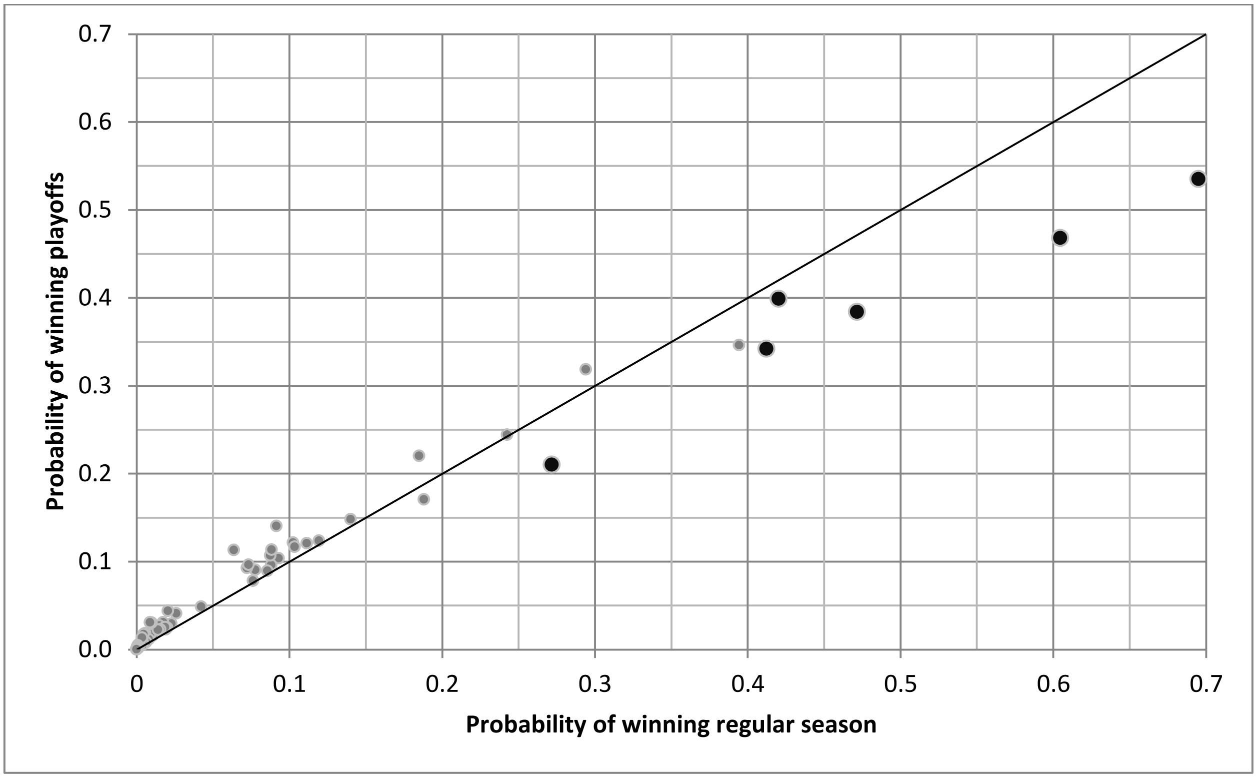

Figure 3: Probability of winning regular season vs. playoffs; the strongest team in each season is marked by a bigger point

The first important observation is that that all six points representing the strongest teams in each season are below the diagonal line, so the additional playoff stage decreased their probability of becoming the champion. This is also confirmed by computing average probabilities across all six seasons; on average, the strongest team had a 48 percent chance of winning the regular season, but just 39 percent chance of winning the playoffs, so the additional playoff stage decreased their

---

similar otherwise. Theoretically, an average team that would become stronger in away games and weaker in home games could keep their regular season chances constant, while increasing their chances in the playoffs; however, this seems to have a negligible impact in the dataset.

probability of winning the competition by 9 percentage points. This result is in line with the finding of Longley and Lacey (2012) for the NHL, but also shows that it would be inaccurate to assume that the strongest team is the same as the regular season winner.

The difference between the two probabilities is especially large for very dominant teams – Sparta Praha in the 2011/12 season (represented by the rightmost point) had a 69 percent chance of winning the regular season (and did actually win), but just a 53 percent chance of winning the playoffs (and did not actually win), so the additional playoff stage decreased their probability of becoming a champion by 16 percentage points. On the other hand, all points representing weaker teams (less than a 15 percent probability of winning the regular season) are above the 45-degree line, so such teams' chances of becoming a champion are helped by the additional playoff stage.

To analyze the change in championship probability from adding the playoff stage in more detail, it is useful to look at these changes for the strongest team in each season (i.e. with the highest probability of winning the regular season), the second-strongest team (i.e. with the second-highest probability of winning the regular season), the third-strongest team, and so on. The maximum (bar top), average (black line), and minimum (bar bottom) changes for each level of team strength are presented in Figure 4.

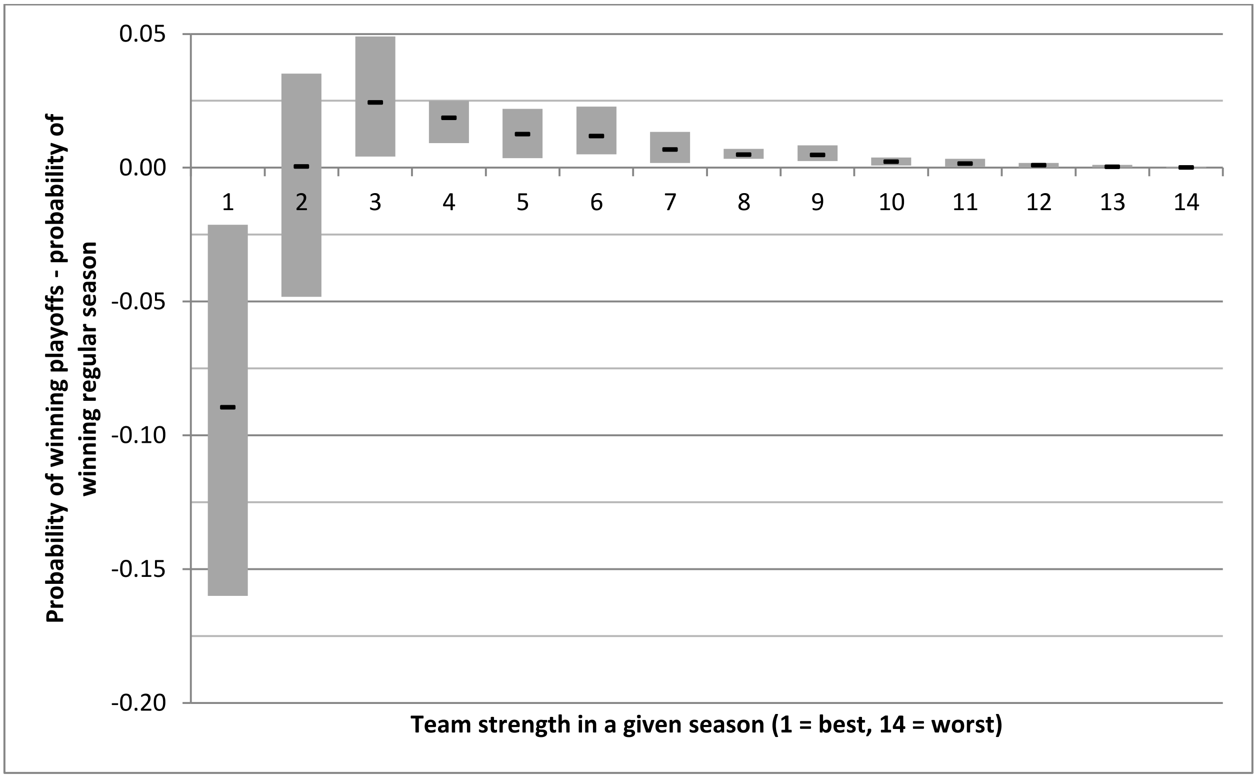

Figure 4: Relative team strength vs. change in championship probability from adding playoffs

The graph shows that the additional playoff stage always decreased the championship probability of the strongest team in the analyzed dataset, sometimes helped and sometimes hurt the second-strongest team,[42] and always helped all the other, weaker teams. The teams that benefited the most were the third-strongest to sixth-strongest teams (much weaker than the strongest team, but still above average). On the other hand, the weakest teams were only negligibly affected, since both their probabilities – winning the regular season and winning the playoffs – were close to zero.

To summarize, the additional playoff stage decreases the championship chances of the strongest team (especially if it is very dominant), increases the championship chances for the other teams (especially if they are significantly weaker than the strongest team, but still above average), and thus increases the seasonal uncertainty. A very similar pattern emerges for different individual

[42] The second-strongest team was helped when being much weaker than the strongest team and hurt when being just a little weaker than the strongest team.

game models, different team strength distributions,[43] or when comparing the probabilities of finishing in the top 2 in the regular season against the probabilities of reaching the playoff final. The additional playoff stage also increases seasonal uncertainty measured by Herfindahl–Hirschman Index or Gini coefficient.[44]

A natural question is *why* the playoffs increase seasonal uncertainty, especially considering that the combination of reseeding before each round and home ice advantage should strongly favor teams that do better in the regular season and are therefore seeded higher. First, let's investigate how big this advantage for stronger teams actually is.

Szemberg et al. (2012) observed that higher-seeded teams tend to win the playoffs much more often. However, this would happen even if the playoff stage design did not favor higher-seeded teams at all, since higher-seeded teams also tend to be stronger. To determine how the regular season final rank influences the probability of winning the playoffs, it is therefore necessary to keep the team strength constant. This is impossible based on just observational data, but easy using the simulation approach; for each team in each season, the probability of winning the playoffs given a particular regular season final rank can be estimated as the relative frequency of winning the playoffs in a subset of simulations where the team reached that rank. Figure 5 shows how attaining a specific seed influences the probability of winning the playoffs for four selected teams of various strengths (dotted lines) and averaged across all teams and seasons (solid line), while bars represent simulated probabilities of a given seed winning the playoffs in the whole dataset (no matter which specific team it is).[45]

---

[43] A simplified individual game model without the modification for pulling the goaltender and giving each team 50 percent probability of winning any extra time underestimates the number of games decided in regulation time, but leads to almost identical championship probabilities. Further increasing the strength of the strongest team in each season (by multiplying both attack strength parameters and dividing both defense strength parameters by a number greater than one) confirms that as a team becomes more dominant, the playoff stage decreases its championship chances by more percentage points.

[44] See Humphreys (2002) for an overview of these measures as applied to uncertainty of outcome.

[45] The standard errors for probability estimates averaged across all seasons (solid line and bars) are negligible because they are calculated from all 6,000,000 simulations. For individual team/season combinations, the standard errors range from a fraction of a percentage point for team/seed combinations that happen often (e.g. a strong team

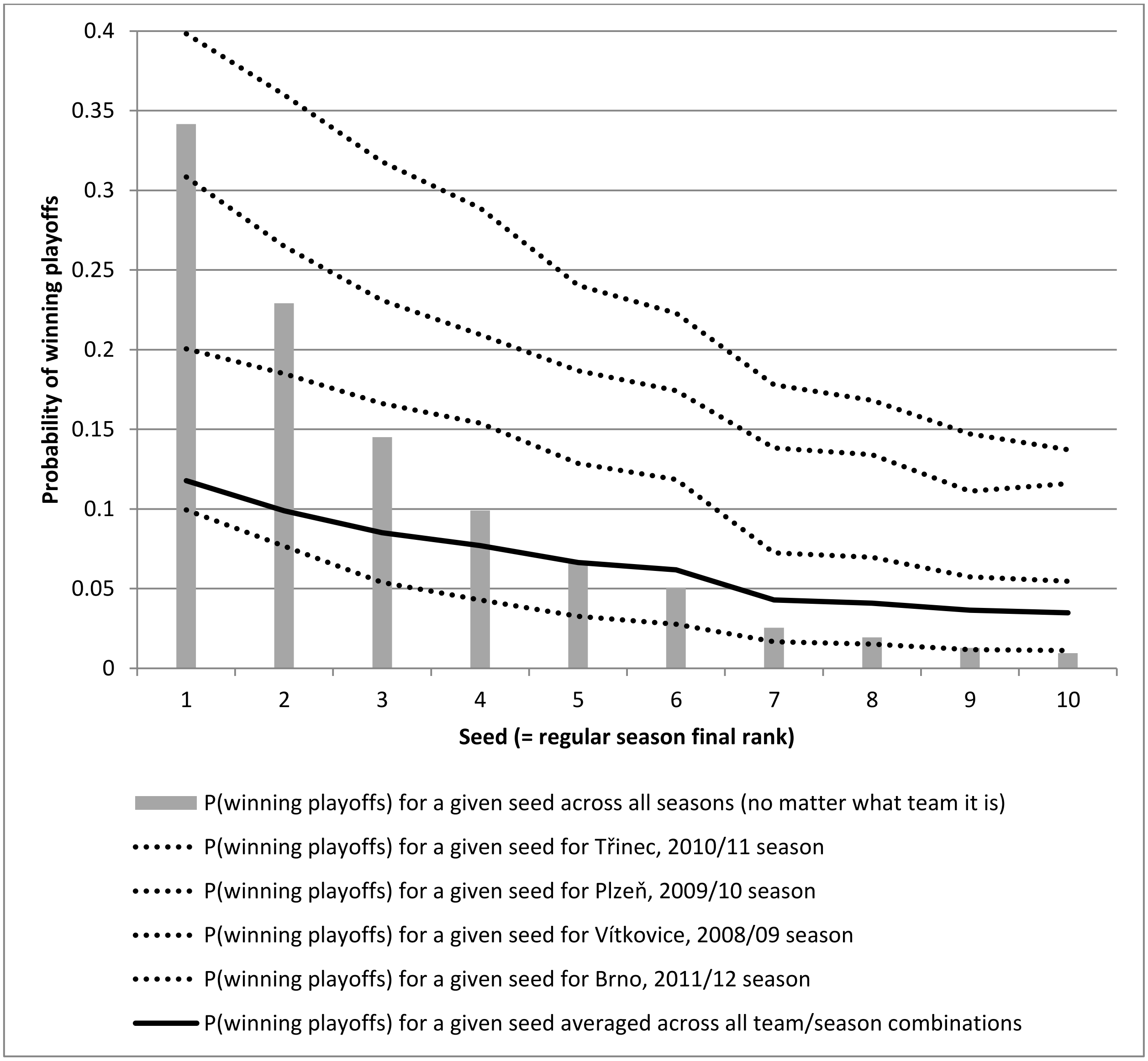

Figure 5: Probability of winning playoffs given a specific seed

It is obvious that as the seed gets worse, the championship probability goes down much more slowly when controlling for team strength, so the observed pattern of higher-seeded teams winning much more often can be mostly explained by these teams simply being stronger. Nevertheless, a better regular season result still provides a significant advantage when keeping

winning the best seed) to one to two percentage points for team/seed combinations that happen rarely (e.g. a strong team obtaining seed 9 or 10).

the team strength constant; on average, obtaining the best seed roughly triples the championship probability compared to the worst seed.

Based on the analysis above, it is clear that the playoff stage design actually does heavily favor higher-seeded teams. To understand why it still decreases the probability that the strongest team wins, let's go back to the example of Sparta Praha in season 2011/12. As mentioned above, Sparta Praha had a 69 percent probability of winning the regular season and a 53 percent probability of winning the playoffs. The probability of Sparta Praha winning the playoffs can be expressed as a product of four different numbers – they had a 99.97 percent probability of qualifying for the quarterfinals (either directly or from the preliminary round); if they qualified, they had an 88 percent probability of advancing to the semifinals; if they did, they had an 82 percent probability of progressing to the final; and if they did, they had a 74 percent probability of winning the whole competition. Therefore, the lower probability of winning the playoffs is not caused by Sparta Praha not being a clear favorite in each round, but rather by even small probabilities of elimination accumulating over multiple rounds.

The additional playoff stage also decreases the probability of the strongest team winning the competition for some reasonable alternative designs. For example, the quarterfinals, semifinals, and final would have to use a best-of-fifteen instead of best-of-seven system (i.e. eight instead of four wins to eliminate the other team) to approximately neutralize the impact of the playoffs on seasonal uncertainty. This would make the playoff stage much longer, decrease the importance of a single game, and likely lower the interest of spectators. Similarly, if only top four teams qualified for the playoffs, the regular season finish would be less interesting and the strongest team would still have a lower probability of winning the playoffs (42 percent) than winning the regular season (48 percent).[46]

The Monte Carlo simulation approach also allows directly investigating the impact of alternative seeding mechanisms. For example, if teams were seeded randomly in each round instead of the current system of highest-ranked teams in the regular season facing lowest-ranked teams, the probability of the strongest team winning the playoffs would go down from 39 to 35 percent. This

[46] These and all the subsequent results are based on additional sets of 100,000 simulations using the same team strengths and seasons, but different tournament designs and/or assumptions.

again confirms that the current seeding system favors the stronger (and usually higher-seeded) teams. A similar decrease in probability of winning the playoffs would result from using the seeding system where the highest surviving seed would play the second-highest surviving seed, the third-highest surviving seed would play the fourth-highest surviving seed, and so on.[47]

As stated in the previous section, all the results so far are based on two assumptions; first, team strengths are constant over the whole season; second, teams always play as well as possible and do not lose regular-season games on purpose to improve their chances in the playoffs. The rest of this section analyzes how relaxing these assumptions influences the results.

To simulate the impact of possible strength fluctuations, the whole season is split into three roughly equal parts – the first half of the regular season, the second half of the regular season, and the playoff stage.[48] Each simulation starts with the original home and away attack and defense strengths for all teams; however, these are not kept constant, but are modified by applying different independent random shocks in each part of the season. Specifically, both home and away attack parameters of team *i* in season part *j* of simulation *k* are multiplied by $exp(ShockSize * AttackShock_{ijk})$, where $AttackShock_{ijk}$ is a random number drawn from the standard normal distribution and the *ShockSize* parameter determines the magnitude of strength fluctuations. Similarly, both home and away defense parameters are multiplied by $exp(ShockSize * DefenseShock_{ijk})$, where $DefenseShock_{ijk}$ is again a random number drawn from the standard normal distribution.

In the original model with constant team strengths (equivalent to $ShockSize = 0$), the strongest team had a 48 percent probability of winning the regular season and 39 percent probability of winning the playoffs. Increasing the *ShockSize* parameter introduces additional noise into the simulation, so both probabilities must go down; they are 46 and 37 percent for $ShockSize = 0.05$ and 41 and 33 percent for $ShockSize = 0.1$. The key conclusion is that even when allowing for strength fluctuations, the probability that the strongest team wins the regular season is still

[47] Such a seeding system is unlikely to be used (unlike random seeding) because it would actively discourage teams from playing as well as possible during the regular season (e.g. seed number 3 is clearly better than seed number 2; it offers the home ice advantage and likely a weaker first-round opponent).

[48] All teams play 52 games in the regular season and at most 26 games in the playoff stage (including the preliminary round).

substantially larger than the probability of winning the playoffs; in fact, the ratio between these two probabilities stays almost exactly the same.[49] Also, the impact of this modification on absolute probabilities should be rather small, because even the lower *ShockSize* value produces fluctuations that are more typical of those estimated by Koopman and Lit (2012) for multiple seasons rather than for just one season.

The second assumption that needs to be analyzed is that teams do not behave strategically by deliberately losing regular season games to increase their probability of winning the playoffs. The first option – expending less effort during the regular season to have more energy for the playoffs – does not seem rational in the Extraliga, because teams finishing in top 6 in the regular season avoid the preliminary round and enjoy about one and a half weeks of rest before the quarterfinals (quite a long period considering that teams usually play three games per week during the regular season). Consequently, any additional benefit from expending less effort at the end of the regular season would be negligible, while the cost of getting a worse seed would be substantial (as was shown in Figure 5).[50]

A more promising strategy could be to deliberately lose some games at the end of the regular season to avoid a strong opponent in the first (or preliminary) round of the playoffs. To simulate the impact of such a strategy, the strongest team in each season moves down one seed at the end of the regular season if it means getting a significantly weaker opponent (using the strength ranking based on the probability of winning the regular season) in the first (or preliminary) round of the playoffs. The strategy is not used if it would mean losing the home ice advantage (moving from seed 4 to 5 or from 8 to 9; in such cases, the opponent would actually stay the same) or having to enter the preliminary round (moving from seed 6 to 7). If the strongest team is supposed to face a preliminary-round winner, it is assumed that a stronger team always advances from this preliminary round. This strategic option is more powerful than anything available in reality; it assumes that the strongest team can perfectly predict the final table and that the other

[49] As *ShockSize* increases to infinity, the original team strengths are completely overshadowed by random shocks and all probabilities converge to 1/14 (i.e. all teams have the same chance of winning both the regular season and the playoffs). However, such a scenario is clearly unrealistic.

[50] Also, if this strategy worked, most teams would use it, thus effectively neutralizing it.

teams cannot use any similar strategy,[51] so it represents an upper bound on how much the strongest team could increase its probability of winning the playoffs.

There are multiple versions of the strategy depending on how much weaker the new opponent has to be for the strongest team to move down one seed – a logical expectation is that as the required minimum strength difference increases, the strategy is employed less frequently, but becomes more efficient if actually used. This is confirmed by simulation results; if any difference in strengths is enough (e.g. moving down one seed if it means facing at most the 10th strongest instead of the 9th strongest team), the strategy is utilized in 43 percent of all simulated seasons, but the probability that the strongest team wins the playoffs actually decreases from 39.0 to 38.3 percent, indicating the strategy is counterproductive. If the minimum strength difference is raised to 5 (e.g. moving down one seed if it means facing at most the 7th strongest instead of the 2nd strongest team), the strategy starts having a slightly positive effect, but is utilized in only 11 percent of all simulated seasons, so the probability that the strongest team wins the playoffs increases by less than 0.1 percentage points.[52] The reason why this type of strategic behavior leads to a negligible improvement at best is simple; moving down a seed might mean a weaker first-round opponent, but the later-round opponents will tend to be stronger and also more likely to have the home ice advantage.

2.5 Conclusion

As shown in the previous section, the additional playoff stage in the Extraliga lowers the average probability that the strongest team becomes a champion from 48 to 39 percent and thus increases

[51] The first-round opponent of any team is very hard to predict before the end of the regular season, because the point differences between teams tend to be small; for example, the average difference in the data set between the 5th place and the 10th place at the end of the regular season is less than 12 points. If other teams actively tried to avoid playing against the strongest team, using the strategy would become even more complicated.

[52] Even higher minimum strength differences again decrease the probability of winning the playoffs because the strategy is practically never used (e.g. in less than 1 percent of all seasons if the minimum difference is 9). Allowing the strongest team to move down multiple seeds (instead of one seed) also provides only a negligible improvement to the chance of winning the playoffs.

seasonal uncertainty. A similar result was obtained by Longley and Lacey (2012) for the NHL, but the Monte Carlo simulation approach used in this article enables a deeper analysis; first, it shows that the more dominant the strongest team is, the more their probability of winning the competition is decreased by the additional playoff stage; second, it demonstrates that the third-strongest to the sixth-strongest teams profit most from adding the playoffs; third, it shows that obtaining the best seed (compared to the worst seed) roughly triples the championship probability; fourth, it indicates that the additional playoff stage would also increase seasonal uncertainty for reasonable alternative designs; fifth, it shows that the results hold even if the assumptions of constant team strengths and no strategic behavior are relaxed.

The higher seasonal uncertainty makes the Extraliga competition more attractive – the supporters of the strongest team cannot be so sure about the final outcome and the fans of weaker teams have a stronger hope of celebrating the championship title. The fact that securing a higher seed significantly increases championship chances makes the regular season finish interesting for fans of almost all teams. The higher seasonal uncertainty is also likely to translate into a more even distribution of all types of revenues and thus a higher competitive balance. In a positive feedback loop, this should further increase seasonal uncertainty. Therefore, the current Extraliga tournament design promotes uncertainty of outcome, which seems to be preferred by both spectators and competition organizers.

The Monte Carlo simulation approach could also be applied to other team competitions, including those with unbalanced schedules and teams split into conferences/divisions;[53] however, for those similar to the Extraliga, such as other top European ice hockey competitions, the results

[53] Of course, applying the framework presented in this article to a competition with unbalanced schedule would be more complicated. Considering the example of the Major League Soccer (top US and Canadian soccer league), estimating team strengths would be more difficult, because the number of teams changes every several years and the playing schedule (the number of home versus away matches against each team) changes every season. This means that the necessary set of equations would be more complex and different every season. Also, the probability of winning the regular season (Supporters' Shield) or the playoffs (MLS Cup) would not only depend on the relative team strength, but also on the conference the team is in; even if there were no difference between team strengths between Eastern and Western Conferences, the Western Conference currently (2012/13) consists of fewer teams (9 compared to 10), so should *ceteris paribus* offer a better chance of getting into the MLS Cup final. On the other hand, the individual match model would not need to include the modification for pulling the goaltender.

are likely to be similar as well. Another promising avenue of research would be to analyze tournament designs in terms of whether strategic behavior during the regular season (i.e. losing on purpose) can meaningfully increase the probability of winning the playoffs; while this does not seem to be the case in the Extraliga, the situation could be different in competitions with different seeding mechanisms or fewer playoff rounds.

2.6 Appendix A: Estimating team strengths

This section describes how team strengths are estimated from the actual regular season results. This is done by setting the total expected numbers of regulation-time goals scored and conceded by each team in its home and away games equal to the corresponding actual values in a given season. For example, team 1 is expected to score $2 * HomeAttack_1 * AwayDefense_2$ goals in its two home games against team 2, $2 * HomeAttack_1 * AwayDefense_3$ goals in its two home games against team 3 … $2 * HomeAttack_1 * AwayDefense_{14}$ goals in its two home games against team 14. The sum of these expressions is set equal to the total number of goals that team 1 actually scored in all its home games. Eventually, this leads to the following set of 14 equations (one equation for each team *i*, where $i = 1 \dots 14$):

$$2 * HomeAttack_i * \sum_{j=1\dots14, j\neq i} AwayDefense_j = Total\ home\ goals\ scored\ by\ team\ i$$

Similar sets of equations are also put together for goals conceded in home games and goals scored and conceded in away games:

$$2 * HomeDefense_i * \sum_{j=1\dots14, j\neq i} AwayAttack_j = Total\ home\ goals\ conceded\ by\ team\ i$$

$$2 * AwayAttack_i * \sum_{j=1\dots14, j\neq i} HomeDefense_j = Total\ away\ goals\ scored\ by\ team\ i$$

$$2 * AwayDefense_i * \sum_{j=1\dots14, j\neq i} HomeAttack_j = Total\ away\ goals\ conceded\ by\ team\ i$$

In the resulting system, there are 56 equations and 56 variables; however, the equations are not independent, since the total number of home goals scored by all teams equals the total number of away goals conceded by all teams and the total number of home goals conceded by all teams equals the total number of away goals scored by all teams. Therefore, there are infinitely many solutions; these can be obtained from each other by multiplying all attack parameters by a positive number and dividing all defense parameters by the same number. Because all these solutions provide exactly the same predictions, this is not a problem and any solution will do.[54] As an example, Table 10 shows the regulation-time total scores in home and away matches for all teams in the 2011/12 season and also the corresponding set of estimated team strengths.

Team	Total score (regulation time)		Estimated team strengths			
			Home		Away	
	Home	Away	Attack	Defense	Attack	Defense
Sparta Praha	76:46	79:53	1.148	0.994	2.274	1.778
Plzeň	102:64	72:75	1.580	1.373	2.119	2.591
Pardubice	93:60	77:69	1.430	1.295	2.255	2.360
Liberec	72:58	62:77	1.115	1.228	1.809	2.578
České Budějovice	66:59	58:64	1.008	1.243	1.694	2.127
Vítkovice	81:67	55:80	1.259	1.408	1.622	2.704
Zlín	60:55	41:60	0.912	1.135	1.190	1.982
Brno	71:45	65:86	1.110	0.955	1.867	2.878
Kladno	73:58	51:75	1.128	1.212	1.487	2.513
Třinec	65:68	77:67	0.996	1.469	2.279	2.225
Karlovy Vary	75:67	57:81	1.166	1.411	1.681	2.721
Slavia Praha	62:78	71:73	0.955	1.674	2.127	2.418
Litvínov	73:71	56:96	1.153	1.494	1.660	3.222
Mladá Boleslav	70:71	46:83	1.090	1.475	1.362	2.774

Table 10: Regulation-time total scores and estimated team strengths, 2011/12 season

Given the estimated strengths, it is possible to calculate the expected number of regulation-time goals scored by each team in any match-up. For instance, if Sparta Praha plays at home against Mladá Boleslav, Sparta Praha is expected to score Sparta's *HomeAttack* * Mladá Boleslav's *AwayDefense* = 1.148 * 2.774 = 3.185 goals, while Mladá Boleslav is expected to score Sparta's

[54] Another option, used in Maher (1982), would be to impose additional constraints on parameter values.

HomeDefense * Mladá Boleslav's *AwayAttack* = 0.994 * 1.362 = 1.354 goals.[55] The total number of goals Sparta Praha is expected to score if they play two home matches against each opponent equals 76, i.e. the actual number of regulation-time goals scored in home matches in the 2011/12 season.

2.7 Appendix B: Model verification

This section shows that the model produces realistic results by comparing the aggregate simulation statistics against the corresponding actual results for each season or all seasons together. First, the total number of regular season goals scored in each season (including extra time) is compared against the total number of goals in the corresponding set of 1,000,000 simulations based on that season. Since team strengths are actually estimated from total goals, the median (50th percentile) number of goals in a simulated season should be very close to the actual total number of goals (the only reason for these two numbers being different is a random number of games going into extra time in the actual season). This comparison is presented in Table 11.

Season	Total goals (including extra time)			
	Simulation percentiles			Actual value
	5th	50th	95th	
2006/07	2051	2126	2202	2123
2007/08	1966	2039	2113	2057
2008/09	2024	2099	2175	2096
2009/10	1986	2060	2135	2057
2010/11	1950	2023	2096	2012
2011/12	1924	1996	2070	1995

Table 11: Simulated vs. actual total number of goals (including extra time)

In each season, the difference between the simulated median and the actual number of total goals is less than 1 percent; the average difference across all seasons is 0.02 percent. Therefore, the individual game model does not seem to be biased in terms of the total number of goals (including extra time).

[55] The correction for pulling the goaltender is calibrated to not change the expected numbers of goals.

In the second test, the simulated relative frequency of each type of game result (home team regulation/extra time win/loss) in a regular season is compared against the actual relative frequency. Because there are only 364 games in one season, all seasons are pooled together for the total number of 2,184 games to increase the test power. The relative frequencies across all seasons are shown in Table 12.

	Home team win		Home team loss	
	Regulation	Extra time	Extra time	Regulation
Simulated relative frequency	0.4915	0.1231	0.0988	0.2866
Actual relative frequency	0.5023	0.1200	0.1058	0.2720

Table 12: Simulated vs. actual relative frequencies of game result types, N = 2,184

For each type of result, the difference between the simulated and the actual relative frequency is within 2 percentage points and the distribution of actual result types is not statistically significantly different from the simulated distribution at $\alpha = 0.05$ (chi-square goodness-of-fit test, p-value = 0.334). Consequently, the model does not seem to be biased in terms of the result type.

Third, the total minimum and maximum numbers of points in each regular season (i.e. the points obtained by the team that finished last and the winner)[56] are compared against the total maximum and minimum numbers of points in the corresponding set of 1,000,000 simulations. Table 13 presents simulated point percentiles and the corresponding actual values.

[56] In the 2010/11 season, three teams were deducted points due to invalid player registration forms. These deductions are not taken into account in this test.

Season	Minimum points				Maximum points			
	Simulation percentiles			Actual value	Simulation percentiles			Actual value
	5th	50th	95th		5th	50th	95th	
2006/07	20	31	44	34	97	105	116	100
2007/08	25	38	51	40	99	109	120	106
2008/09	48	58	66	59	90	97	107	93
2009/10	44	54	62	58	96	105	116	106
2010/11	33	46	56	41	96	103	114	96
2011/12	43	53	61	61	99	109	121	107

Table 13: Simulated vs. actual minimum/maximum points

The average difference between the simulated median number of points and the actual number is 3.75 points (a little more than the difference between winning and losing a single game). Each interval between the 5th and 95th point percentiles can be thought of as a 90-percent confidence interval on the prediction of the actual number of points; there are 12 such intervals and the actual value lies on the interval boundary in two cases and never outside. Therefore, the simulation also seems to produce realistic regular season point distributions.

The last test compares the simulated relative frequencies of best-of-seven playoff series results (i.e. the quarterfinals + the semifinals + the finals) from the point of view of the higher-seeded team against the actual relative frequencies. Because there are only seven such results per season, all seasons are again pooled together for the total number of 42 series results. The simulated and actual relative frequencies across all seasons are shown in Table 14 (the first number in the result represents the number of games won by the higher-seeded team and the second number represents the number of games won by the lower-seeded team).

	Quarter/semi/final playoff series result							
	4-0	4-1	4-2	4-3	3-4	2-4	1-4	0-4
Simulated relative frequency	0.1004	0.2116	0.1512	0.1950	0.0963	0.1393	0.0679	0.0383
Actual relative frequency	0.0714	0.2857	0.1190	0.2143	0.0714	0.1429	0.0476	0.0476

Table 14: Simulated vs. actual relative frequencies of playoff series results, N = 42

The highest difference between the simulated and actual relative frequency is 7 percentage points (4-1 result); this result was predicted to happen about 9 times, but it actually happened 12 times. However, this is completely natural given the small sample size; the actual result type distribution is not statistically significantly different from the simulated distribution at $\alpha = 0.05$ (chi-square goodness-of-fit exact test, p-value = 0.937). It is also possible to look at the simulated versus the actual relative frequency of the higher-seeded team eliminating the lower-seeded team (4-0 + 4-1 + 4-2 + 4-3 results); again, the actual relative frequency of 0.6905 is not statistically significantly different from the simulated relative frequency of 0.6582 at $\alpha = 0.05$ (two-tailed t-test, p-value = 0.661).

3 Does match uncertainty increase attendance? A non-regression approach

The uncertainty of outcome hypothesis predicts that more balanced sports matches should attract higher attendances, but the empirical evidence is mixed at best. First, this article shows that some inconsistent findings in the literature could be explained by wrongly specified regressions. Second, a new approach to analyzing the effect of match uncertainty is proposed. Using data about nine seasons of the English Championship, the article shows that in a pair of matches where both home teams are slight favorites, a switch of the corresponding away teams would decrease the total attendance by several percent, while the opposite is true if both home teams are underdogs or strong favorites. These results suggest that attendance demand is a bell-shaped function of match balance that is maximized if teams of the same quality play against each other.

3.1 Introduction

Do more balanced sports matches attract higher attendances? The uncertainty of outcome hypothesis (Rottenberg, 1956; Neale, 1964) certainly predicts so, but the empirical evidence is mixed at best. So far, the link between match uncertainty and attendance has been examined by regressing individual match attendance (or its logarithm) on variables representing qualities of both teams, other variables influencing attendance (ticket price, team rivalry, distance between teams, weather…), and a variable measuring how the match is balanced. 18 such studies reviewed in Borland and McDonald (2003) investigated different sports (mostly soccer and baseball), used different ways of measuring team quality (team ranks or points/goals per game) and match uncertainty (difference in team ranks or points per game; absolute value of betting spread; quadratic specification of home win probability derived from betting odds), and arrived at different results; some studies found that higher match uncertainty increases attendance, some found the opposite, some found that attendance increases with home win probability (and possibly starts decreasing if home win probability is higher than 0.6-0.7), some found no significant effect.

Similarly contradictory results can also be found in more recent research. Buraimo and Simmons (2008) modeled English Premier League attendance and concluded that attendance is minimized if home win probability derived from betting odds equals about 0.35. Buraimo and Simmons (2009) obtained a similar result for Spanish soccer. However, Benz et al. (2009) found that for one model specification, German Bundesliga attendance (excluding season tickets) was maximized for home team win probability equal to 0.53. Contradictory results for German Bundesliga were obtained by Pawlowski and Anders (2012); in one regression specification, attendance decreased if home team was a favorite rather than outsider; in another specification; higher match uncertainty decreased attendance. Coates and Humphreys (2011) claimed that the previous inconsistent results were due to linear or quadratic specifications of match uncertainty; their results for the NHL indicate that the attendance increases if the home team is a strong favorite or a slight underdog.

This article makes two contributions. First, three simple simulated data sets with no impact of match uncertainty on attendance are used to show that many commonly used regression specifications produce different (and wrong) results about the link between match uncertainty and attendance. This could explain the inconsistent findings in the literature, especially if the actual impact of match uncertainty is weak or nonexistent. Second, a new approach to analyzing the effect of match uncertainty on attendance is proposed. Using data about nine seasons of the English Championship, the article shows that in a pair of matches where both home teams are slight favorites, a switch of the corresponding away teams would decrease the total attendance. On the other hand, if both home teams are underdogs or strong favorites, switching the away teams would increase the total attendance. However, the magnitude of such attendance changes is quite small (several percent). These results are consistent with the uncertainty of outcome hypothesis and suggest that attendance demand is a bell-shaped function of match balance that is maximized if teams of the same quality play against each other.

3.2 Data

The proposed method of measuring the impact of match uncertainty on attendance is demonstrated on the data set consisting of nine regular seasons (2004/05-2012/13) of the second-highest English soccer league; English Championship. This competition still attracts a lot of spectators, but attendances only rarely come close to the stadium capacity, so the attendance demand for each match is directly observable.

In each season of the Championship, 24 teams play one home and one away match against each other, so there are 552 matches in each season and 4,968 matches in the whole Championship data set. The relevant data for each match are its attendance, which was downloaded from the website worldfootball.net, and the corresponding betting odds, which were obtained from the website football-data.co.uk.[57] The betting odds were converted in a standard way to home win, draw and away win probabilities and these probabilities were averaged across different bookmakers.[58]

For each match, match balance was calculated as the home win probability plus one half of the draw probability; this variable is similar to home win probability used in many previous articles, but has the advantage of being exactly 0.5 for perfectly balanced matches with each team having the same probability of winning. The descriptive statistics for variables *Attendance* and *MatchBalance* are provided in Table 15.

[57] The attendance data were downloaded on July 7th, 2013. One missing attendance figure was obtained from the website www.11v11.com. The betting odds on home win, draw, and away win were downloaded on June 10th, 2013, and provided by major bookmakers William Hill, Bet&Win, and Interwetten. Although some betting odds were missing, there was at least one set of betting odds for each match.

[58] To convert betting odds into probabilities, they are first inverted. The sum of these inverted numbers is more than one to allow for bookmaker's profit, so the inverted numbers are divided by this sum to obtain the home win, draw, and away win probabilities.

	Average	Min	Percentiles					Max
			0.1	0.25	0.5	0.75	0.9	
Attendance	17,632	1,211	9,492	12,822	17,219	22,267	25,652	52,181
MatchBalance	0.5743	0.2641	0.4758	0.5182	0.5699	0.6340	0.6806	0.8083

Table 15: Descriptive statistics of the Championship data set, N = 4,968

Both the average and the median match balance values are close to 0.57; this number is higher than 0.5 because of the home team advantage. Most match balance values (80%) are concentrated in the 0.48-0.68 interval, which means that it is hard to say much about what happens to match attendance outside this interval.

3.3 The pitfalls of using the regression approach

This section shows that many commonly used regression specifications produce misleading results about the relationship between match uncertainty and attendance. This is demonstrated on three simple simulated data sets with no impact of match uncertainty on attendance.

To construct each data set, let's assume there are 24 teams in a competition (the same as in the Championship data set) and team qualities are uniformly distributed on the interval [0; 1]. This means that team *i*'s quality $Quality_i = (i - 1)/23$; the *Quality* variable corresponds to the normalized rank or points per game used in other studies. All teams play one home and one away match against each other, generating one complete season of 552 matches. Let's further assume that each team attracts a fixed number of spectators to its home matches and that there is a different (smaller) fixed number of spectators that travel with the team to its away matches. Both numbers are increasing functions of team quality.[59]

There is no special reason why the relationship between team quality and the number of spectators should be linear; in fact, an obvious non-linear relationship between team rank and points per game guarantees that it is not the case in at least some previous studies. The data sets

[59] A similar assumption was made in Peel and Thomas (1992). Their estimation results also show that match attendance increases with qualities of both teams (measured by team ranks) with the home team's quality having a stronger influence.

cover the three simplest cases: attendance is an exactly linear function of quality (Data set 1), attendance is a concave function of quality (Data set 2), and attendance is a convex function of quality (Data set 3). Therefore, the first data set satisfies the linearity assumption, while the other two data sets represent two simplest deviations from this assumption.

To produce plausible total attendance numbers (similar to the Championship data set), the home and away spectator numbers $Home_i$ and $Away_i$ attracted by team *i* are set equal to the following expressions:

Data set 1 (linear): $Home_i = 5{,}000 + 25{,}000 * Quality_i$, $Away_i = 5{,}000 * Quality_i$

Data set 2 (concave): $Home_i = 5{,}000 + 25{,}000 * Quality_i^{0.8}$, $Away_i = 5{,}000 * Quality_i^{0.8}$

Data set 1 (convex): $Home_i = 5{,}000 + 25{,}000 * Quality_i^{1.25}$, $Away_i = 5{,}000 * Quality_i^{1.25}$

The attendance of a match between teams *i* and *j* ($Attendance_{ij}$) is simply the sum of spectators attracted by the home team ($Home_i$) and spectators travelling with the away team ($Away_j$):

Data set 1 (linear): $Attendance_{ij} = 5{,}000 + 25{,}000 * Quality_i + 5{,}000 * Quality_j$

Data set 2 (concave): $Attendance_{ij} = 5{,}000 + 25{,}000 * Quality_i^{0.8} + 5{,}000 * Quality_j^{0.8}$

Data set 1 (convex): $Attendance_{ij} = 5{,}000 + 25{,}000 * Quality_i^{1.25} + 5{,}000 * Quality_j^{1.25}$

As stated above, fans care only about the quality of their own team, so there is no causal relationship between match uncertainty and attendance. This is also clear from the attendance formulas – they are additively separable (there is no interaction between qualities of both teams). Therefore, any valid method of measuring the impact of match uncertainty on attendance should conclude that the impact is zero. To test whether this is true for common regression specifications, a variable $MatchBalance_{ij}$ is defined in the following way for each match using the logistic function:

$$MatchBalance_{ij} = 1/(1 + exp(Quality_j – Quality_i – 0.25))$$

The match balance variable is an increasing function of home team's quality and a decreasing function of away team's quality, so a higher value indicates that the home team is more likely to win (the number 0.25 provides the home advantage). The match balance values in each simulated data set range from 0.32 to 0.78 with the average of 0.56, closely mimicking the match balance values in the Championship data set that were calculated as the home win probability plus one half of the draw probability. The logistic form guarantees that the match balance variable is always between 0 and 1 for any possible difference in team qualities.

An additional variable $MatchUncertainty_{ij}$ measures how close a specific match is to the balance of 0.5 (i.e. both teams being equally likely to win):

$$MatchUncertainty_{ij} = 1 - 2 * |MatchBalance_{ij} - 0.5|$$

If a match is perfectly balanced, $MatchUncertainty_{ij}$ equals 1; on the other hand, if one team is sure to win, $MatchUncertainty_{ij}$ goes down to 0. This variable is analogical to variables such as the difference in team ranks, difference in points per game (possibly adjusted for home team advantage), or the absolute value of betting spread used in other studies.

Researchers using a regression approach to measure the impact of match uncertainty on attendance choose from various regression specifications. Probably the simplest one copies the attendance-generating formula of Data set 1 and adds the *MatchUncertainty* variable:

$$Attendance_{ij} = \beta_0 + \beta_1 * Quality_i + \beta_2 * Quality_j + \beta_3 * MatchUncertainty_{ij} + \varepsilon$$

This simple regression specification can be modified by replacing attendance with its logarithm (since many variables are expected to influence attendance by a given percentage instead of by a given number of spectators); by replacing home team quality with a set of dummies for each home team (home fixed effects); by also replacing away team quality with a set of dummies (all fixed effects); or by replacing the *MatchUncertainty* variable with a quadratic specification of the *MatchBalance* variable (i.e. $\beta_3 * MatchBalance_{ij} + \beta_4 * MatchBalance_{ij}^2$). The estimated effects of match uncertainty on attendance for all three simulated data sets using twelve possible regression specifications are summarized in Table 16.[60]

[60] All models were estimated with heteroskedasticity-consistent standard errors. Results are reported as significant if p-value < 0.05.

	Data set 1 (linear)	Data set 2 (concave)	Data set 3 (convex)
Attendance No fixed effects Uncertainty	Zero effect	Higher uncertainty increases attendance	Higher uncertainty decreases attendance
Log of attendance No fixed effects Uncertainty	Insignificant	Higher uncertainty increases attendance	Higher uncertainty decreases attendance
Attendance Home fixed effects Uncertainty	Zero effect	Higher uncertainty increases attendance	Higher uncertainty decreases attendance
Log of attendance Home fixed effects Uncertainty	Higher uncertainty decreases attendance	Higher uncertainty decreases attendance	Higher uncertainty decreases attendance
Attendance All fixed effects Uncertainty	Zero effect	Zero effect	Zero effect
Log of attendance All fixed effects Uncertainty	Higher uncertainty decreases attendance	Higher uncertainty decreases attendance	Higher uncertainty decreases attendance
Attendance No fixed effects Quadratic balance	Zero effect	Attendance maximized if balance = 0.41	Attendance minimized if balance = 0.54
Log of attendance No fixed effects Quadratic balance	Insignificant	Insignificant	Attendance minimized if balance = 0.15
Attendance Home fixed effects Quadratic balance	Zero effect	Attendance maximized if balance = 1.30	Attendance minimized if balance = 1.01
Log of attendance Home fixed effects Quadratic balance	Attendance minimized if balance = 0.18	Attendance minimized if balance = 0.11	Attendance minimized if balance = 0.23
Attendance All fixed effects Quadratic balance	Zero effect	Zero effect	Zero effect
Log of attendance All fixed effects Quadratic balance	Attendance minimized if balance = 0.25	Attendance minimized if balance = 0.25	Attendance minimized if balance = 0.25

Table 16: Estimated effect of match uncertainty on attendance for simulated data sets

Since there is no actual relationship between match uncertainty and attendance in the simulated data sets, most regression results are incorrect. This is caused by a not exactly linear relationship between attendance (or its logarithm) and team quality variables; fitting a linear model leads to a specific pattern of residuals that is then captured by match uncertainty or balance variables (which are themselves determined by team qualities). The only specification that produces correct results for all three data sets includes fixed effects for both home and away teams and a non-logarithmic attendance. On the other hand, the logarithmic form of attendance mostly leads to the conclusion that higher uncertainty decreases attendance or that attendance is a convex function of balance (minimized for some specific balance value). This could explain similar surprising results in the literature, such as Pawlowski and Anders (2012), Buraimo and Simmons (2008), or Buraimo and Simmons (2009), especially if the actual impact of match uncertainty on attendance is weak or nonexistent.[61]

Although the specification with fixed effects for both home and away teams and a non-logarithmic attendance provides correct results for all three simulated data sets, it is still problematic, because it imposes a specific functional form on the relationship between match balance and attendance. There is no theoretical reason why this relationship should be linear or quadratic – Coates and Humphreys (2011) hypothesized an asymmetric relationship (fans preferring matches where home teams are strong favorites or slight underdogs), but it could easily be S-shaped (most fans do not care about match uncertainty, but some fans will attend a match only if the home team is sufficiently favored) or bell-shaped (most fans do not care about match uncertainty, but some fans will attend a match only if it is balanced enough). This could be fixed by using a nonparametric estimation, ranging from dummy variables for various match balance intervals to LOESS (local regression). However, these approaches generally require much bigger data sets to get sufficiently precise estimates.

[61] Of course, no result reported in the literature can be exactly the same as the corresponding cell in Table 16. First, there is no causal relationship between match uncertainty and attendance in the simulated data sets, but there is likely to be some relationship in reality. Second, the nonlinear relationship between team quality and attendance is likely to be more complex than just a simple convex or concave exponential function. Third, attendance is completely deterministic in the simulated data sets, but stochastic in reality.

Even when using a correct regression specification, results are hard to interpret; they do not say how attendance would change if match uncertainty changed and all the other variables stayed constant, since match uncertainty cannot change without also changing team qualities. Some authors address this by taking the estimated attendance demand function and asking what would happen if the league structure changed (Dobson et al., 2001) or if teams were more evenly balanced (Forrest and Simmons, 2002; Buraimo and Simmons, 2009). Such simulations can be very complicated, because researchers cannot simply change a value of one variable (e.g. team quality or match balance), but have to generate different values for all other related variables.

As shown above, commonly used simple regression specifications lead to different (and mostly incorrect) conclusions about the relationship between match uncertainty and attendance. The variety of results reported in Table 16 is in fact similar to the variety found in the literature for real data. Therefore, the inconsistent findings in the literature could easily be caused by misspecified regressions, especially if the actual impact of match uncertainty on attendance is weak or nonexistent.

Clearly, a different approach could be useful. Such an approach should fulfill three criteria; first, it should not find any effect of match uncertainty on attendance in any of the simulated data sets; second, it should not assume any specific functional form of the match uncertainty-attendance relationship; third, its results should be easy to interpret. Exactly such an approach is described in the next section.

3.4 A non-regression approach

This section presents an approach to examining the link between match uncertainty and attendance that does not use a regression and whose results are easy to interpret. The proposed method is demonstrated on the Championship data set described above. The main idea is that although it is not possible to change uncertainty of a match between two fixed teams, it is possible to change uncertainties of matches between two fixed sets of teams by pairing them in different ways and then analyze what happens to the total attendance of such match combinations.

The approach (as applied to the Championship data set) starts with all combinations of two home teams (H_1, H_2) and two away teams (A_1, A_2) in a given season. Since there are 24 teams in a season, there are 24 * 23/2 possible home team pairs and 22 * 21/2 possible away team pairs, giving 24 * 23 * 22 * 21/4 = 63,756 combinations per season and 63,756 * 9 = 573,804 combinations for the whole nine-season data set. Each combination of two home teams and two away teams can be matched in two different ways: H_1-A_1 + H_2-A_2 or H_1-A_2 + H_2-A_1. Those two possible pairs of matches will be different in terms of both match balance and total attendance.

If H_1 and A_1 are similarly strong teams and H_2 and A_2 are similarly weak teams, it is interesting to ask whether a pair of matches where similar teams play against each other (H_1-A_1 + H_2-A_2) tends to have a higher total attendance than a pair of matches where teams of opposite strengths play against each other (H_1-A_2 + H_2-A_1). To answer this question, only those combinations of home and away teams are selected where balances of both matches in one pair (called the balanced pair) are close to 0.57 (both the average and the median value of match balance in the data set), while the one match balance in the second pair (called the unbalanced pair) is much higher than 0.57 and the other is much lower than 0.57. More formally, the conditions for selecting a combination of match pairs are the following (α and β are parameters, $\beta \geq \alpha > 0$):

Balanced pair: $0.57 - \alpha \leq \textit{Both match balances} \leq 0.57 + \alpha$

Unbalanced pair: $\textit{One match balance} < 0.57 - \beta < 0.57 + \beta < \textit{The other match balance}$

Decreasing α and increasing β creates a bigger contrast between the pairs of matches, but decreases the number of combinations of match pairs that are selected, so the exact values should be chosen depending on the size of the data set. The specific values of α and β chosen for the Championship data set are $\alpha = 0.03$ and $\beta = 0.09$, leading to 4,075 combinations of match pairs selected for further analysis with one such combination from the 2012/13 season provided as an example in Table 17.

	Match	Match balance	Attendance	Total attendance
Balanced pair	Leicester – Watford	0.5958	25,091	46,802
	Wolves – Bristol	0.5574	21,711	
Unbalanced pair	Leicester – Bristol	0.7184	22,529	41,100
	Wolves – Watford	0.4455	18,571	

Table 17: An example of match balances and attendances in two possible match pairs

In this example, Leicester (ultimately finished 6th) and Watford (finished 3rd) are relatively strong teams, while Wolves (finished 23rd) and Bristol (finished 24th) represent relatively weak teams. In the balanced pair, similarly-strong teams play against each other and both match balances are between 0.54 (0.57 – 0.03) and 0.60 (0.57 + 0.03), while in the unbalanced pair, the away teams are switched, one match balance is above 0.66 (0.57 + 0.09), and the other match balance is below 0.48 (0.57 – 0.09). In the example, the actual total attendance of the balanced pair was higher than the total attendance of the unbalanced pair, indicating that if teams of the same strength play against each other, the total attendance is higher.

Looking at all 4,075 similar combinations, the total attendance of the balanced pair was higher in 2,349 cases (57.64%) and lower in 1,726 cases (42.36%).[62] However, the total attendance of all balanced pairs was just 1.5% higher than the total attendance of unbalanced pairs.

How to interpret these findings? For all three completely deterministic data sets introduced in the previous section, switching away teams would keep the total attendance exactly the same. In reality, attendance also depends on many other factors, but without a causal link between match uncertainty and attendance, switching away teams would still have no systematic impact on total attendance and the probability that the balanced pair is more attended would be 50%. Because the 57.64% result for the balanced pair is statistically significantly different from 50% ($p = 0.002$),[63]

[62] There was no case of both total attendances being exactly the same; such a case could be dropped when calculating the percentages.

[63] To obtain the p-value, the whole test described above was applied to 50,000 versions of the original data set. In each version, actual attendances were replaced with different random values drawn from the standard normal distribution (representing other factors influencing attendance besides match uncertainty). In only 118 cases (0.2%), the total attendance of one pair type (balanced or unbalanced) was higher in at least 57.64% of all cases.

it can be concluded that match uncertainty indeed influences attendance; however, the magnitude of the effect seems to be quite small.

The findings above also tell us something about the shape of the relationship between match balance and attendance;[64] if the balanced pair tends to have a higher attendance than the unbalanced pair and the average match balance in both pairs is about the same (true in the data), attendance should be a concave function of match balance for match balance values around 0.57. To infer the shape of the relationship in other regions, the whole test above can be repeated for any match balance value (denoted *TestedBalance*) different from 0.57. Only the combinations of match pairs are selected that fulfil the following conditions:

Balanced pair: *TestedBalance* – $\alpha \leq$ *Both match balances* $\leq$ *TestedBalance* + α

Unbalanced pair: *One match balance* < *TestedBalance* – β < *TestedBalance* + β < *The other match balance*

Again, it is possible to calculate the proportion of combinations where the total attendance of the balanced pair was higher. The results for *TestedBalance* values between 0.48 and 0.68 are provided in Figure 6.[65]

[64] In the three simulated data sets, a relationship between match balance and attendance would be represented by adding a function of $MatchBalance_{ij}$ variable into the attendance equation, for example $Attendance_{ij} = 5{,}000 + 25{,}000 * Quality_i + 5{,}000 * Quality_j + f(MatchBalance_{ij})$.

[65] The 0.48 – 0.68 interval ranges from the 10^{th} to the 90^{th} percentile of match balance values in the Championship data set. The test was done for all *TestedBalance* values that are multiples of 0.001, i.e. 0.480, 0.481, 0.482 … 0.679, 0.680. The number of usable team combinations for each *IdealBalance* value is always more than 500 (more than 3,000 on average). The α and β parameters were kept the same as in the previous test, i.e. $\alpha = 0.03, \beta = 0.09$.

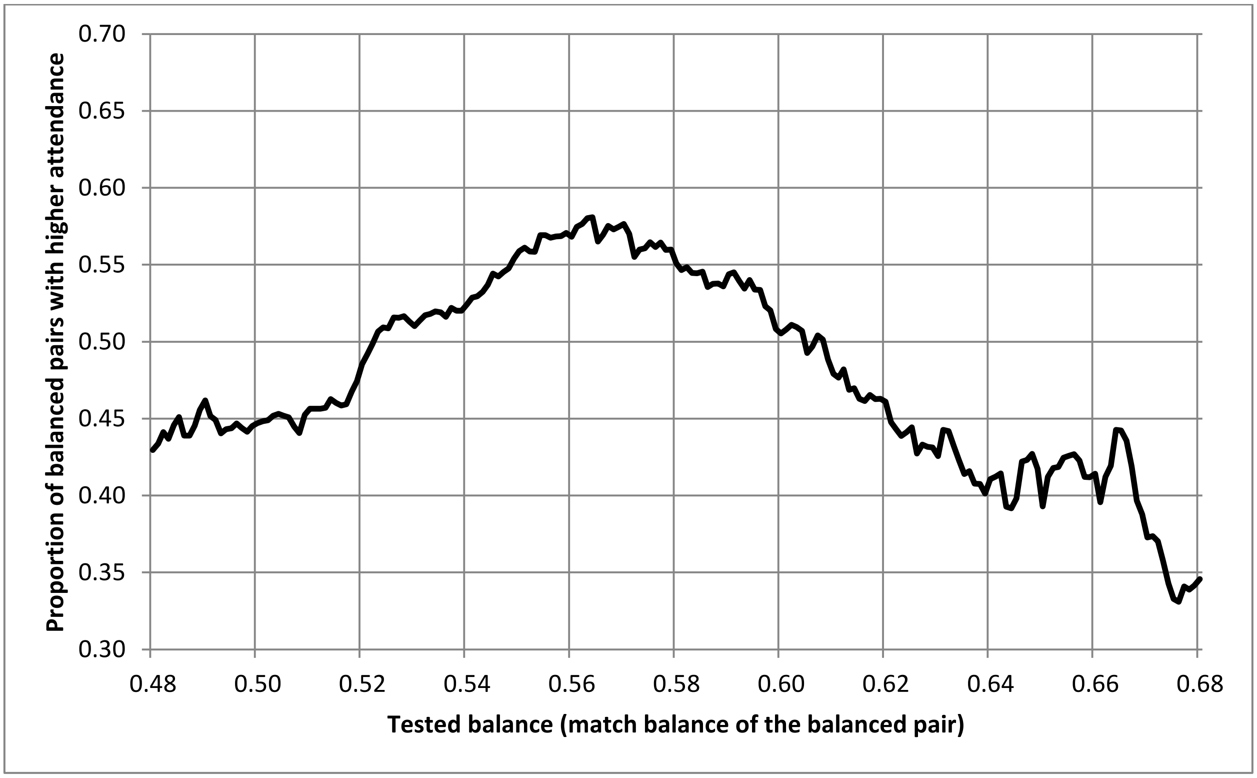

Figure 6: Proportion of balanced pairs with higher attendance for various TestedBalance values

The graph shows that if the match balance in the balanced pair is between 0.53 and 0.61 (i.e. teams of the same quality play against each other), switching the away teams that would unbalance both matches would tend to decrease the total attendance. Therefore, the relationship between match balance and attendance should be concave in this range. On the other hand, if both matches in a pair have approximately the same balance that lies outside the 0.53 – 0.61 interval (i.e. both home teams are underdogs or strong favorites), an away-team switch that would unbalance both matches would increase the total attendance. The relationship between match balance and attendance should be concave in these regions. In all cases, the average attendance change would be several percent at most. The probability that the balanced pair of matches is more attended is highest for match balance values close to 0.57, i.e. when teams of similar strengths play against each other. Therefore, a plausible relationship between match balance and attendance consistent with these results would be approximately bell-shaped, maximized at match balance around 0.57, and having inflection points close to 0.53 and 0.61.[66]

3.5 Discussion

This article has showed that many regression specifications commonly used in the literature can produce misleading results about the link between match uncertainty and attendance. Even if the regression equation is correctly specified, the results are hard to interpret, because it is not possible to change match balance without also changing team qualities.

After that, the article has proposed a new approach to examining the link between match uncertainty and attendance that does not rely on regression. Unlike commonly used regression specifications, the proposed method correctly does not find any link between match uncertainty and attendance if the attendance demand is an additively separable function of team qualities (such as in the three simulated data sets). The non-regression approach also does not assume any specific functional form of the match uncertainty-attendance relationship.

Using data about nine seasons of the English Championship, the proposed method shows that in a pair of matches where both home teams are slight favorites, switching the away teams would decrease the total attendance, while the opposite is true if both home teams are underdogs or strong favorites. However, the impact of such team switches on attendance is just several percent at most. The results are consistent with the uncertainty of outcome hypothesis and suggest that attendance demand is a bell-shaped function of match balance that is maximized if teams of the same quality play against each other (in such matches, home teams are slightly favored due to home advantage). One possible explanation of such a shape could be that there are two groups of potential spectators with different preferences; fans in the first group (seasonal ticket holders, hardcore fans) do not care about match uncertainty and attend all matches if they have free time

[66] The results are robust to changing α and β (e. g. $\alpha = 0.04$, $\beta = 0.06$), discarding team combinations where the change in total attendance between match pairs is too low (e.g. less than 10%), discarding team combinations where the difference between attendances of balanced pair matches is too high (e.g. more than 10 or 20%), or restricting the analysis to a subset of seasons.

and no better opportunities, while fans in the second group (occasional spectators) choose to attend only the most interesting matches with one criterion being a proper match balance.

The above results can be directly applied to tournament design; to increase the total attendance of a competition while keeping the number of home and away matches of each team constant, a higher proportion of matches should be played between evenly matched teams. This could be achieved by splitting teams into groups based on team quality instead of on region[67] or by making the tournament design more similar to the Swiss system commonly used in chess. However, the potential attendance increase would likely be small.

Both the small effect size and the bell shape of the attendance demand function further support the claim that some inconsistent results in the previous research could have been caused by misspecified regressions. However, the proposed non-regression approach also has two limitations that result from using historical attendance figures to estimate attendance demand. First, the approach assumes that attendance demand is actually observable and not right-censored (i.e. the stadium is not close to capacity). This is not a problem in the Championship data set, but would be a problem in the Premier League or other top European soccer competitions. A possible solution would be to discard such team combinations where any attendance is close to the corresponding stadium capacity. Second, the non-regression method assumes that other factors influencing match attendance besides team qualities (e.g. day of the week, TV broadcast, distance between teams, weather, or match importance) are not strongly correlated with match balance. If this assumption does not hold, the results will be biased.

The variable most likely to be correlated with match balance is match importance, i.e. how much a given match result influences the probability of a given season outcome, such as promotion or relegation. If matches between equal-quality teams tend to be more important and higher importance increases attendance, the higher attendance of matches between equal-quality teams would be partly explained by match importance, not by match uncertainty, and the actual effect size would be even lower (the regression approach suffers from the same omitted variable bias

[67] On the other hand, splitting teams into regional groups would lead to a higher proportion of matches between regional rivals and lower travelling distance between teams. Both of these factors tend to increase attendance (García and Rodríguez, 2009).

problem if match importance is not properly controlled for). Again, a possible solution would be to throw away such team combinations where any match is above some importance threshold.[68] For competition organizers, the distinction between match importance and match uncertainty might not even be relevant; if more matches between equal-quality teams increase attendance, the exact mechanism does not matter.

There are several possible avenues of further research. First, the results presented in this article are for one specific competition, so the proposed non-regression approach should be applied to soccer competitions in different countries and to different sports to see whether the results stay the same. Second, more attention should be paid to preference-revealing fan behavior during the match. Anecdotally, fans start leaving the stadium prematurely if the score difference is big, especially if the home team is badly losing. The article by Tainsky et al. (2013) is a nice example of this approach applied to TV ratings of NCAA football. Third, fan preferences could be revealed in short series of matches that are essentially one longer match, such as those in the NHL playoffs or European soccer cups; a lower attendance when one team is practically sure to advance to the next round would confirm the uncertainty of outcome hypothesis. Fourth, fans could simply be asked about their preferences related to individual match uncertainty similarly to the stated preferences approach applied to the overall competitive balance by Pawlowski and Budzinski (2013) and Pawlowski (2013).

[68] An overview and comparison of methods for calculating match importance can be found in the first article in this dissertation called “Using Monte Carlo simulation to calculate match importance: The case of English Premier League.” Relatively simple solutions for dealing with match importance would be to throw away the second half of each season or all team combinations where at least one team was ultimately promoted or relegated. The latter modification applied to the Championship data set does not substantially change the original results.

4 What causes the favorite-longshot bias? Further evidence from tennis[69]

In sports betting markets, bets on favorites tend to have a higher expected value than bets on longshots. This article uses a data set of almost 45,000 professional single tennis matches to show that the favorite-longshot bias is much stronger in matches between lower-ranked players, in later-round matches, and in high-profile tournaments. These results cannot be solely explained by bettors being locally risk-loving or overestimating chances of longshots, but are consistent with bookmakers protecting themselves against both better informed insiders and the general public exploiting new information.

4.1 Introduction

In sports betting markets, bets on favorites usually have a higher expected value (lose less money) than bets on longshots (Sauer, 1998; Cain et al., 2003; Direr 2013). There are three types of explanations for this so-called favorite-longshot bias (Snowberg and Wolfers, 2010; Makropoulou and Markellos, 2011; Rossi, 2011). The first explanation claims that bettors are local risk-lovers and bookmakers take advantage by lowering the odds on longshots. According to the second explanation, bettors overestimate winning probabilities of longshots and bookmakers again take advantage of this psychological bias. The third explanation is based on information asymmetry; bookmakers could potentially lose a lot of money if they underestimate longshots and this mispricing is exploited by either better informed insiders or by the general

[69] A shorter version of this article (without the appendix) was published in Applied Economics Letters, 2014, Volume 21, Issue 2, pp. 90-92, doi: 10.1080/13504851.2013.842628.

public reacting faster than bookmakers to new information. Therefore, bookmakers offer lower odds on longshots to protect themselves against this type of loss.

To distinguish between these competing explanations, this article uses a data set of almost 45,000 professional single tennis matches to show that the favorite-longshot bias is much more pronounced in matches between lower-ranked players, in later-round matches, and in high-profile tournaments. These results, as discussed later, are consistent with the information asymmetry explanation. The favorite-longshot bias in tennis was already analyzed by Forrest and McHale (2007), but they had a much smaller data set, did not test the effect of players' ranks or tournament round, and did not find any difference for high-profile tournaments.

4.2 Data

The data set consists of results of 44,871 professional men's and women's single tennis matches with valid betting odds.[70] The decimal betting odds on each player's win were converted to implied probabilities of winning in the standard way by calculating their inverse values. Since the two resulting numbers for each match add up to more than one to allow the bookmaker to have profit, they have to be both divided by their sum. Because the two possible bets on each match are not independent (implied probabilities add up to one, exactly one bet pays off), only one (chosen randomly) is included in the final data set. Therefore, there are 44,871 observed bets with an implied probability of the player winning (the variable *ImpliedProbability*) and a corresponding match result (the variable *Result* that equals one if the player won and zero if the player lost).

To test how the favorite-longshot bias differs across various types of matches, the following dummy variables are defined: *LowerRank* equals one in 12,878 matches where both players were outside of top 50 in ATP/WTA rankings, zero otherwise; *LaterRound* equals one in 24,189

[70] The data set was downloaded from the website tennis-data.co.uk on June 22nd, 2013. The men's tennis matches start in 2002; the women's matches start in 2007. The betting odds are the latest available odds by the bookmaker Bet365. Originally, there were 48,042 matches, but 3,171 matches (6.6%) were discarded due to missing odds or a withdrawal of one player before the match started.

matches that were not in the first round (lowest round in the data set), zero otherwise; and *HighProfile* equals one in 8,962 matches in a high-profile tournament (Grand Slam, ATP World Tour Finals, or WTA Tour Championships), zero otherwise.

4.3 Model and results

To test whether the favorite-longshot bias exists in the market as a whole, the following standard linear probability model is employed:

$$Result = \beta_0 + \beta_1 * ImpliedProbability + \varepsilon$$

In the absence of bias (null hypothesis), the coefficient values would be $\beta_0 = 0$ and $\beta_1 = 1$, while the standard favorite-longshot bias would be indicated by $\beta_0 < 0$ and $\beta_1 > 1$. The estimation results[71] in Table 18 show that the favorite-longshot bias is indeed present in the investigated data set; the winning probability implied by the betting odds is higher than the actual probability in case of longshots and lower than the actual probability in case of favorites.

	Coefficient	Standard error
Constant	-0.0293***	0.0044
ImpliedProbability	1.0594***	0.0077

Table 18: The Favorite-Longshot Bias in the Whole Market, N = 44,871

To investigate whether the favorite-longshot bias differs across various types of matches, the model is expanded in the following way:

$$Result = \beta_0 + \beta_1 * ImpliedProbability + \beta_2 * LowerRank + \beta_3 * LowerRank * ImpliedProbability + \beta_4 * LaterRound + \beta_5 * LaterRound * ImpliedProbability + \beta_6 * HighProfile + \beta_7 * HighProfile * ImpliedProbability + \varepsilon$$

[71] The estimation method in the whole article is OLS with heteroskedasticity-robust standard errors. One star indicates p-value < 0.1, two stars p-value < 0.05, three stars p-value < 0.01.

In case of no difference among various types of matches (the null hypothesis), $\beta_2 \ldots \beta_7 = 0$, while the overall bias would still be captured by $\beta_0 < 0$ and $\beta_1 > 1$. The estimation results for the expanded model are presented in Table 19.

	Coefficient	Standard error
Constant	0.0032	0.0085
ImpliedProbability	1.0051	0.0148
LowerRank	-0.0539***	0.0132
LowerRank * ImpliedProbability	0.0918***	0.0244
LaterRound	-0.0233**	0.0092
LaterRound * ImpliedProbability	0.0358**	0.0162
HighProfile	-0.0361***	0.0094
HighProfile * ImpliedProbability	0.0704***	0.0160

Table 19: The Favorite-Longshot Bias across Various Types of Matches, N = 44,871

The coefficients show the favorite-longshot bias is much stronger in matches between lower-ranked players, in later-round matches, and in matches in high-profile tournaments, while it is practically nonexistent in the other matches. These results are robust across different model specifications and data subsamples (see Appendix). They have also been confirmed by comparing average implied probabilities with relative frequencies of winning over different probability ranges for different types of matches (similarly to Forrest and McHale 2007). A graphical analysis also confirms that the relationship between the implied and actual probability of a win is approximately linear.

4.4 Discussion

The results seem to be contradictory; on the one hand, the favorite-longshot bias is stronger in later-round matches and in matches in high-profile tournaments, i.e. in matches that are likely to attract high betting volumes; on the other hand, the favorite-longshot bias is also more pronounced in matches between lower-ranked players, which are likely to exhibit low betting volumes. This pattern cannot be explained solely by people being local risk-lovers or overestimating chances of longshots; if all bettors had the same preferences or biases, the type of match should not matter at all. Even if the risk-loving preferences (or the corresponding bias) were exhibited only by occasional bettors, thus causing the stronger favorite-longshot bias in matches that are likely to attract high betting volumes, it would not explain why the bias is also more pronounced in matches between lower-ranked players. Therefore, at least one part of the explanation must lie in the information asymmetry.

Forrest and McHale (2007) argued that in Grand Slam tournaments, players are more motivated and less likely to underperform, so the role of private information should be much smaller. Consequently, if the favorite-longshot bias was a defense of bookmakers against better informed insiders, it should be smaller in high-profile tournaments. However, according to the results in this article, the bias is actually larger. This is hard to explain as a defense against insider trading; besides players being more motivated, the proportion of insiders among all bettors is also likely to be smaller, not larger, in high-profile tournaments.

The most plausible explanation of the results seems to be a combination of two information asymmetry approaches: Matches between lower-ranked players are harder to predict, since public information is limited and private information about players' motivation or health problems could play a large role; therefore, it makes sense for the bookmaker to set lower odds on the longshot to minimize possible losses. On the other hand, private information should not play such a big role in later tournament rounds and high-profile tournaments, but in such matches the bookmaker faces a different kind of risk; the general public could react faster than the bookmaker to newly available information. Combined with a high volume of bets, this could mean a considerable loss, so the bookmaker again protects itself by setting lower odds on the longshot.

Of course, the information asymmetry explanation does not rule out that the other alternatives, i.e. risk-loving preferences or overestimating small probabilities of winning, also play a role. Clearly, more research is needed. One possible direction would be to test more thoroughly whether the stronger favorite-longshot bias in high-profile tournaments also exists in other individual or team sports (or even in tennis doubles); if the above explanation is correct, the effect in team sports should be smaller, since the impact of new information (e.g. a minor sickness of a player) is likely to have less influence on the expected result.

4.5 Appendix

This appendix compares the estimation results for the expanded model on the whole data set (already presented in Table 19 above) with estimation results for three specific subsamples. The first subsample consists of only men's matches, the second subsample includes only women's matches, and the third subsample is the full data set without ATP World Tour Finals and WTA Tour Championships. The estimation results are summarized in Table 20.

	All data	Men only	Women Only	No tour finals
Constant	0.0032 (0.0085)	0.0029 (0.0108)	-0.0104 (0.0135)	-0.0029 (0.0083)
ImpliedProbability	1.0051 (0.0148)	1.0045 (0.0190)	1.0057 (0.0237)	1.0048 (0.0149)
LowerRank	-0.0539*** (0.0132)	-0.0544*** (0.0166)	-0.0376* (0.0216)	-0.0469*** (0.0132)
LowerRank * ImpliedProbability	0.0918*** (0.0244)	0.0988*** (0.0309)	0.0799** (0.0398)	0.0919*** (0.0244)
LaterRound	-0.0233** (0.0092)	-0.0307*** (0.0116)	-0.0019 (0.0149)	-0.0136 (0.0091)
LaterRound * ImpliedProbability	0.0358** (0.0162)	0.0500** (0.0206)	0.0101 (0.0263)	0.0365** (0.0162)
HighProfile	-0.0361*** (0.0094)	-0.0425*** (0.0117)	-0.0192 (0.0156)	-0.0388*** (0.0093)
HighProfile * ImpliedProbability	0.0704*** (0.0160)	0.0784*** (0.0200)	0.0589** (0.0267)	0.0712*** (0.0161)
Number of observations	44,871	29,136	15,735	44,624

Table 20: The Favorite-Longshot Bias across Different Subsamples

In all cases, the coefficients of *LowerRank*, *LaterRound*, and *HighProfile* variables are negative, while the coefficients of these variables interacted with *ImpliedProbability* are positive. Although the coefficients for only women's matches seem a bit closer to zero than those for only men's matches, there is no statistically significant difference ($p < 0.05$) in the values of any specific coefficient between these two subsamples. Therefore, the results reported in the article are not specific to just one type of tennis matches.

5 The Fibonacci strategy revisited: Can you really make money by betting on soccer draws?[72]

This article investigates the strategy of betting on soccer draws using the Fibonacci sequence. In the previous literature, this strategy has been found to be both simple and profitable, indicating that the soccer betting market is not efficient. The strategy is tested both in a simulated market and on a real data set of almost 60,000 European soccer matches. Contrary to the previous findings, all tested versions of the Fibonacci betting strategy are found to lose money.

5.1 Introduction

When investigating market efficiency, economists often turn to sports betting markets, since each asset (placed bet) has a certain value at a specific time (after the match). There are two types of efficiency typically studied in sports betting markets – strong and weak efficiency (Thaler and Ziemba, 1988). In a strongly efficient market, each bet has the same negative expected value – for example, a $1 bet on any match result can be expected to pay back just 90 cents. In a weakly efficient market, bets might have different expected values, but these are still always negative.

There is ample evidence that sports betting markets are not strongly efficient – for example, bets on favorites and home teams lose less money than bets on longshots and away teams (Sauer, 1998). Some authors also claim to have found profitable strategies, mostly for betting on European soccer (e.g. Kuypers, 2000; Goddard and Asimakopoulos, 2004; Vlastakis et al., 2009),

[72] This article was accepted for publication in the Journal of Gambling Business and Economics.

but these strategies usually rely on hard-to-implement models and identify only a small number of profitable betting opportunities. One notable exception is the Fibonacci betting strategy, first proposed by Archontakis and Osborne (2007), which is claimed to be both simple and profitable, although risky.

The Fibonacci betting strategy is designed for betting on soccer results. It is based on the Fibonacci sequence (1, 1, 2, 3, 5, 8, 13…), where the first two numbers equal one and each successive number is the sum of the two previous numbers. The strategy works as follows: bet $1 (the first number in the sequence) on a draw, if losing, bet $1 (the second number) on a draw in the next match, if losing again, bet $2 (the third number) on a draw in the next match, and so on until a draw actually occurs; after that, start the whole sequence from beginning. Archontakis and Osborne (2007) proved that each sequence of bets ending in a draw is profitable if draw odds are always at least 2.618 (usually true). The authors also tested the Fibonacci strategy on 32 games in 2002 FIFA World Cup and found that it would have generated a profit.

The Fibonacci betting strategy was later tested by Demir et al. (2012) on a sample of 32 seasons of top European soccer competitions and found profitable in all 32 cases. The strategy was also found to be profitable in a simple simulated strongly efficient market using 1,000 simulations. The authors characterize the Fibonacci betting strategy as “simple and profitable” (p. 30), but requiring a lot of capital if draws fail to occur for a long time.

This article first investigates the behavior of the proposed strategy in a simulated strongly efficient market and shows that it actually is not and cannot be profitable in such a market. However, under certain conditions the strategy could still be profitable in a real market, so it is tested on a data set of almost 60,000 European soccer matches and also found to be losing money.

5.2 Simulated strongly efficient market

This section replicates one version of a simulated strongly efficient market used in Demir et al. (2012). In this market, draws are independent events, the probability of each draw is 0.3, and the betting odds offered on each draw are 3. In such a market, each $1 bet has the expected payout of 0.3 * 3 = $0.9, so the expected value of such a bet is -10 cents.

To evaluate the Fibonacci betting strategy, the betting must actually stop at some point in time. One option, used in both Archontakis and Osborne (2007) and Demir et al. (2012), is to stop betting after X matches. However, this could generate huge losses if X is high and no draws occur. A second, more realistic option is to stop betting if the total profit is at least $X or less than or equal to -$X. This corresponds to the gambler willing to risk $X and wanting to earn at least this amount – something that a profitable strategy should be able to do more often than half the time. Table 21 shows the results for three different settings for each option; each set of results is based on 10,000,000 computer simulations.

	Stop betting after X matches			Stop betting if profit ≥ $X or ≤ -$X		
	X = 10	X = 20	X = 40	X = 10	X = 100	X = 1,000
Maximum number of bets	10	20	40	24	166	1,208
Average number of bets	10	20	40	11.2373	75.4538	451.5707
Maximum single bet	55	6,765	102,334,155	8	89	987
Maximum profit	22	2,585	39,088,170	13	134	1,377
Minimum profit	-143	-17,710	-267,914,295	-17	-188	-1,986
Relative frequency of positive profit	0.7386	0.8628	0.9316	0.4476	0.4273	0.4071
Relative frequency of negative profit	0.2340	0.1299	0.0675	0.5524	0.5727	0.5929
Average sum of bets	28.3961	165.6527	2366.8091	22.1124	342.2596	4,267.2651
Average sum of winnings	25.5603	148.9913	2139.7781	19.9050	308.0609	3,840.0997
Average profit	-2.8358	-16.6614	-227.0310	-2.2074	-34.1986	-427.1654
Profit margin	-0.0999	-0.1006	-0.0959	-0.0998	-0.0999	-0.1001

Table 21: Fibonacci strategy in a strongly efficient market, 10,000,000 simulations for each setting

The first option of stopping after X matches produces highly asymmetrical returns; it has a high probability of generating a small profit and a low probability of generating a large loss. The second option provides more symmetrical results, but the strategy brings a positive profit in less than 50 percent of the cases. The key result is that for each setting, the average sum of bets is higher than the average sum of winnings, so the average profit is negative. This is also easy to prove theoretically: If a gambler bets X_1 on match number 1, X_2 on match number 2 … X_n on match number n, the expected winnings are 0.3 * 3 * X_1, 0.3 * 3 * X_2 … 0.3 * 3 * X_n, so the expected sum of winnings = 0.9 * sum of bets and the expected profit margin = (expected sum of winnings – sum of bets)/sum of bets = -0.1 (close to the simulated value for all settings). Both the simulation results and the theoretical proof contradict the findings in Demir et al. (2012); however, they stopped betting after 150 matches and used only 1,000 simulations – not enough to properly explore the whole range of possible outcomes.[73]

5.3 Real market

Although the Fibonacci strategy is not and cannot be profitable in a strongly efficient market, it could still be profitable in a real market under the following two conditions: first, some bets on draws have positive expected values; second, the amounts bet on such matches are high enough to more than compensate for expected losses from the other bets. This could happen if bookmakers underestimated the probability of a draw after a long string of non-drawn matches.

To test whether the Fibonacci strategy is profitable in a real betting market, this article uses data from 171 completed seasons of 19 top European soccer competitions that took place from 2004/05 to 2012/13. The data set contains 59,725 match results with valid betting odds.[74]

The Fibonacci strategy is simulated in the following way: for each match in the data set, there are 1,000 bettors that start their betting on this match. Each bettor then continues betting on draws in

[73] The highly asymmetrical returns for stopping after X matches are the complicating factor; for the profit margin to converge, the simulated sample should contain a sufficient number of even the worst-case outcomes of no draws at all. For stopping after 40 matches, the probability of such an outcome is $(1 – 0.3)^{40} \approx 6.4 * 10^{-7}$, so even 10,000,000 simulations used in this article are barely enough for this specific setting.

[74] The 19 competitions are the top Belgian, top 2 German, top 4 English, top 2 French, top Greek, top 2 Italian, top Dutch, top Portuguese, top 2 Scottish, top 2 Spanish, and top Turkish league. The data set was downloaded from the website football-data.co.uk on June 10th, 2013, and contained 61,646 match results; 1,921 matches (3 %) did not have associated valid betting odds, so they were discarded. The betting odds were quoted by a major British bookmaker William Hill.

the closest available match in the same competition, but only on one match on the same day. If there are more matches played on the same day, there are two alternative settings: first, the bettor chooses randomly from all matches on that day; second, the bettor chooses randomly from all matches with the highest betting odds on a draw on that day (used in Demir et al., 2012). After the end of the season, the bettor continues betting on the next season of the same competition. At the end of the last season, the bettor goes back in time to the first season of the same competition. The betting ends after 20 matches (one setting) or if the total profit is at least $100 or less than or equal to -$100 (another setting). Therefore, there are 4 combinations of settings and 59,725 * 1,000 = 59,725,000 simulations for each setting. The simulation results are summarized in Table 22.

	Choose randomly from same-day matches		Choose randomly from same-day matches with highest draw odds	
	Stop betting after 20 matches	Stop betting if profit ≥ $100 or ≤ -$100	Stop betting after 20 matches	Stop betting if profit ≥ $100 or ≤ -$100
Maximum number of bets	20	375	20	168
Average number of bets	20	57.8339	20	44.8804
Maximum single bet	6,765	89	6,765	89
Maximum profit	43,175	668.95	29,645	666.95
Minimum profit	-17,710	-189	-17,710	-189
Relative frequency of positive profit	0.8560	0.4314	0.8417	0.4170
Relative frequency of negative profit	0.1438	0.5686	0.1581	0.5830
Average sum of bets	270.0189	289.6924	397.3635	257.7025
Average sum of winnings	232.6497	259.6695	341.4649	227.3980
Average profit	-37.3692	-30.0230	-55.8986	-30.3044
Profit margin	-0.1384	-0.1036	-0.1407	-0.1176

Table 22: Fibonacci strategy in a real market, 59,725,000 simulations for each setting

For all four combinations of settings, the Fibonacci strategy has a negative average profit and therefore loses money. In fact, the estimated profit margins do not really outperform the profit margin of the simplest possible strategy of betting $1 on a draw in each match in the data set (-0.1130). Again, this result contradicts the findings in Archontakis and Osborne (2007) and Demir et al. (2012); however, their results were based on extremely limited numbers of trials (1 and 32, respectively).

5.4 Conclusion

In this article, the Fibonacci strategy for betting on soccer has been tested both in a simulated strongly efficient market and on a data set of almost 60,000 European soccer matches. All tested versions of the strategy lose money in both simulated and real markets. The previous positive results were likely caused by a very low number of trials. In conclusion, the Fibonacci betting strategy, previously presented as both simple and profitable, is indeed simple, but not profitable.

Afterword

The first three articles in this dissertation deal with various types of outcome uncertainty and how they relate to match attendance demand. The main contribution of the first article is the new method of calculating match importance based on Monte Carlo simulation approach. Unlike the previous approaches, this method does not require ex-post information and can be used for any type of season outcome. The presented method is also useful for calibrating less complex algorithms, such as modified mathematical certainty, leading to better estimates of the impact of match importance on attendance.

The second article is the first to apply the Monte Carlo simulation framework to the question how an additional playoff stage impacts seasonal uncertainty. Using the Czech ice hockey Extraliga as an example, the article shows that the additional playoff stage decreases the probability that the strongest team becomes the champion and thus increases seasonal uncertainty. Compared to the previous approaches, the Monte Carlo simulation allows for deeper analysis of various what-if scenarios, alternative tournament designs, and strategic team behavior.

The third article analyzes the link between match uncertainty and attendance and makes two contributions; first, it shows that the inconsistent findings in the literature could be explained by wrongly specified regressions; second, it proposes a new, non-regression approach to analyzing the effect of match uncertainty. The results show that attendance demand is maximized if teams of the same quality play against each other. Based on this finding, the total attendance could be moderately increased if teams were split into groups based on team quality instead of on region or by making the tournament design more similar to the Swiss system commonly used in chess.

The last two articles in this dissertation investigate efficiency of sports betting markets. The fourth article uses tennis betting data to distinguish between competing explanations for the so-called favorite-longshot bias. Unlike the previous articles, it focuses on how the bias changes in different types of matches. The results show that the favorite-longshot bias is much stronger in matches between lower-ranked players, in later-round matches, and in high-profile tournaments.

These results cannot be solely explained by bettors being locally risk-loving or overestimating chances of longshots, but are consistent with bookmakers protecting themselves against both better informed insiders and the general public exploiting new information.

The last article tests the strategy of betting on soccer draws using the Fibonacci sequence that has been previously found to be both simple and profitable, thus refuting even weak market efficiency. Using a bigger data set, many more simulations, and a better criterion of profitability, the article finds that all tested versions of the strategy actually lose money.

References

Archontakis, F., & Osborne, E. (2007). Playing It Safe? A Fibonacci Strategy for Soccer Betting. *Journal of Sports Economics, 8*(3), 295-308.

Baimbridge, M., Cameron, S., & Dawson, P. (1996, August). Satellite Television and the Demand for Football: A Whole New Ball Game? *Scottish Journal of Political Economy, 43*(3), 317-333.

Benz, M.-A., Brandes, L., & Franck, E. (2009, April). Do Soccer Associations Really Spend on a Good Thing? Empirical Evidence on Heterogeneity in the Consumer Response to Match Uncertainty of Outcome. *Contemporary Economic Policy, 27*(2), 216-35.

Borland, J., & Lye, J. (1992, September). Attendance at Australian Rules Football: A Panel Study. *Applied Economics, 24*(9), 1053-58.

Borland, J., & MacDonald, R. (2003). Demand for Sport. *Oxford Review of Economic Policy, 19*(4), 478-502.

Buraimo, B., & Simmons, R. (2008). Do Sports Fans Really Value Uncertainty of Outcome? Evidence from the English Premier League. *International Journal of Sport Finance*(3), 146-155.

Buraimo, B., & Simmons, R. (2009). A tale of two audiences: Spectators, television viewers and outcome uncertainty in Spanish football. *Journal of Economics and Business*(61), 326-338.

Cain, M., Law, D., & Peel, D. (2000). The favourite-longshot bias and market efficiency in UK football betting. *Scottish Journal of Political Economy, 47*(1), 25-36.

Cain, M., Law, D., & Peel, D. (2003). The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets. *Bulletin of Economic Research, 55*(3), 263-273.

Cairns, J. A. (1987). Evaluating changes in league structure: the reorganization of the Scottish Football League. *Applied Economics*(19), 259-275.

Coates, D., & Humphreys, B. R. (2011, June). Game Attendance and Competitive Balance in the National Hockey League. *University of Alberta Department of Economics Working Paper 2011-08*.

Demir, E., Danis, H., & Rigoni, U. (2012). Is the Soccer Betting Market Efficient? A Cross-Country Investigation Using the Fibonacci Strategy. *The Journal of Gambling Business and Economics, 6*(2), 29-49.

Direr, A. (2013). Are betting markets efficient? Evidence from European Football Championships. *Applied Economics, 45*(3), 343-356.

Dixon, M. J., & Coles, S. G. (1997). Modelling Association Football Scores and Inefficiencies in the Football Betting Market. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 46*(2), 265-280.

Dobson, S. M., & Goddard, J. A. (1992, October). The Demand for Standing and Seated Viewing Accommodation in the English Football League. *Applied Economics, 24*(10), 1155-63.

Dobson, S., Goddard, J., & Wilson, J. O. (2001). League Structure and Match Attendances in English Rugby League. *International Review of Applied Economics, 15*(3), 335-351.

Feddersen, A., Humphreys, B., & Soebbing, B. (2012, July). Contest Incentives in European Football. *University Of Alberta, Department of Economics Working Paper Series, Working Paper No. 2012-13*.

Forrest, D., & McHale, I. (2007). Anyone for Tennis (Betting)? *The European Journal of Finance, 13*(8), 751-768.

Forrest, D., & Simmons, R. (2002). Outcome Uncertainty and Attendance Demand in Sport: The Case of English Soccer. *Journal of the Royal Statistical Society - Series D (The Statistician), 51*(2), 229-241.

Fort, R., & Quirk, J. (1995, September). Cross-Subsidization, Incentives and Outcomes in Professional Team Sports Leagues. *Journal of Economic Literature, 33*(3), 1265-1299.

García, J., & Rodríguez, P. (2002, February). The Determinants of Football Match Attendance Revisited: Empirical Evidence from the Spanish Football League. *Journal of Sports Economics, 3*(1), 18-38.

García, J., & Rodríguez, P. (2009, September). Sports Attendance: A Survey of the Literature 1973-2007. *Rivista di Diritto ed Economia dello Sport*(2), 111-151.

Goddard, J. (2005). Regression Models for Forecasting Goals and Match Results in Association Football. *International Journal of Forecasting*(21), 331-340.

Goddard, J., & Asimakopoulos, I. (2004). Forecasting Football Results and the Efficiency of Fixed-odds Betting. *Journal of Forecasting, 23*(1), 51-66.

Goossens, D. R., Beliën, J., & Spieksma, F. C. (2012, April). Comparing league formats with respect to match importance in Belgian football. *Annals of Operations Research, 194*(1), 223-240.

Humphreys, B. R. (2002, May). Alternative Measures of Competitive Balance in Sports Leagues. *Journal of Sports Economics, 3*(2), 133-148.

Jennett, N. (1984). Attendances, Uncertainty of Outcome and Policy in Scottish League Football. *Scottish Journal of Political Economy, 31*(2), 176-198.

Koopman, S. J., & Lit, R. (2012). A Dynamic Bivariate Poisson Model for Analysing and Forecasting Match Results in the English Premier League. *Tinbergen Institute Discussion Paper*(TI 2012-099/III), 1-30.

Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics, 32*(11), 1353-1363.

Longley, N., & Lacey, N. J. (2012, October). The "Second" Season: The Effects of Playoff Tournaments on Competitive Balance Outcomes in the NHL and NBA. *Journal of Sports Economics, 13*(5), 471-493.

Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica, 36*(3), 109-118.

Makropoulou, V., & Markellos, R. N. (2011, September). Optimal Price Setting in Fixed-Odds Betting Markets under Information Uncertainty. *Scottish Journal of Political Economy, 58*(4), 519-536.

Neale, W. C. (1964, February). The Peculiar Economics of Professional Sports: A Contribution to the Theory of the Firm in Sporting Competition and in Market Competition. *The Quarterly Journal of Economics, 78*(1), 1-14.

Paul, R. J. (2003). Variations in NHL Attendance: The Impact of Violence, Scoring, and Regional Rivalries. *American Journal of Economics and Sociology, 62*(2), 345-64.

Pawlowski, T. (2013). Testing the Uncertainty of Outcome Hypothesis in European Professional Football: A Stated Preference Approach. *Journal of Sports Economics*. doi:10.1177/1527002513496011

Pawlowski, T., & Anders, C. (2012). Stadium attendance in German professional football – the (un)importance of uncertainty of outcome reconsidered. *Applied Economics Letters, 19*(16), 1553-1556.

Pawlowski, T., & Budzinski, O. (2013). The Monetary Value of Competitive Balance for Sport Consumers: A Stated Preference Approach to European Professional Football. *International Journal of Sport Finance, 8*(2), 112-123.

Peel, D. A., & Thomas, D. A. (1992). The Demand for Football: Some Evidence on Outcome Uncertainty. *Empirical Economics, 17*(2), 323-331.

Rossi, M. (2011). Match Rigging and the Favorite Long-Shot Bias in the Italian Football Betting Market. *International Journal of Sport Finance, 6*(4), 317-334.

Rottenberg, S. (1956, June). The Baseball Players' Labor Market. *Journal of Political Economy, 64*(3), 242-258.

Rue, H., & Salvesen, Ø. (2000). Prediction and Retrospective Analysis of Soccer Matches in a League. *Journal of the Royal Statistical Society. Series D (The Statistician), 49*(3), 399-418.

Sauer, R. D. (1998, December). The Economics of Wagering Markets. *Journal of Economic Literature, 36*(4), 2021-2064.

Scarf, P. A., & Yusof, M. M. (2011). A numerical study of tournament structure and seeding policy for the soccer World Cup Finals. *Statistica Neerlandica, 65*(1), 43-57.

Scarf, P. A., Yusof, M. M., & Bilbao, M. (2009). A numerical study of designs for sporting contests. *European Journal of Operational Research*(198), 190-198.

Scarf, P., & Shi, X. (2008). The importance of a match in a tournament. *Computers and Operations Research, 35*(7), 2406-2418.

Schilling, M. F. (1994, October). The Importance of a Game. *Mathematics Magazine, 67*(4), 282-288.

Simmons, R., & Forrest, D. (2006, August). New issues in attendance demand: The case of the English football league. *Journal of Sports Economics, 7*(3), 247-266.

Snowberg, E., & Wolfers, J. (2010). Explaining the Favorite-Long Shot Bias: Is it Risk-Love or Misperceptions? *Journal of Political Economy, 118*(4), 723-746.

Szemberg, S., Merk, M., & Steiss, A. (2012, March 2). *What's the value of being 1st?* Retrieved February 15, 2013, from International Ice Hockey Federation: http://www.iihf.com/nc/home-of-hockey/news/news-singleview/recap/6495.html

Szymanski, S. (2001, February). Income Inequality, Competitive Balance and the Attractiveness of Team Sports: Some Evidence and a Natural Experiment from English Soccer. *The Economic Journal*(111), F69-F84.

Szymanski, S. (2003, December). The Economic Design of Sporting Contests. *Journal of Economic Literature, 41*(4), 1137-1187.

Szymanski, S., & Késenne, S. (2004, March). Competitive Balance and Gate Revenue Sharing in Team Sports. *The Journal of Industrial Economics, 52*(1), 165-177.

Tainsky, S., Kerwin, S., Xu, J., & Zhou, Y. (2013). Will the real fans please remain seated? Gender and television ratings for pre-game and game broadcasts. *Sport Management Review*. doi:10.1016/j.smr.2013.04.002

Taylor, B. A., & Trogdon, J. G. (2002, January). Losing to Win: Tournament Incentives in the National Basketball Association. *Journal of Labor Economics, 20*(1), 23-41.

Thaler, R. H., & Ziemba, W. T. (1988, Spring). Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries. *The Journal of Economic Perspectives, 2*(2), 161-174.

Thomas, A. C. (2007). Inter-arrival Times of Goals in Ice Hockey. *Journal of Quantitative Analysis in Sports, 3*(3).

Vlastakis, N., Dotsis, G., & Markellos, R. N. (2009). How Efficient is the European Football Betting Market? Evidence from Arbitrage and Trading Strategies. *Journal of Forecasting*(28), 426-444.

Vrooman, J. (1995, April). A General Theory of Professional Sports Leagues. *Southern Economic Journal, 61*(4), 971-990.