A comparison of penalty shootout designs in soccer

László Csató[*]

Institute for Computer Science and Control (SZTAKI)
Laboratory on Engineering and Management Intelligence, Research Group of Operations
Research and Decision Systems

Corvinus University of Budapest (BCE)
Department of Operations Research and Actuarial Sciences

Budapest, Hungary

17th May 2022

𝕬𝖑𝖑𝖊𝖘 𝖊𝖗𝖘𝖈𝖍𝖊𝖎𝖓𝖙 𝖘𝖔 𝖊𝖎𝖓𝖋𝖆𝖈𝖍, 𝖆𝖑𝖑𝖊 𝖊𝖗𝖋𝖔𝖗𝖉𝖊𝖗𝖑𝖎𝖈𝖍𝖊𝖓 𝕶𝖊𝖓𝖓𝖙𝖓𝖎𝖘𝖘𝖊 𝖊𝖗𝖘𝖈𝖍𝖊𝖎𝖓𝖊𝖓 𝖘𝖔 𝖋𝖑𝖆𝖈𝖍, 𝖆𝖑𝖑𝖊 𝕶𝖔𝖒𝖇𝖎𝖓𝖆𝖙𝖎𝖔𝖓𝖊𝖓 𝖘𝖔 𝖚𝖓𝖇𝖊𝖉𝖊𝖚𝖙𝖊𝖓𝖉, 𝖉𝖆�𝖘 𝖎𝖓 𝖁𝖊𝖗𝖌𝖑𝖊𝖎𝖈𝖍𝖚𝖓𝖌 𝖉𝖆𝖒𝖎𝖙 𝖚𝖓𝖘 𝖉𝖎𝖊 𝖊𝖎𝖓𝖋𝖆𝖈𝖍𝖘𝖙𝖊 𝕬𝖚𝖋𝖌𝖆𝖇𝖊 𝖉𝖊𝖗 𝖍ö𝖍𝖊𝖗𝖊𝖓 𝕸𝖆𝖙𝖍𝖊𝖒𝖆𝖙𝖎𝖐 𝖒𝖎𝖙 𝖊𝖎𝖓𝖊𝖗 𝖌𝖊𝖜𝖎𝖘𝖘𝖊𝖓 𝖜𝖎𝖘𝖘𝖊𝖓𝖘𝖈𝖍𝖆𝖋𝖙𝖑𝖎𝖈𝖍𝖊𝖓 𝖂ü𝖗𝖉𝖊 𝖎𝖒𝖕𝖔𝖓𝖎𝖊𝖗𝖙.[1]

(Carl von Clausewitz: *Vom Kriege*)

Abstract

Penalty shootout in soccer is recognized to be unfair because the team kicking first in all rounds enjoys a significant advantage. The so-called Catch-Up Rule has been suggested recently to solve this problem but is shown here not to be fairer than the simpler deterministic Alternating (ABBA) Rule that has already been tried. We introduce the Adjusted Catch-Up Rule by guaranteeing the first penalty of the possible sudden death stage to the team disadvantaged in the first round. It outperforms the Catch-Up and Alternating Rules, while remains straightforward to implement. A general measure of complexity for penalty shootout mechanisms is also provided as the minimal number of binary questions required to decide the first-mover in a given round without knowing the history of the penalty shootout. This quantification permits a two-dimensional evaluation of any mechanism proposed in the future.

Keywords: sports rules; soccer; penalty shootout; mechanism design; fairness

MSC class: 60J20, 91A05, 91A80

JEL classification number: C44, C72, Z20

---

[*] E-mail: *laszlo.csato@sztaki.hu*

[1] "*All appears so simple, all the requisite branches of knowledge appear so plain, all the combinations so unimportant, that, in comparison with them, the easiest problem in higher mathematics impresses us with a certain scientific dignity.*" (Source: Carl von Clausewitz: *On War*, Book 1, Chapter 7 – Friction in War, translated by Colonel James John Graham, London, N. Trübner, 1873. http://clausewitz.com/readings/OnWar1873/TOC.htm)

1 Introduction

Fairness has several interpretations in sports, one basic desideratum being the interpretation of the Aristotelian Justice principle: higher-ability competitors should win with a higher probability alongside the equal treatment of equals. In particular, we address the problem of penalty shootouts in soccer (association football) from this point of view.

According to the current rulebook of soccer, Laws of the Game 2019/20, "*when competition rules require a winning team after a drawn match or home-and-away tie, the only permitted procedures to determine the winning team are: a) away goals rule; b) two equal periods of extra time not exceeding 15 minutes each; c) kicks from the penalty mark*" (IFAB, 2019, Section 10). In the ultimate case of item *c)*, a coin is tossed to decide the goal at which the kicks will be taken. Then the referee tosses a coin again, the winner decides whether to take the first or second kick, and five kicks are taken alternately by both teams (if, before both teams have taken five kicks, one has scored more goals than the other could score, even if it were to complete its five kicks, no more kicks are taken). If the scores are still level after five rounds, the kicks continue in the *sudden death* stage until one team scores a goal more than the other from the same number of kicks. Following Brams and Ismail (2018), we will refer to this rule as the *Standard (ABAB) Rule*.

Since most penalties are successful in soccer, the player taking the second kick is usually under greater mental pressure, especially from the third or fourth penalties onward, when a miss probably means the loss of the match. Consequently, the team kicking first in a penalty shootout is recognized to win significantly more frequently than 50 percent of the time (Apesteguia and Palacios-Huerta, 2010; Palacios-Huerta, 2014; Da Silva et al., 2018; Rudi et al., 2019), indicating the unfairness of the Standard ($ABAB$) Rule.

Therefore, three alternative mechanisms for penalty shootouts will be considered here:

- *Alternating (ABBA) Rule*: the order of the first two penalties (AB) is mirrored in the next two (BA), and this sequence continues even in the possible sudden death stage of the shootout (the sixth round of penalties is started by team B, the seventh by team A, and so on).

- *Catch-Up Rule* (Brams and Ismail, 2018): the order of the penalties in a given round, including the sudden death, is the mirror image of the previous round except if the first team failed and the second scored in the previous round when the order of the teams remains unchanged.

- *Adjusted Catch-Up Rule*: the first five rounds of penalties, started by team A, are kicked according to the Catch-Up Rule, but team B is guaranteed to be the first kicker in the sudden death stage (sixth round).

Note that the Adjusted Catch-Up Rule combines the other two mechanisms: it coincides with the Catch-Up Rule in the first five rounds and with the Alternating ($ABBA$) Rule in the sudden death stage.

The three designs will be compared concerning not only fairness but also simplicity because achieving fairness has a price in increasing complexity.

The contribution of our research resides in the following points:

1. We find that the Catch-Up Rule, promoted by Brams and Ismail (2018), does not outperform the simpler and already tried Alternating ($ABBA$) Rule under the assumptions of the same authors, which substantially reduces the importance of a central result of Brams and Ismail (2018);

2. We show that the proposed Adjusted Catch-Up Rule is fairer than both alternative penalty shootout designs;

3. We suggest the first general complexity measure of penalty shootout mechanisms in the literature that remains consistent with the view of the decision-makers. It is based on the minimal number of binary questions required to decide the first-mover in a given round of the penalty shootout without knowing its history.

The paper is organized as follows. Section 2 discusses the problem of penalty shootouts, while Section 3 analyzes the fairness of the three penalty shootout designs presented above. Quantification of the complexity of an arbitrary rule is provided in Section 4. Section 5 offers some concluding thoughts.

2 An overview of penalty shootouts

Soccer is typically a game with a low number of scores, hence ties, even the result of 0-0, are relatively common. Since in knockout (elimination) tournaments only one team advances to the next round, these ties should be broken.

Before 1970, soccer matches that were tied after extra time were either decided by a coin toss or replayed. However, events in the 1968 European football championship led FIFA, the international governing body of association football, to try penalty shootouts (Anbarcı et al., 2019).[2] In the following decades, penalty shootout has become the standard tie-breaking procedure in knockout tournaments.

Besides, penalty shootout may be a specific tie-breaking rule in round-robin tournaments. For example, in the group stage of the 2020 UEFA European Football Championship, if two teams, which have the same number of points and the same number of goals scored and conceded, play their last group match against each other and are still equal at the end of that match, their final rankings are determined by kicks from the penalty mark, provided that no other teams within the group have the same number of points on completion of all group matches (UEFA, 2018, Article 20.02).[3]

In the 1988-89 season of the Argentinian League, all drawn matches went to penalties without extra time, when the winner of the shootout obtained two points, and the loser one point (Palacios-Huerta, 2014, Section 10). A similar rule was applied in the 1994-95 Australian National Soccer League, except that a regular win was awarded by four points (Kendall and Lenten, 2017, Section 3.9.7).

2.1 On the fairness of penalty shootouts

Penalty shootouts have inspired many academic researchers to investigate the issue of fairness as they offer excellent natural experiments despite a substantial rule change in their implementation: before June 2003, the team that won the random coin toss had to take the first kick, and after July 2003, the winner of the coin toss can choose the order of kicking. The stakeholders also feel potential problems, more than 90% of coaches and

---

[2] The tied semifinal (after extra time) between Italy and the Soviet Union was decided by a coin toss for Italy. The final between Italy and Yugoslavia ended in a draw of 1-1 even after 30 minutes extra time, thus it was replayed two days later.

[3] Despite the restrictive conditions, it remains not only a theoretical possibility: in the elite round of 2016 UEFA European Under-17 Championship qualification, Poland overtook Ireland in Group 7, and Belgium was ranked higher than Spain in Group 8 due to this particular rule.

players asked in a survey want to go first, mainly because they attempt to put psychological pressure on the other team (Apesteguia and Palacios-Huerta, 2010).[4]

Denote the two teams by A and B, where A is the first kicker. According to Apesteguia and Palacios-Huerta (2010), team A wins with the probability of 60.5% based on 129 pre-2003 penalty shootouts and with the probability of 59.2% based on 269 shootouts that include post-2003 cases. The advantage of the first-mover is statistically significant. However, using a superset of their pre-2003 sample with 540 shootouts, Kocher et al. (2012) report this value to be only 53.3% and insignificant. Palacios-Huerta (2014) further expands the database to 1001 penalty shootouts played before 2012 to get a 60.6% winning probability for the first team. Vandebroek et al. (2018) explain this disagreement with insufficient sample sizes, they find that even a relatively small but meaningful lagging-behind effect (the team having less score succeeds with only a 70% probability instead of 75%) cannot be reliably identified if only 500 penalty shootouts are considered.

Da Silva et al. (2018) collect 232 penalty shootout situations and get a 59.48% winning probability for team A, which is statistically significant. On the other hand, Arrondel et al. (2019) show no advantage based on 252 French penalty shootouts. However, their results reveal that the probability of scoring is negatively affected by the stake (the impact of my scoring on the expected probability that my team will eventually win) and the difficulty of the situation (the ex-ante probability of my team eventually losing). Finally, Rudi et al. (2019) investigate 1635 penalty shootouts, which leads to a statistically significant 54.86% winning probability for team A. Although this is closer in magnitude to the value presented by Kocher et al. (2012) than to the findings of Apesteguia and Palacios-Huerta (2010) and Palacios-Huerta (2014), the larger sample size enables a more precise estimation and higher statistical power to detect the possible advantage.

Similar problems may arise in other sports such as handball, ice hockey, or water polo (Anbarcı et al., 2019). Cohen-Zada et al. (2018) and Da Silva et al. (2018) find that the $ABBA$ pattern does not favor the player who serves first in a tennis tiebreak. According to González-Díaz and Palacios-Huerta (2016), the player drawing the white pieces in the odd games of a multi-stage chess contest has about 60% chance to win the match. Therefore, since the World Chess Championship 2006, the colors are reversed halfway through in the match containing twelve scheduled games as one player plays with the white pieces in the 1st, 3rd, 5th, 8th, 10th, 12th games according to the $ABABAB|BABABA$ sequence.

To summarize, while the empirical evidence remains somewhat controversial, it seems probable that the team kicking the first penalty enjoys an advantage, which is widely regarded as unfair. This fact is also recognized by the IFAB (International Football Association Board), the rule making body of soccer: Laws of the Game 2017/18 explicitly says in its section discussing the future that the IFAB will consult widely on a number of important Law-related topics, including "*a potentially fairer system of taking kicks from the penalty mark*" (IFAB, 2017).[5]

2.2 Alternative mechanisms for penalty shootouts

The IFAB has decided to test the *Alternating (ABBA) Rule*. The trial was initially scheduled at the 2017 UEFA European Under-17 Championship and the 2017 UEFA

---

[4] An interesting exception was a quarterfinal of the 2018 FIFA World Cup when the Croatian team captain *Luka Modrić* chose to kick the second penalties despite winning the coin toss against Russia (Mirror, 2018).

[5] This sentence appears in the same place in Laws of the Game 2018/19 (IFAB, 2018), but is missing from Laws of the Game 2019/20 (IFAB, 2019).

Women's Under-17 Championship, organized in May 2017 (UEFA, 2017b), and was extended to the 2017 UEFA European Under-19 Championship and the 2017 UEFA Women's Under-19 Championship in the following month (UEFA, 2017a). The first implementation of the new system was a penalty shootout between Germany and Norway in the Women's Under-17 Championship semifinal on 11 May 2017 (Thomson Reuters, 2017).

This mechanism was applied in the 2017 FA Community Shield, too, where Arsenal, the winner of the 2017 FA Cup Final, won after an $ABBA$ penalty shootout against Chelsea, the champions of the 2016/17 Premier League. There was even a controversy in the Dutch KNVB Cup in 2017 when a referee erroneously employed the Alternating ($ABBA$) rule during a penalty shootout, hence it should be replayed three weeks after (Mirror, 2017).

However, the 133rd Annual Business Meeting (ABM) of the IFAB agreed that the Alternating ($ABBA$) rule will no longer be a future option for competitions due to "*the absence of strong support, mainly because the procedure is complex*" (FIFA, 2018).

Academic researchers have proposed some further rules to increase fairness (Anbarcı et al., 2019; Brams and Ismail, 2018; Echenique, 2017; Palacios-Huerta, 2012). Our point of departure is the *Catch-Up Rule* (Brams and Ismail, 2018), which takes into account the results of penalties in the preceding round to allow the team performing worse to catch up. Assume that team A kicks first in a particular round, thus it is advantaged. In the next round, team B will kick first except if A fails and B succeeds.

We suggest a slight improvement in this mechanism. Note that the penalty shootout is essentially composed of two parts, the first five rounds, and the possible sudden death stage. Therefore, it makes sense to balance the advantage of the first-mover by making it disadvantaged at the beginning of the sudden death. Formally, if team A starts the shootout, then team B will kick first in the sixth round, provided that it is reached. Under the original Catch-Up Rule, it is possible that A kicks first in the sixth round, for instance, when it leads by 4-3 after four rounds, but A fails and B succeeds in the fifth round of penalty kicks. This variant of the Catch-Up Rule, which *a priori* fixes the first-mover in the sudden death, is called the *Adjusted Catch-Up Rule*.

Table 1 illustrates how the four rules work. The Red team is the first kicker, ✔ means a successful, and ✗ indicates an unsuccessful penalty. Since the result after five rounds is 3-3, the sudden death stage starts: the Red team kicks first in the sixth round according to the Catch-Up Rule as the Blue team was the first-mover in the previous round, but the Blue team kicks first in the sixth round when the Adjusted Catch-Up Rule is used because it was disadvantaged in the first round.

3   The analysis of three penalty shootout designs

Following the literature on penalty shootouts, fairness means that no team enjoys an advantage because of winning or losing the coin toss.

Definition 1. *Fairness*: A penalty shootout mechanism is *fair* if the probability of winning does not depend on the outcome of the coin toss.

Consequently, in our mathematical model, a mechanism is called *fairer* than another if the probability of winning the match conditional on winning the coin toss is closer to 0.5 for equally skilled teams.

Table 1: An example of penalty shootout rules

Rule	ABAB		ABBA		Catch-Up		Adj. Catch-Up	
Team	Red	Blue	Red	Blue	Red	Blue	Red	Blue
1st kick	✔		✔		✔		✔	
2nd		✔		✔		✔		✔
3rd	✗			✗		✗		✗
4th		✗	✗		✗		✗	
5th	✔		✔		✔		✔	
6th		✔		✔		✔		✔
7th	✔			✗		✗		✗
8th		✗	✔		✔		✔	
9th	✗		✗			✔		✔
10th		✔		✔	✗		✗	
11th	✔			✔	✔			✔
12th		✔	✔			✔	✔	
13th	✔		✔			✗	✔	
14th		✗		✗	✔			✗

The standard *ABAB* rule will not be discussed here because it has already been investigated in Brams and Ismail (2018) — and is markedly unfair.

3.1 Fairness: a simple model which solely depends on the order

Table 2: Penalty shootout success rates per round

	First kicker	Second kicker
Round 1	0.79	0.72
Round 2	0.82	0.77
Round 3	0.77	0.64
Round 4	0.74	0.68
Round 5	0.74	0.67

Source: Apesteguia and Palacios-Huerta (2010, p. 2558)

Apesteguia and Palacios-Huerta (2010, p. 2558) provide empirical probabilities for scoring a penalty on each round, presented in Table 2. It can be seen that the team kicking first in a given round always succeeds with a higher probability. Hence, following Brams and Ismail (2018), we use the reasonable assumption that the probability of a successful penalty depends only on whether the team kicks first or second in a round: the advantaged team has a probability p of scoring, and the disadvantaged team has a probability $q(\leq p)$ of scoring. Similarly to Brams and Ismail (2018), our baseline choice is $p = 3/4$ and $q = 2/3$, which are close to the empirical success rates given in Table 2,

especially in the last three rounds. This corresponds to about a 60% chance of winning for the first-mover as observed in practice by Apesteguia and Palacios-Huerta (2010) and Palacios-Huerta (2014).

To illustrate the model, Brams and Ismail (2018) analyze the Catch-Up Rule for a penalty shootout over only two rounds and derive that $p = 3/4$ and $q = 2/3$ result in:

- the probability of team A winning is $P^2(A) = 41/144 \approx 0.285$;[6]

- the probability of team B winning is $P^2(B) = 39/144 \approx 0.270$;

- the probability of a tie is $P^2(T) = 64/144 \approx 0.444$.

If there is a tie after two rounds, the shootout goes to sudden death. Assume that team A kicks first and let $W(A)$ be the probability of winning for team A in the sudden death stage. The Catch-Up, Adjusted Catch-Up, and Alternating ($ABBA$) Rules coincide in this stage, the calculations of Brams and Ismail (2018) remain valid, that is,

$$W(A) = \frac{1 - q + pq}{2 - p - q + 2pq}. \tag{1}$$

For $p = 3/4$ and $q = 2/3$, one gets $W(A) = 10/19 \approx 0.526$.

If the penalty shootout is played over two rounds before sudden death, the probability of a tie is $P^2(T) = 64/144$. Under the Catch-Up Rule, A kicks first in the third round with a probability of $58/144 \approx 0.403$, while B kicks first in the third round with a probability of $6/144 \approx 0.042$ because team B will kick first only in the case of the following sequence: A fails, B scores, A scores, B fails, which has a probability of $(1-p)qp(1-q)$. Consequently, the probability that team A wins is

$$Q^2(A) = P^2(A) + \frac{58}{144} \times \frac{10}{19} + \frac{6}{144} \times \frac{9}{19} = \frac{1413}{2736} \approx 0.516.$$

On the other hand, the Adjusted Catch-Up Rule guarantees the first penalty in the sudden death for team B, hence the probability that team A wins under this mechanism is

$$Q^2(A) = P^2(A) + \left(\frac{58}{144} + \frac{6}{144} \right) \times \frac{9}{19} = \frac{1355}{2736} \approx 0.495.$$

A more detailed discussion of the Alternating ($ABBA$) Rule is provided because it is missing from Brams and Ismail (2018) but can contribute to a better understanding of the model. There are three ways for team A to win a penalty shootout over two rounds:

I) 2-0: *A scores on both rounds while B fails to score on both rounds*
   On the first round, A succeeds and B fails with probability $p(1-q)$. On the second round, B kicks first and fails, while A kicks second and succeeds with probability $(1-p)q$. The joint probability of this outcome over the two rounds is $p(1-q)(1-p)q$.

II) 2-1: *A scores on both rounds while B fails to score on one of these rounds*
   There are two subcases:

- *B scores on the first round*
  On this round, both teams succeed with probability pq. On the second round, B kicks first and fails, while A kicks second and scores with probability $(1-p)q$. The joint probability over both rounds is $pq(1-p)q$.
- *B scores on the second round*
  On the first round, A succeeds and B fails with probability $p(1-q)$. On the second round, B kicks first and succeeds, after which A also scores, with probability pq. The joint probability over both rounds is $p(1-q)pq$.

Hence the probability of the outcome 2-1 is

$$pq(1-p)q + p(1-q)pq.$$

III) 1-0: *A scores on one round while B fails to score on both rounds*
There are two subcases:

- *A scores on the first round*
  On this round, A succeeds and B fails with probability $p(1-q)$. On the second round, both teams fail with probability $(1-p)(1-q)$. The joint probability over the two rounds is $p(1-q)(1-p)(1-q)$.

- *A scores on the second round*
  On the first round, both teams fail with probability $(1-p)(1-q)$. On the second round, B kicks first and fails, after which A succeeds, with probability $(1-p)q$. The joint probability over the two rounds is $(1-p)(1-q)(1-p)q$.

Thus the probability of the outcome 1-0 is

$$p(1-q)(1-p)(1-q) + (1-p)(1-q)(1-p)q.$$

The assumption $p = 3/4$ and $q = 2/3$ implies that:

- the probability of team A winning is $P^2(A) = 41/144 \approx 0.285$;

- the probability of team B winning is $P^2(B) = 41/144 \approx 0.285$;

- the probability of a tie is $P^2(T) = 62/144 \approx 0.431$.

Unsurprisingly, this rule leads to equal winning probabilities for the two teams over two rounds as two is an even number.

The Alternating ($ABBA$) Rule provides the first penalty in the sudden death for team A because it is the third round, hence the probability that team A wins is

$$Q^2(A) = P^2(A) + \frac{62}{144} \times \frac{10}{19} = \frac{1399}{2736} \approx 0.511.$$

To summarize, while all three alternative designs tend to equalize the winning probabilities compared to the Standard ($ABAB$) Rule, the Adjusted Catch-Up Rule seems to be the closest to fairness. In particular, the Catch-Up and Alternating ($ABBA$) Rules give $100 \times (0.516/0.484 - 1) \approx 6.8\%$ and 4.64% advantage for the team kicking the first penalty, respectively, while the Adjusted Catch-Up Rule results in an advantage of only 1.92% for *the other team* in a penalty shootout over two rounds with sudden death.

Table 3: The probability that A wins including sudden death ($p = 3/4$ and $q = 2/3$)

	Catch-Up Rule	Adjusted Catch-Up Rule	Alternating ($ABBA$) Rule
1 Round	0.526	0.526	0.526
2 Rounds	0.516	0.495	0.511
3 Rounds	0.518	0.515	0.519
4 Rounds	0.513	0.501	0.508
5 Rounds	**0.514**	**0.509**	**0.515**
6 Rounds	0.512	0.504	0.507
7 Rounds	0.512	0.507	0.513
8 Rounds	0.511	0.504	0.506

The winning probabilities of the advantaged team, which kicks the first penalty, are shown in Table 3 for penalty shootouts lasting eight or fewer predetermined rounds followed by sudden death when $p = 3/4$ and $q = 2/3$. Note that the probabilities for the Catch-Up Rule have already been reported in Brams and Ismail (2018) up to five rounds.

All three methods, especially the Alternating ($ABBA$) Rule, exhibit a small odd-even effect since their bias is greater for an odd number of predetermined rounds. As expected, they make the contest fairer if the number of rounds increases. The simplest Alternating ($ABBA$) Rule is better than the Catch-Up Rule for an even number of rounds, while the latter has a marginal advantage for an odd number of rounds.

However, the Adjusted Catch-Up Rule consistently outperforms both of them. The smallest imbalance can be observed for a penalty shootout played over four rounds, followed by sudden death if the shootout remains unresolved. In this case, the team kicking first has only 0.58% more chance to win under the Adjusted Catch-Up Rule.

Until now, we have investigated only the case of $p = 3/4$ and $q = 2/3$. Figure 1 plots the winning probabilities of team A using the presented rules for different values of p as a function of q, where $0.5 \leq q \leq p$ since the penalties in soccer are usually successful. It shows that the order of these designs with respect to fairness is not influenced by the particular parameters chosen: the Catch-Up and the Alternating ($ABBA$) Rules remain almost indistinguishable, and the Adjusted Catch-Up Rule turns out to be the best as before. Furthermore, all mechanisms are fairer if p is closer to q, according to our intuition.

Unfortunately, there is no hope to analytically derive conditions for p and q which make the Adjusted Catch-Up Rule fairer compared to the other designs even in this simple mathematical model. The reason is that the five rounds of penalties imply $2^{10} = 1024$ different cases, and the probability of each is given by a formula containing the product of ten items from the set of p, q, $(1 - p)$, and $(1 - q)$. Nevertheless, Figure 1 supports this conjecture by reinforcing the lack of non-linear effects.

3.2 Fairness: empirical round dependent scoring probabilities

The three rules can also be compared in the view of the empirical round dependent probabilities from Table 2. Since success rates in the sudden death stage are uncertain due to the small sample size, it is assumed that our former mathematical model holds after five rounds with the fixed probabilities p and q. Figure 2 presents the results of these calculations. While the Catch-Up Rule is closer to fairness based on the empirical data than the Alternating ($ABBA$) Rule, the Adjusted Catch-Up Rule remains the winner.

Figure 1: The probability that A wins a shootout over five rounds including sudden death

<div align="center">

Catch-Up Rule ⋯⋯ Adjusted Catch-Up Rule —— Alternating ($ABBA$) Rule – –

</div>
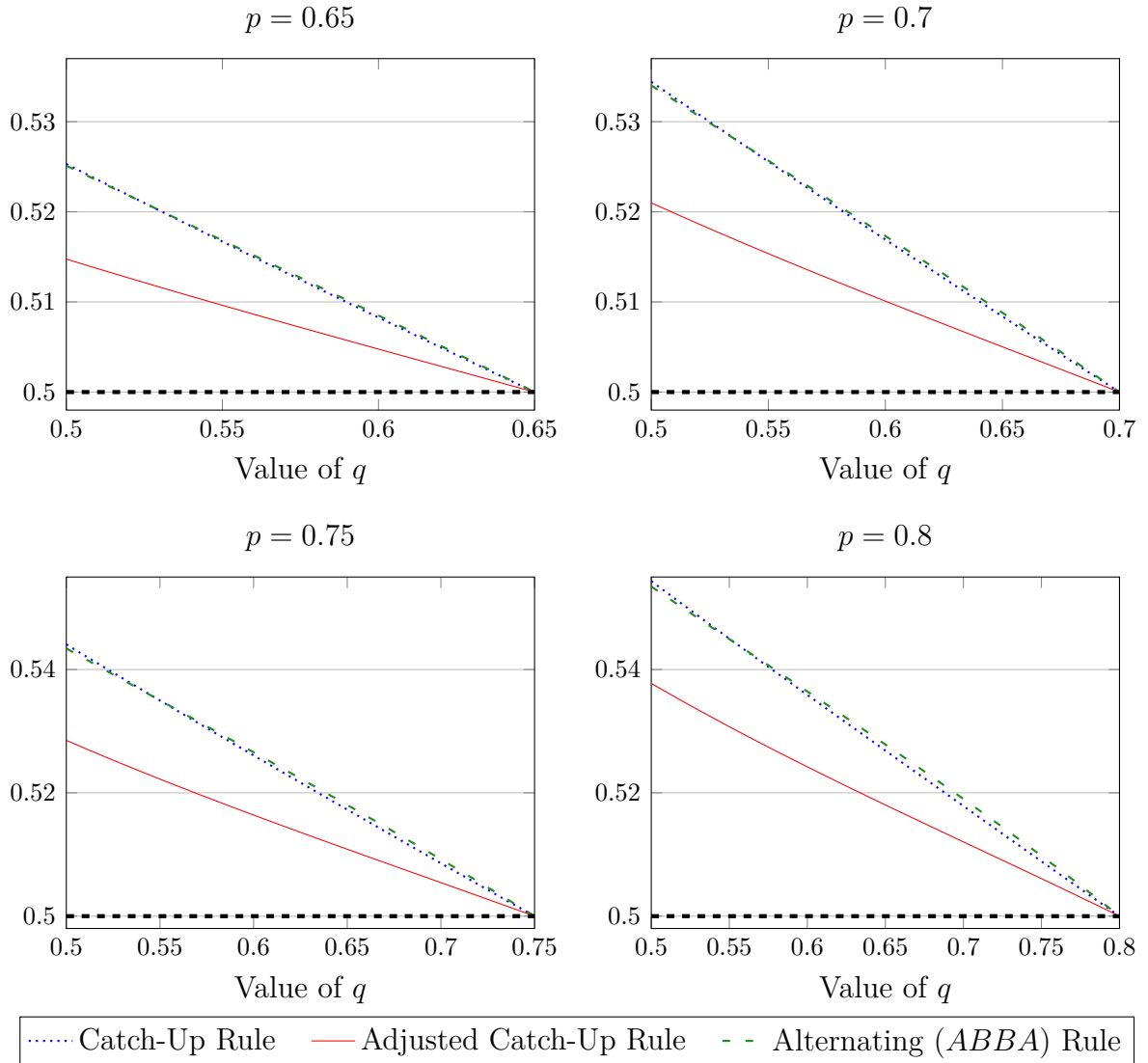
We have attempted to determine the scoring probabilities p and $q (\leq p)$ in the sudden death stage which make the Adjusted Catch-Up Rule fairer than the other two mechanisms. Formally, suppose that the following values are known:

- $P^5(A)$: the probability that A wins a penalty shootout over five rounds without sudden death under the Catch-Up Rule;

- $P_A^5(T)$: the probability that a penalty shootout over five rounds is tied under the Catch-Up Rule and A kicks the sixth penalty according to the Catch-Up Rule;

- $P_B^5(T)$: the probability that a penalty shootout over five rounds is tied under the Catch-Up Rule and B kicks the sixth penalty according to the Catch-Up Rule.

Furthermore, denote by α the probability of winning the sudden death by the team that kicks first in this stage. Formula (1) implies $0.5 \leq \alpha$ because of the assumption $q \leq p$ to incorporate the psychological effect, which is probably even stronger in the sudden death.

Then the overall probability of winning for team A under the Catch-Up Rule is

$$P^5(A) + P_A^5(T) \times \alpha + P_B^5(T) \times (1 - \alpha), \tag{2}$$

Figure 2: The empirical probability that A wins a penalty shootout over five rounds including sudden death
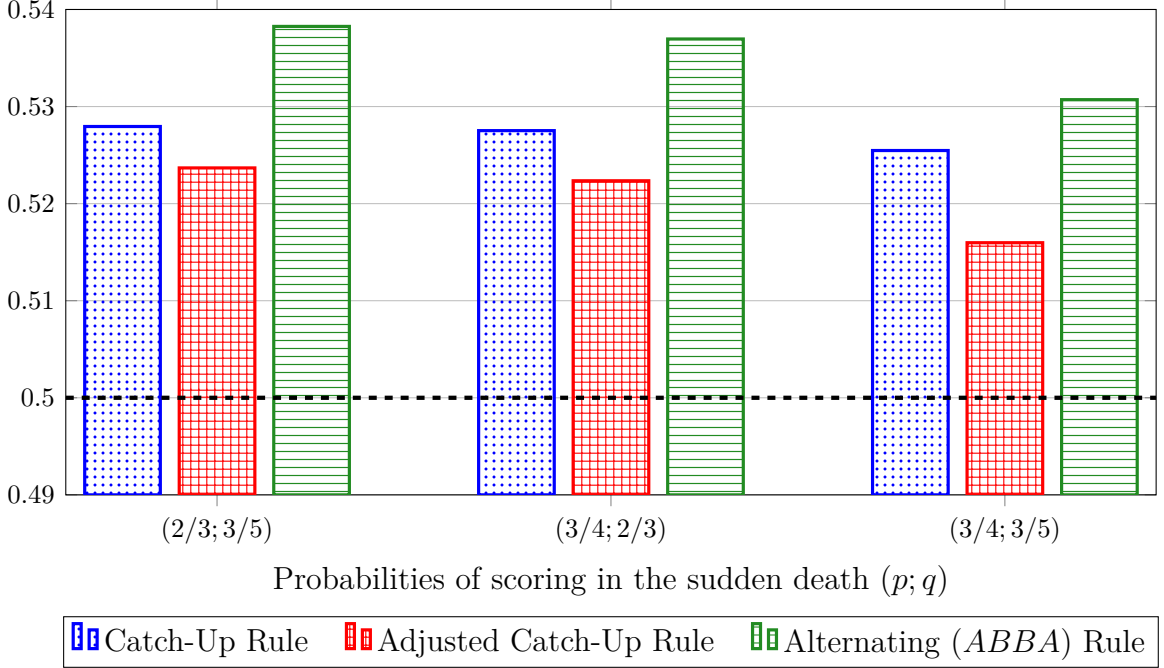
while the overall probability of winning for team A under the Adjusted Catch-Up Rule is

$$P^5(A) + \left(P_A^5(T) + P_B^5(T) \right) \times (1 - \alpha). \tag{3}$$

The Adjusted Catch-Up Rule is fairer than the Catch-Up Rule if the value of (3) is closer to 0.5 than the value of (2). By using the round dependent empirical scoring probabilities of Table 2, this results in $0.5 \leq \alpha \leq \alpha(CU) \approx 0.6569$. Thus the Adjusted Catch-Up Rule becomes fairer than the Catch-Up Rule if

$$0.5 \leq \frac{1 - q + pq}{2 - p - q + 2pq} \leq \alpha(CU) \iff \frac{1 - 2\alpha(CU) + \alpha(CU)p}{1 - p - \alpha + 2\alpha(CU)p} \leq q \leq p.$$

An analogous calculation leads to the conclusion that the Adjusted Catch-Up Rule is fairer than the $ABBA$ Rule if $0.5 \leq \alpha \leq \alpha(ABBA) \approx 0.6252$.

The range of values $(p; q)$, $q \leq p$ for the scoring probabilities in sudden death that makes the Adjusted Catch-Up Rule fairer than the other two penalty shootout designs with the empirical results from Table 2 are plotted in Figure 3. Our proposal outperforms the Catch-Up Rule in the region indicated by the blue vertical lines, while it is preferred to the Alternating ($ABBA$) Rule in the region indicated by the green horizontal lines (the latter is a subset of the former). Since any reasonable value of q lies between these bounds, the Adjusted Catch-Up Rule is the closest to fairness among the three designs with the empirical round dependent success rates of Apesteguia and Palacios-Huerta (2010).

3.3 Beyond fairness: expected length and strategy-proofness

In the model above, the expected length of the sudden death stage is $1/(p + q - 2pq)$, the same for all mechanisms (Brams and Ismail, 2018). The Catch-Up and Adjusted Catch-Up Rules differ only in which team kicks the first penalty of the sudden death.

Figure 3: The fixed scoring probabilities in sudden death which guarantee that the Adjusted Catch-Up Rule is fairer than the other penalty shootout designs
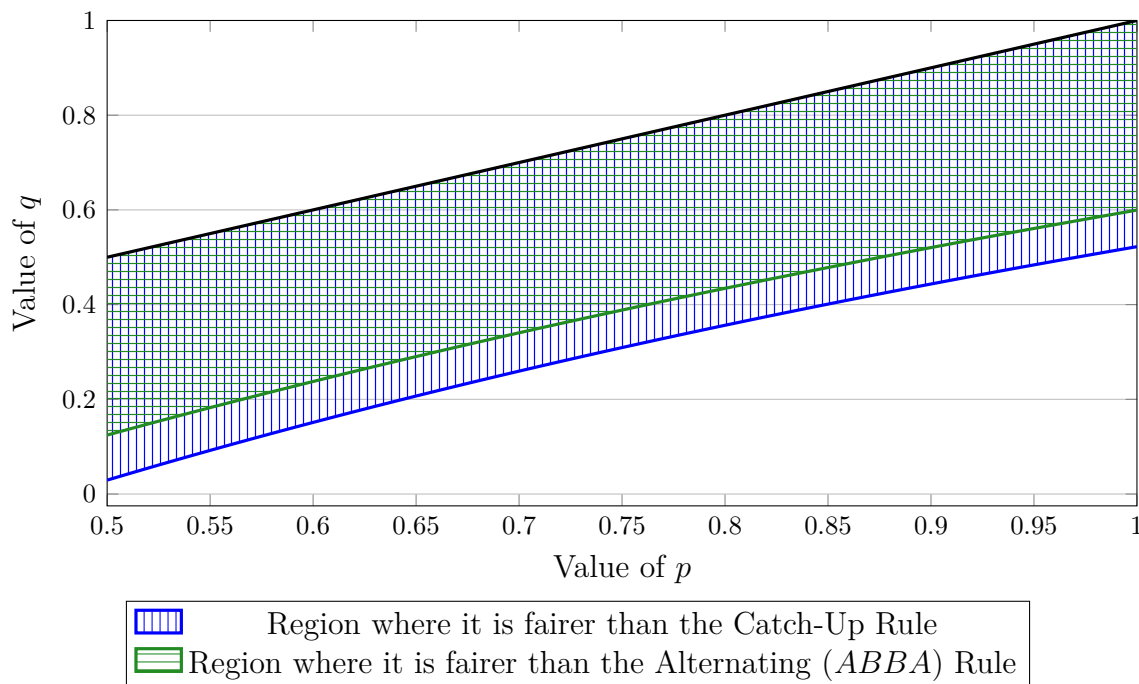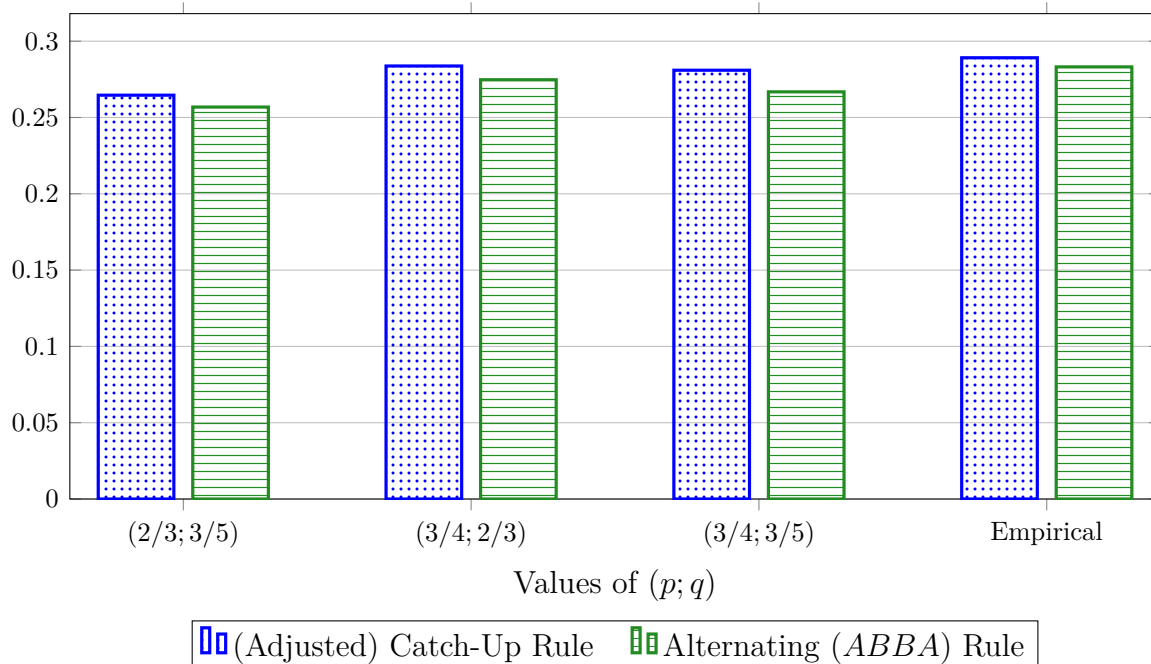

Region where it is fairer than the Catch-Up Rule
Region where it is fairer than the Alternating ($ABBA$) Rule

Figure 4: The probability that a penalty shootout over five rounds goes to sudden death

(Adjusted) Catch-Up Rule   Alternating ($ABBA$) Rule

However, the probability of reaching this stage is greater with the (Adjusted) Catch-Up Rule than with the Alternating ($ABBA$) Rule as Figure 4 illustrates based on some particular values of p and q, as well as the empirical round dependent success rates given in Table 2. Consequently, the former mechanisms can make the penalty shootout somewhat more exciting.

It has been presented recently that certain sports rules do not satisfy incentive compat-

ibility, that is, a team might be strictly better off by exerting a lower effort (Csató, 2018, 2019; Dagaev and Sonin, 2018; Kendall and Lenten, 2017; Vong, 2017). The Alternating ($ABBA$) Rule is not vulnerable to any kind of strategic manipulation since neither team can influence the order of shooting. According to Brams and Ismail (2018), no team is interested in missing a kick under the Catch-Up Rule if $(p - q) \leq 1/2$, which seems likely to be met in practice. The Adjusted Catch-Up Rule offers fewer opportunities to change the order of the penalties since the first-mover in sudden death is fixed, therefore it also satisfies strategy-proofness if the condition $(p - q) \leq 1/2$ holds.

4 The complexity of penalty shootout designs

Since the IFAB has stopped the trials of the Alternating ($ABBA$) Rule due to its complexity, this should be another important feature of mechanisms for penalty shootouts. The first attempt to quantify their simplicity has been provided in Anbarcı et al. (2015), and has been repeated in Anbarcı et al. (2019). They call a rule simple if it has a stationary machine representation with only two states such that in one state team A kicks first and in the other team B kicks first. However, this measure is not consistent with the decision of the IFAB (FIFA, 2018) because it judges the Standard ($ABAB$) and the Alternating ($ABBA$) Rules to have the same level of complexity.

Rudi et al. (2019) suggest another measure of simplicity but they choose complexity levels somewhat arbitrarily and their approach is not able to classify stochastic mechanisms (such as the Catch-Up Rule), which depend on the outcome of previous penalties.

Thus we provide a procedure that quantifies the complexity of any penalty shootout design, remains intuitive, and is consistent with the recent decision of the IFAB.

Definition 2. *Complexity*: Suppose that the *mathematician* should report the *referee* on which team is the first-mover in the next round of a penalty shootout. The mathematician has initially no information but she can ask binary questions on the history of the shootout including the number of the next round. The *complexity* of any penalty shootout mechanism is the minimal number of questions needed to determine the first kicking team in a given round, taking into account that the questions and their number might depend on the answer(s) to the preceding question(s).

In other words, there is an information asymmetry between the mathematician and the referee as the former knows only the rules, while the latter knows only the history of the shootout.

Definition 2 can be applied to reveal the simplicity of a penalty shootout mechanism.

Proposition 1. *The complexities of some penalty shootout designs are as follows:*

- *Standard ($ABAB$) Rule:* 0*;*

- *Alternating ($ABBA$) Rule:* 1*;*

- *Catch-Up Rule:* 2*;*

- *Adjusted Catch-Up Rule: between* 2 *and* 3*.*

Proof. According to the Standard ($ABAB$) Rule, team A will be the first-mover in the next round of penalties, which is known without asking any question.

The Alternating ($ABBA$) Rule requires the knowledge of the parity (odd: team A, even: team B) of the next round's number.

The Catch-Up Rule can be implemented by asking two questions because it depends on the first kicker in the previous round and on the fact whether the first kicker has failed but the second has scored in the previous round or not.

The Adjusted Catch-Up Rule first requires the knowledge of whether the sudden death stage is reached or not. After that, either the Alternating ($ABBA$) Rule (one question) or the Catch-Up Rule (two questions) is applied. □

Our approach seems to provide reasonable estimates of simplicity. For example, the design consisting of three rounds of $ABBA$ followed by Catch-Up is between 2 and 3: first, the mathematician should know whether the next round is one of the first three or not, and then the appropriate design can be implemented with further one ($ABBA$) or two (Catch-Up) questions. However, the Adjusted Catch-Up Rule is probably simpler than this artificial mechanism because changing the doctrine at the beginning of the sudden death stage can be considered less costly compared to changing the doctrine after three rounds as the rule of aggregation is modified in the sudden death anyway. Hence the Adjusted Catch-Up Rule can be judged only marginally more complex than the Catch-Up Rule.

The application of a more complex mechanism remains questionable unless it yields meaningful gains in fairness and other aspects. The Catch-Up Rule does not seem to be fairer than the Alternating ($ABBA$) Rule based on Table 3 and Figure 1, which reduces the significance of Brams and Ismail (2018)'s proposal. On the other hand, our Adjusted Catch-Up Rule dominates both of them except for a small increase in complexity.

5 Conclusions

Tournament organizers supposedly want to guarantee fairness. However, the standard penalty shootout mechanism in soccer contains a well-known bias favoring the first shooter. This means a problem because an order of actions that provides an ex-post advantage to one team may harm efficiency by decreasing the probability of the stronger team to win. Consequently, there is little excuse to continue the use of the current rule.

We have demonstrated by a mathematical model that the recently suggested Catch-Up Rule is not worth implementing since it is not fairer than the less complex Alternating ($ABBA$) Rule already tried. On the other hand, the Adjusted Catch-Up Rule can be considered as a promising candidate to make penalty shootouts fairer and even more exciting. Finally, the proposed quantification of complexity permits a two-dimensional evaluation of any mechanism recommended in the future.

Acknowledgments

References

Anbarcı, N., Sun, C.-J., and Ünver, M. U. (2015). Designing fair tiebreak mechanisms: The case of FIFA penalty shootouts. Manuscript. DOI: 10.2139/ssrn.2558979.

Anbarcı, N., Sun, C.-J., and Ünver, M. U. (2019). Designing practical and fair sequential team contests. Manuscript. DOI: 10.2139/ssrn.3453814.

Apesteguia, J. and Palacios-Huerta, I. (2010). Psychological pressure in competitive environments: Evidence from a randomized natural experiment. *American Economic Review*, 100(5):2548–2564.

Arrondel, L., Duhautois, R., and Laslier, J.-F. (2019). Decision under psychological pressure: The shooter's anxiety at the penalty kick. *Journal of Economic Psychology*, 70:22–35.

Brams, S. J. and Ismail, M. S. (2018). Making the rules of sports fairer. *SIAM Review*, 60(1):181–202.

Cohen-Zada, D., Krumer, A., and Shapir, O. M. (2018). Testing the effect of serve order in tennis tiebreak. *Journal of Economic Behavior & Organization*, 146:106–115.

Csató, L. (2018). Was Zidane honest or well-informed? How UEFA barely avoided a serious scandal. *Economics Bulletin*, 38(1):152–158.

Csató, L. (2019). UEFA Champions League entry has not satisfied strategyproofness in three seasons. *Journal of Sports Economics*, 20(7):975–981.

Da Silva, S., Mioranza, D., and Matsushita, R. (2018). FIFA is right: the penalty shootout should adopt the tennis tiebreak format. *Open Access Library Journal*, 5(3):1–23.

Dagaev, D. and Sonin, K. (2018). Winning by losing: Incentive incompatibility in multiple qualifiers. *Journal of Sports Economics*, 19(8):1122–1146.

Echenique, F. (2017). ABAB or ABBA? The arithmetics of penalty shootouts in soccer. Manuscript. https://pdfs.semanticscholar.org/85d0/3edc04470d5670266c075f7860c441a17bce.pdf.

FIFA (2018). IFAB's 133rd Annual Business Meeting recommends fine-tuning Laws for the benefit of the game. 22 November. https://www.fifa.com/about-fifa/news/y=2018/m=11/news=ifab-s-133rd-annual-business-meeting-recommends-fine-tuning-laws-for-the-benefit.html.

González-Díaz, J. and Palacios-Huerta, I. (2016). Cognitive performance in competitive environments: Evidence from a natural experiment. *Journal of Public Economics*, 139:40–52.

IFAB (2017). *Laws of the Game 2017/18.* The International Football Association Board. Effective from 1st June 2017. https://football-technology.fifa.com/media/1245/lotg_17_18_en.pdf.

IFAB (2018). *Laws of the Game 2018/19.* The International Football Association Board. Effective from 1st June 2018. http://static-3eb8.kxcdn.com/documents/662/133314_290518_LotG_18_19_EN_SinglePage_150dpi.pdf.

IFAB (2019). *Laws of the Game 2019/20.* The International Football Association Board. Effective from 1st June 2019. http://static-3eb8.kxcdn.com/files/document-category/062019/frRhKJNjSBAtiyt.pdf.

Kendall, G. and Lenten, L. J. A. (2017). When sports rules go awry. *European Journal of Operational Research*, 257(2):377–394.

Kocher, M. G., Lenz, M. V., and Sutter, M. (2012). Psychological pressure in competitive environments: New evidence from randomized natural experiments. *Management Science*, 58(8):1585–1591.

Mirror (2017). Dutch side knocked out of national cup after losing a penalty shoot-out they won three weeks earlier. 12 October. https://www.mirror.co.uk/sport/football/news/dutch-side-knocked-out-national-11331107.

Mirror (2018). Luka Modric chose second penalty despite winning coin toss against Russia – but there's a theory on why he did. 9 July. https://www.mirror.co.uk/sport/football/news/luka-modric-chose-second-penalty-12883102.

Palacios-Huerta, I. (2012). Tournaments, fairness and the Prouhet-Thue-Morse sequence. *Economic Inquiry*, 50(3):848–849.

Palacios-Huerta, I. (2014). *Beautiful game theory: How soccer can help economics.* Princeton University Press, Princeton, New York.

Rudi, N., Olivares, M., and Shetty, A. (2019). Ordering sequential competitions to reduce order relevance: soccer penalty shootouts. Manuscript. DOI: 10.2139/ssrn.3358682.

Thomson Reuters (2017). New penalty system gets usual result as Germany win. 12 May. https://uk.reuters.com/article/uk-soccer-uefa-penalties-idUKKBN18730W.

UEFA (2017a). Comprehensive bidding regulations approved for all finals and final tournaments. 1 June. www.uefa.com/insideuefa/mediaservices/newsid=2474545.html.

UEFA (2017b). Penalty shoot-out trial at UEFA final tournaments. 1 May. http://www.uefa.com/insideuefa/protecting-the-game/refereeing/news/newsid=2463576.html.

UEFA (2018). *Regulations of the UEFA European Football Championship 2018-20.* https://www.uefa.com/MultimediaFiles/Download/Regulations/uefaorg/Regulations/02/54/36/05/2543605_DOWNLOAD.pdf.

Vandebroek, T. P., McCann, B. T., and Vroom, G. (2018). Modeling the effects of psychological pressure on first-mover advantage in competitive interactions: The case of penalty shoot-outs. *Journal of Sports Economics*, 19(5):725–754.

Vong, A. I. K. (2017). Strategic manipulation in tournament games. *Games and Economic Behavior*, 102:562–567.