A comprehensive analysis of soccer penalty shootout designs

László Csató[a] Dóra Gréta Petróczy[b]

21st April 2020

Wenn ein Sachverständiger sein halbes Leben darauf verwendet, einen dunkeln Gegenstand überall aufzuklären, so wird er wohl weiter kommen als derjenige, welcher in kurzer Zeit damit vertraut sein will. Daß also nicht jeder von neuem aufzuräumen und sich durchzuarbeiten habe, sondern die Sache geordnet und gelichtet finde, dazu ist die Theorie vorhanden.[1]

(Carl von Clausewitz: *Vom Kriege*)

Abstract

The standard design of soccer penalty shootouts has received serious criticism due to its bias towards the team that kicks the first penalty. The rule-making body of the sport has decided in 2017 to try an alternative mechanism. Although the adoption of the new policy has stalled, academic researchers have recently suggested some other designs to improve fairness. This paper offers an extensive overview of seven such soccer penalty shootout mechanisms, one of them first defined here. Their fairness are analysed in three different mathematical models of psychological pressure. We also consider the probability of reaching the sudden death stage, as well as the complexity and strategy-proofness of the designs. Some rules are found to be inferior as they do not lead to a substantial gain in fairness compared to simpler mechanisms. Our work has the potential to impact decision-makers who can save resources by choosing only theoretically competitive designs for field experiments.

Keywords: OR in sports; fairness; mechanism design; penalty shootout; soccer

***MSC* class**: 60J20, 91A80

***JEL* classification number**: C44, C72, Z20

[a] Corresponding author. E-mail: *laszlo.csato@sztaki.hu*
Institute for Computer Science and Control (SZTAKI), Laboratory on Engineering and Management Intelligence, Research Group of Operations Research and Decision Systems, Budapest, Hungary
Corvinus University of Budapest (BCE), Department of Operations Research and Actuarial Sciences, Budapest, Hungary
[b] E-mail: *doragreta.petroczy@uni-corvinus.hu*
Corvinus University of Budapest (BCE), Department of Finance, Budapest, Hungary

[1] "*If a man of expertness spends half his life in the endeavour to clear up an obscure subject thoroughly, he will probably know more about it than a person who seeks to master it in a short time. Theory is instituted that each person in succession may not have to go through the same labour of clearing the ground and toiling through it, but may find the thing in order, and light admitted on it.*" (Source: Carl von Clausewitz: *On War*, Book 2, Chapter 2—On the Theory of War, translated by Colonel James John Graham, London, N. Trübner, 1873. http://clausewitz.com/readings/OnWar1873/TOC.htm)

1 Introduction

The order of actions in a sequential contest can generate various psychological effects that may change the ex-ante winning probabilities of the contestants. Soccer penalty shootouts, used to decide a tied match in a knockout tournament, offer a natural laboratory to test whether teams with equally skilled players have the same probability to win. Since the team favoured by a coin toss can decide to kick the first or the second penalty in all rounds (before June 2003, this team was automatically the first-mover), psychological pressure may mean that the second-mover has significantly less than 50% chance to win.

Previous research has shown mixed evidence and resulted in an intensive debate. While most authors find that the starting team enjoys an advantage (Apesteguia and Palacios-Huerta, 2010; Palacios-Huerta, 2014; Da Silva et al., 2018; Rudi et al., 2019), some papers do not report such a problem (Kocher et al., 2012; Arrondel et al., 2019). Vandebroek et al. (2018) argue that this disagreement is mainly due to inadequate sample sizes, thus the natural solution would be to design and implement an appropriate field experiment—but this would take years.

Nonetheless, almost all stakeholders recognise that penalty shootouts are potentially *unfair*. According to a survey, more than 90% of coaches and players choose to increase the psychological pressure on the other team by kicking the first penalties (Apesteguia and Palacios-Huerta, 2010). The IFAB (International Football Association Board), the rule-making body of soccer has recently planned to consult "*a potentially fairer system of taking kicks from the penalty mark*" (IFAB, 2017, 2018, Section "The future"). A straightforward alternative, the Alternating ($ABBA$) rule, has been trialled in some matches (Csató, 2020). However, the 133rd Annual Business Meeting (ABM) of the IFAB has decided to stop the experiment due to "*the absence of strong support, mainly because the procedure is complex*" (FIFA, 2018).

In our opinion, with the increasing use of information technology in a soccer game such as the video assistant referee (VAR), the fairness of penalty shootouts will probably emerge as a topic of controversy in the future, and some alternative mechanisms will be tested on the field. Even though the Alternating ($ABBA$) sequence has been found to provide no advantage for any player in a tennis tiebreak (Cohen-Zada et al., 2018), and Monte Carlo simulations show that it substantially mitigates the bias in soccer (Del Giudice, 2019), this is not the only solution to ensure fairness. Palacios-Huerta (2012) has argued to follow the Prouhet-Thue-Morse sequence, where the first n moves are mirrored in the next n. Recent suggestions include the Catch-up rule (Brams and Ismail, 2018), its variant called the Adjusted Catch-up rule (Csató, 2020), and the Behind-first rule (Anbarcı et al., 2019).

We contribute to the topic by evaluating all these mechanisms in three different probabilistic models that reflect the potential advantage of the team kicking the first penalty. Besides summarising probably all findings of the previous literature on penalty shootout rules, this comprehensive review attempts to quantify the fairness of seven different designs and to consider other aspects such as the probability of reaching the sudden death stage, or the complexity of each rule. Wherever possible, analytical proofs are provided for the conjectures derived from numerical calculations as Propositions 1 and 2 illustrate.

From the plethora of interesting findings, the following results are worth underlining. The Catch-up rule turns out to be inferior compared to the straightforward Alternating ($ABBA$) rule as it does not yield any gain in fairness but it is more complex. On the other hand, the Behind-first mechanism—that alternates the shooting order but guarantees the first penalty to the team lagging behind in each round—somewhat overperforms them concerning fairness, which can be improved further by compensating the second-mover in the sudden death based on an idea of Csató (2020). This Adjusted Behind-first rule, introduced in the current work, emerges as a promising alternative to the standard design.

Why is such an extensive but abstract study necessary? Firstly, empirical results are often disputed due to problems with the selection and the size of the datasets. Secondly, the mechanisms proposed recently have never been implemented in practice. Thirdly, since field experiments take a long time to carry out, it would be beneficial to filter out poor alternative mechanisms and save the resources for theoretically attractive policy changes.

The paper proceeds as follows. Section 2 presents the penalty shootout designs to be compared and discusses three ways to mathematically formalise psychological pressure. Our findings are detailed in Section 3. Section 3.1 computes the probability of winning in the sudden death stage, Section 3.2 summarises the theoretical results on the fairness of penalty shootout mechanisms, while Section 3.3 contains the numerical calculations for a large set of parameters. Section 3.4 deals with some questions beyond fairness. Finally, Section 4 offers concluding remarks.

2 Mechanisms and probability models

Denote the team that kicks the first penalty by A, and the other team by B. The penalty shootout consists of five rounds in its regular phase. In each round, both teams kick one penalty. The shooting order in a round can be (1) independent of the outcomes in the previous rounds (*deterministic* rule); (2) influenced by the results of preceding penalties (*stochastic* rule). The scores are aggregated after the five rounds, and the team which has scored more goals than the other wins the match. If the scores are level, the *sudden death* stage starts and continues until one team scores a goal more than the other from the same number of penalties.

2.1 Penalty shootout designs

We investigate three deterministic procedures:

- *Standard (ABAB)* rule: team A kicks the first, and team B kicks the second penalty in each round. This is the official soccer penalty shootout design.
- *Alternating (ABBA)* rule: the order of the teams alternates, the second round (BA) mirrors the first (AB), and this sequence continues without any change, even in the possible sudden death phase.
- $ABBA|BAAB$ rule: the order in the first two rounds is $ABBA$, which is mirrored in the next two ($BAAB$), and this sequence is repeated.

The "double alternating" $ABBA|BAAB$ mechanism is considered as it takes us one step closer to the Prouhet-Thue-Morse sequence than the Alternating ($ABBA$) design. In our opinion, it is unlikely that the administrators want to move further along this line.

There are two stochastic designs, both of them having two variants:

- *Catch-up* rule (Brams and Ismail, 2018): the first kicking team alternates but the shooting order does not change if the first kicker missed and the second succeeded in the previous round.

- *Adjusted Catch-up* rule (Csató, 2020): the first five rounds are designed according to the Catch-up rule, however, team B kicks the first penalty in the sudden death stage (sixth round) regardless of the outcome in the previous round.

- *Behind-first* rule (Anbarcı et al., 2019): the team having less score after some rounds kicks the first penalty in the next round, and the order of the previous round is mirrored if the score is tied.

- *Adjusted Behind-first* rule: the first five rounds are designed according to the Behind-first rule, however, team B kicks the first penalty in the sudden death stage (sixth round) regardless of the outcome in the previous round.

The Adjusted Behind-first mechanism applies the idea underlying the Adjusted Catch-up rule, introduced in Csató (2020), for the Behind-first design. According to our knowledge, it is first defined here.

Table 1: An illustration of the seven penalty shootout mechanisms

		Penalty kicks in the regular phase										Sudden death			
Mechanism	Team	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ABAB	A	✗		✔		✗		✔		✗		✔		✔	
	B		✔		✔		✗		✗		✗		✔		✗
ABBA	A	✗			✔	✗			✔	✗			✔	✔	
	B		✔	✔			✗	✗			✗	✔			✗
ABBA|*BAAB*	A	✗			✔		✗	✔		✗			✔		✔
	B		✔	✔		✗			✗		✗	✔		✗	
Catch-up	A	✗		✔			✗	✔			✗	✔			✔
	B		✔		✔	✗			✗	✗			✔	✗	
Adj. Catch-up	A	✗		✔			✗	✔			✗		✔	✔	
	B		✔		✔	✗			✗	✗		✔			✗
Behind-first	A	✗		✔		✗		✔			✗	✔			✔
	B		✔		✔		✗		✗	✗			✔	✗	
Adj. Behind-first	A	✗		✔		✗		✔			✗		✔	✔	
	B		✔		✔		✗		✗	✗		✔			✗

Table 1 illustrates the seven penalty shootout designs. Since the scores are 3-3 after five penalties, the shootout goes to sudden death, where both teams succeed in the sixth round. However, in the seventh round only team A scores, implying that it wins the match.

Note that the four stochastic mechanisms lead to different shooting orders. Team B kicks the first penalty in the third round under the Catch-up rule because both teams miss in the second round where team A is the first-mover. On the other hand, the Behind-first rule favours team A in the third round as it is lagging in the number of goals. Both designs give the first penalty in the sixth round to team A since team B is the first-mover in the fifth round. However, the Adjusted Catch-up and Behind-first rules compensate team B by kicking first in the sudden death for being disadvantaged in the first round of the shootout.

2.2 Models of psychological pressure

According to Apesteguia and Palacios-Huerta (2010, Figure 2A), the first kicking team scores its penalties with a higher probability in all rounds. A possible reason is that most soccer penalties are successful, thus the player taking the second kick usually faces greater mental pressure. Therefore, following Brams and Ismail (2018) and Csató (2020), our first *model M1* assumes that the first kicker has a probability p of scoring, while the second kicker has a probability q of scoring, where $p \geq q$.

However, this anxiety may be missing if the first kicker fails. The second *model M2* assumes that each player has a probability p of scoring, except for the second shooter after a successful penalty, who scores with probability q, where $p \geq q$.

Finally, the psychological pressure can come from lagging in the number of goals as Vandebroek et al. (2018) argue. Consequently, the third *model M3* is defined such that each player has a probability p of scoring, except for the kicker from the team having fewer scores, who succeeds with probability q, where $p \geq q$.

Example 2.1. Consider a penalty shootout, which stands at 2-3 with the first-mover in the fourth round lagging behind. The probability models above have the following implications (recall that $p \geq q$).

- Model M1: The 7th penalty is scored with probability p. The 8th penalty is scored with probability q.
- Model M2: The 7th penalty is scored with probability p. The 8th penalty is scored with probability p if the 7th penalty was unsuccessful, and with probability q if the 7th penalty was successful.
- Model M3: The 7th penalty is scored with probability q. The 8th penalty is scored with probability p.

All of the above models are mainly consistent with previous observations. Empirical testing of their validity is beyond the scope of the current work.

3 Results

The fairness of the soccer penalty shootout is usually interpreted such that no team should have an advantage because of winning or losing the coin toss. Therefore, a mechanism is called *fairer* than another if the probability of winning is closer to 0.5 for two equally skilled teams.

3.1 The probability of winning in the sudden death

The first five rounds of a shootout can be represented by a finite sequence of binary numbers as each penalty is either scored or missed. However, the sudden death has no definite end, thus it is necessary to calculate the winning probabilities in this phase by hand. With a slight abuse of notation, denote by $W(A)$ the winning probability of the team kicking the first penalty in the sudden death.

In model M1, Brams and Ismail (2018, p. 192) derives for the Standard ($ABAB$) rule that

$$W_1^S(A) = \frac{p(1-q)}{p+q-2pq}.$$

Models M2 and M3 are equivalent in the sudden death, where only the second team can have less goals scored. Hence $W_2^S(A) = p(1-q) + [pq + (1-p)(1-p)]\, W_2^S(A)$, which leads to

$$W_2^S(A) = \frac{p(1-q)}{2p - pq - p^2}.$$

The $ABBA$, (Adjusted) Catch-up, and (Adjusted) Behind-first mechanisms coincide in the sudden death where they imply an alternating order of kicking.

In model M1, according to Brams and Ismail (2018, p. 192):

$$W_1^R(A) = \frac{1 - q + pq}{2 - p - q + 2pq}.$$

In models M2 and M3, $W_2^R(A) = p(1-q) + [pq + (1-p)(1-p)] \left[1 - W_2^R(A)\right]$ as the first penalty in the second round of the sudden death is kicked by team B, that is,

$$W_2^R(A) = \frac{1 - p + p^2}{2 - 2p + pq + p^2}.$$

Finally, the sudden death of the $ABBA|BAAB$ rule is the most complicated one. Here there are two different cases: (a) the winning probability is $W(AA)$ when team A kicks the first penalty in the first two rounds of this phase, which is followed by two rounds with team B being the first kicker; and (b) the winning probability is $W(AB)$ when team A kicks the first penalty in the first round of the sudden death, continued with two rounds where team B is the first kicker.

Consider model M1. The first round of the sudden death is won by team A with probability $p(1-q)$. The sudden death reaches its second round with probability $pq + (1-p)(1-q)$. It is won by the team kicking the first penalty in the second round with probability $p(1-q)$, while it is won by the team kicking the second penalty in the second round with probability $(1-p)q$. Consequently,

$$W_1(AA) = p(1-q) + [pq + (1-p)(1-q)]\, p(1-q) + [pq + (1-p)(1-q)]^2 \,[1 - W_1(AA)]$$

$$W_1(AA) = \frac{p(1-q) + (1-p-q+2pq)\, p(1-q) + (1-p-q+2pq)^2}{1 + (1-p-q+2pq)^2}.$$

Analogously,

$$W_1(AB) = p(1-q) + [pq + (1-p)(1-q)]\, (1-p)q + [pq + (1-p)(1-q)]^2 \,[1 - W_1(AB)]$$

$$W_1(AB) = \frac{p(1-q) + (1-p-q+2pq)\, (1-p)q + (1-p-q+2pq)^2}{1 + (1-p-q+2pq)^2}.$$

Consider models M2 and M3, which contain two differences compared to model M1: (1) in any round, the winning probability of the second kicker is $(1-p)p$ instead of $(1-p)q$; and (2) the probability of reaching the next round is $pq + (1-p)^2$ instead of $pq + (1-p)(1-q)$. Hence, similar calculations show that

$$W_2(AA) = \frac{p(1-q) + (1-2p+pq+p^2)\, p(1-q) + (1-2p+pq+p^2)^2}{1 + (1-2p+pq+p^2)^2},$$

and

$$W_2(AB) = \frac{p(1-q) + (1-2p+pq+p^2)\, (1-p)p + (1-2p+pq+p^2)^2}{1 + (1-2p+pq+p^2)^2}.$$

3.2 Theoretical findings

Despite the relatively simple mathematical models, it is non-trivial to obtain analytical statements as five rounds of penalties mean $2^{10} = 1024$ different scenarios, and the probability of each contains ten items from the set of p, q, $(1-p)$, and $(1-q)$. In practice, the shootout is finished if one team has scored more goals than the other could score, but this consideration does not decrease significantly the number of cases. We summarise two results from the previous literature and prove a third for the stochastic designs in model M3.

According to Echenique (2017), a crucial condition of the psychological advantage enjoyed by the first-mover is the following: the probability that the first shooter scores and the second shooter misses is greater than the probability that the first shooter misses and the second scores. Models M1 and M2 satisfy this requirement as $p > q$ implies $p(1-q) > (1-p)q$ and $p(1-q) > (1-p)p$. Echenique (2017, Proposition 2) states that the condition is sufficient and necessary for the first shooter advantage under the Standard ($ABAB$) and the Alternating ($ABBA$) rules. Furthermore, the Alternating ($ABBA$) mechanism is always fairer than the Standard ($ABAB$) mechanism if it holds.

Regarding model M3, the winning probability of the first team does not depend on the number of rounds in the regular phase under the Standard ($ABAB$) model (Vandebroek et al., 2018, p. 735).

The following result is first presented here.

Proposition 1. *The Catch-up and Behind-first rules lead to the same winning probabilities in model M3.*

Proof. The probability of each possible outcome is shown to be the same under the Catch-up and Behind-first rules. This probability is the product of the individual probabilities in every round.

Assume that the Catch-up and Behind-first rules differ in the probability of the kth round. Hence the order of the teams in the kth round should be different under the two rules, CD for the Catch-up and DC for the Behind-first. If the scores of the teams are different at the beginning of the kth round, then the team lagging behind has the probability q of scoring, and the other team has the probability p of scoring. The commutative property of multiplication means that the probability of the kth round is independent of the shooting order, which contradicts the assumption above. To conclude, the teams should be tied at the beginning of the kth round and their shooting under should be different under the two mechanisms.

Consider the case when the shooting order in the $(k-1)$th round was CD by the Catch-up rule. Consequently, team C missed, while team D succeeded in the $(k-1)$th round as otherwise, the Catch-up rule would imply the alternated order DC in the kth round. Therefore, team D was lagging behind (by one goal) at the beginning of the $(k-1)$th round, thus the shooting order in the $(k-1)$th round was DC by the Behind-first rule. It is a contradiction since no team is lagging behind at the beginning of the kth round and the order of the previous round was DC under the Behind-first rule, thus it should be CD in the kth round according to this mechanism.

Consider the case when the shooting order in the $(k-1)$th round was DC by the Catch-up rule. Then there are three possibilities:

- *Team C scored and team D failed in the $(k-1)$th round*
  This contradicts the assumption that the shooting order in the kth round is CD

under the Catch-up rule as this mechanism implies the unchanged order DC in the kth round.

- *Team C failed and team D scored in the $(k-1)$th round*
  Since the teams are tied at the beginning of the kth round, team D was lagging in the score at the beginning of the $(k-1)$th round. The shooting order in the $(k-1)$th round was DC under the Behind-first rule, contradicting the shooting order DC in the kth round by the Behind-first rule as this mechanism implies an alternated order if the score is tied, which is the case for the kth round.

- *Both teams failed or both teams scored in the $(k-1)$th round*
  Since the scores were level at the beginning of the $(k-1)$th round, too, the shooting order was CD under the Behind-first rule. The repetition of the arguments above results in by backward induction that the teams should have been tied at the beginning of the penalty shootout, which is trivial, and the shooting order in the first round should have been different under the two mechanisms, which is impossible.

The proof is completed as the assumption leads to a contradiction in each of the possible cases discussed above. □

Corollary 3.1. *The Adjusted Catch-up and Adjusted Behind-first rules imply the same winning probabilities in model M3.*

Proof. The Adjusted Catch-up and Adjusted Behind-first rules coincide with the Catch-up and Behind-first rules, respectively, in the regular stage. Both of them give the first penalty of the sudden death to team B, and they follow an alternating order in this phase. □

3.3 The analysis of fairness

The probability of winning can be accurately determined by a computer code for any values of p and q to gain an insight into the fairness of the presented penalty shootout mechanisms.

First, similarly to Brams and Ismail (2018), we use the values $p = 3/4$ and $q = 2/3$ that are close to the empirical scoring probabilities. Table 2 reveals the probability of winning for team A that kicks the first penalty. The Standard ($ABAB$) rule is the most unfair, and it deviates from equality even further as the number of rounds in the regular stage grows. Some rules, especially the $ABBA|BAAB$, can favour the second-mover team B in models M1 and M2 but not in M3. The Alternating ($ABBA$), (Adjusted) Catch-up, and (Adjusted) Behind-first designs exhibit a moderate odd-even effect: they are less fair when the number of rounds is odd, which restricts the opportunities to balance the advantage enjoyed by team A. There are similar cycles with the length of four under the $ABBA|BAAB$ rule.

The calculations reflect the theoretical contributions: (a) the Alternating ($ABBA$) mechanism outperforms the Standard ($ABAB$) mechanism in models M1 and M2 (Echenique, 2017); (b) the fairness of the Standard ($ABAB$) rule is not influenced by the number of rounds in model M3 (Vandebroek et al., 2018); (c) the (Adjusted) Catch-up and the (Adjusted) Behind-first designs lead to the same winning probabilities in model M3 (Proposition 1). All rules are fairer in model M2 compared to model M1 because the preference towards the first shooter is weaker in the former model.

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

(a) Model M1: the second player has a scoring probability q

Model M1	Number of rounds							
Mechanism	1	2	3	4	**5**	6	7	8
$ABAB$	0.600	0.608	0.618	0.628	**0.637**	0.645	0.653	0.661
$ABBA$	0.526	0.511	0.519	0.508	**0.515**	0.507	0.513	0.506
$ABBA\|BAAB$	0.513	0.494	0.489	0.504	**0.509**	0.497	0.492	0.503
Catch-up	0.526	0.516	0.518	0.513	**0.514**	0.512	0.512	0.511
Adjusted Catch-up	0.526	0.495	0.515	0.501	**0.509**	0.504	0.507	0.504
Behind-first	0.526	0.516	0.516	0.512	**0.512**	0.510	0.510	0.508
Adjusted Behind-first	0.526	0.495	0.512	0.500	**0.506**	0.501	0.503	0.501

(b) Model M2: the second player has a scoring probability q if the first player succeeds

Model M2	Number of rounds							
Mechanism	1	2	3	4	**5**	6	7	8
$ABAB$	0.571	0.578	0.586	0.593	**0.600**	0.606	0.612	0.618
$ABBA$	0.520	0.508	0.514	0.506	**0.511**	0.505	0.509	0.504
$ABBA\|BAAB$	0.510	0.496	0.492	0.503	**0.506**	0.497	0.494	0.502
Catch-up	0.520	0.513	0.514	0.510	**0.511**	0.509	0.509	0.508
Adjusted Catch-up	0.520	0.497	0.510	0.502	**0.506**	0.503	0.505	0.503
Behind-first	0.520	0.513	0.513	0.510	**0.509**	0.508	0.507	0.507
Adjusted Behind-first	0.520	0.497	0.509	0.501	**0.505**	0.502	0.503	0.501

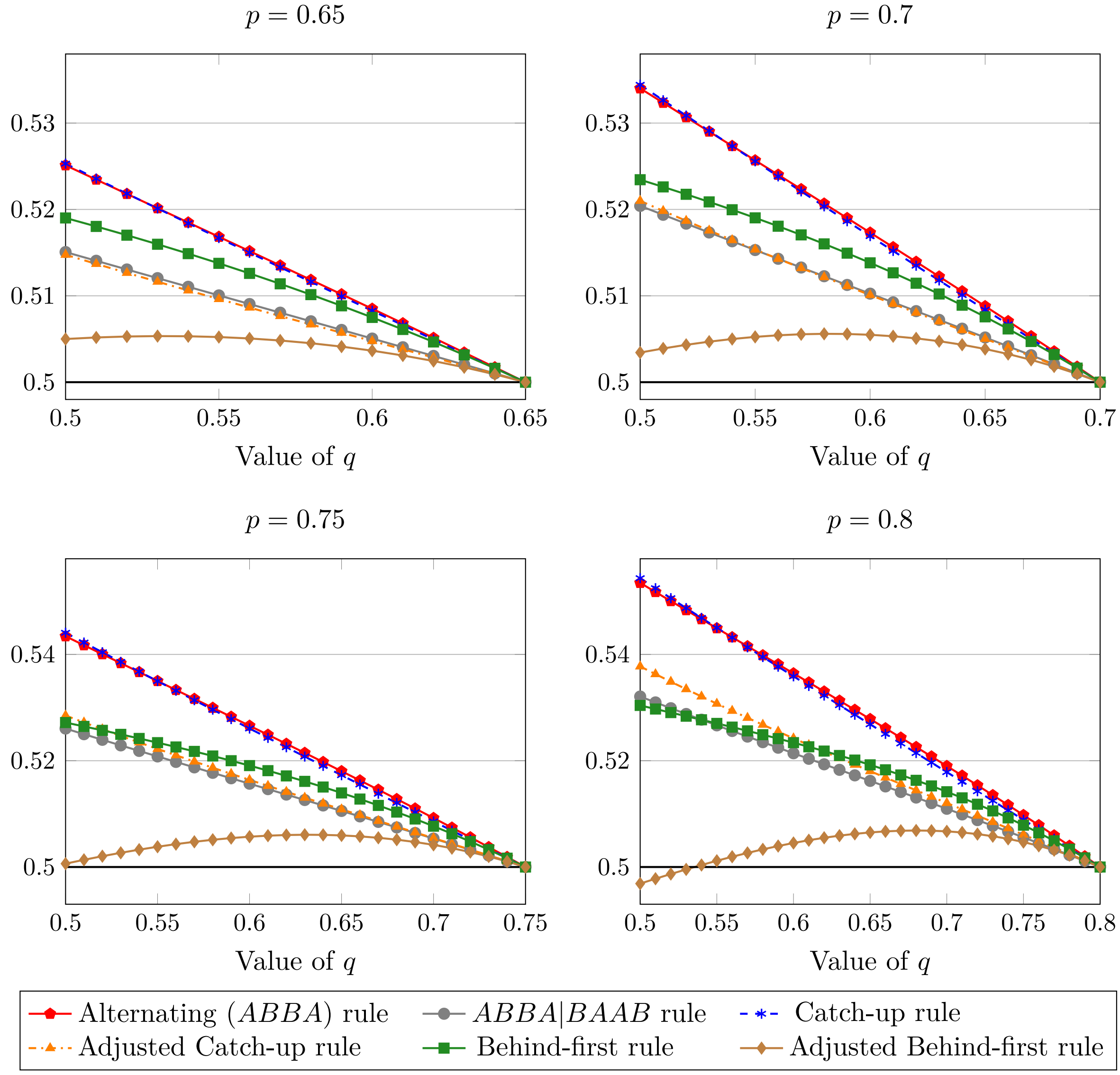
(c) Model M3: the team lagging behind has a scoring probability q

Model M3	Number of rounds							
Mechanism	1	2	3	4	**5**	6	7	8
$ABAB$	0.571	0.571	0.571	0.571	**0.571**	0.571	0.571	0.571
$ABBA$	0.520	0.516	0.514	0.514	**0.512**	0.512	0.511	0.511
$ABBA\|BAAB$	0.510	0.504	0.507	0.510	**0.507**	0.504	0.507	0.508
Catch-up	0.520	0.515	0.515	0.513	**0.513**	0.512	0.511	0.511
Adjusted Catch-up	0.520	0.501	0.513	0.506	**0.510**	0.507	0.508	0.508
Behind-first	0.520	0.515	0.515	0.513	**0.513**	0.512	0.511	0.511
Adjusted Behind-first	0.520	0.501	0.513	0.506	**0.510**	0.507	0.508	0.508

The Catch-up rule is not fairer than the already tried Alternating ($ABBA$) rule, and the Behind-first rule does not perform considerably better than them. However, the minor amendment proposed by Csató (2020) in the first round of the sudden death consistently makes the stochastic designs fairer. Thus they become a competitive alternative to the deterministic $ABBA|BAAB$ that aims to implement the Prohuet-True-Morse sequence, especially in model M1.

In order to extend these findings, Figures 1–3 plot the winning probability of team A as the function of parameter q—that has a different meaning in each model—for four

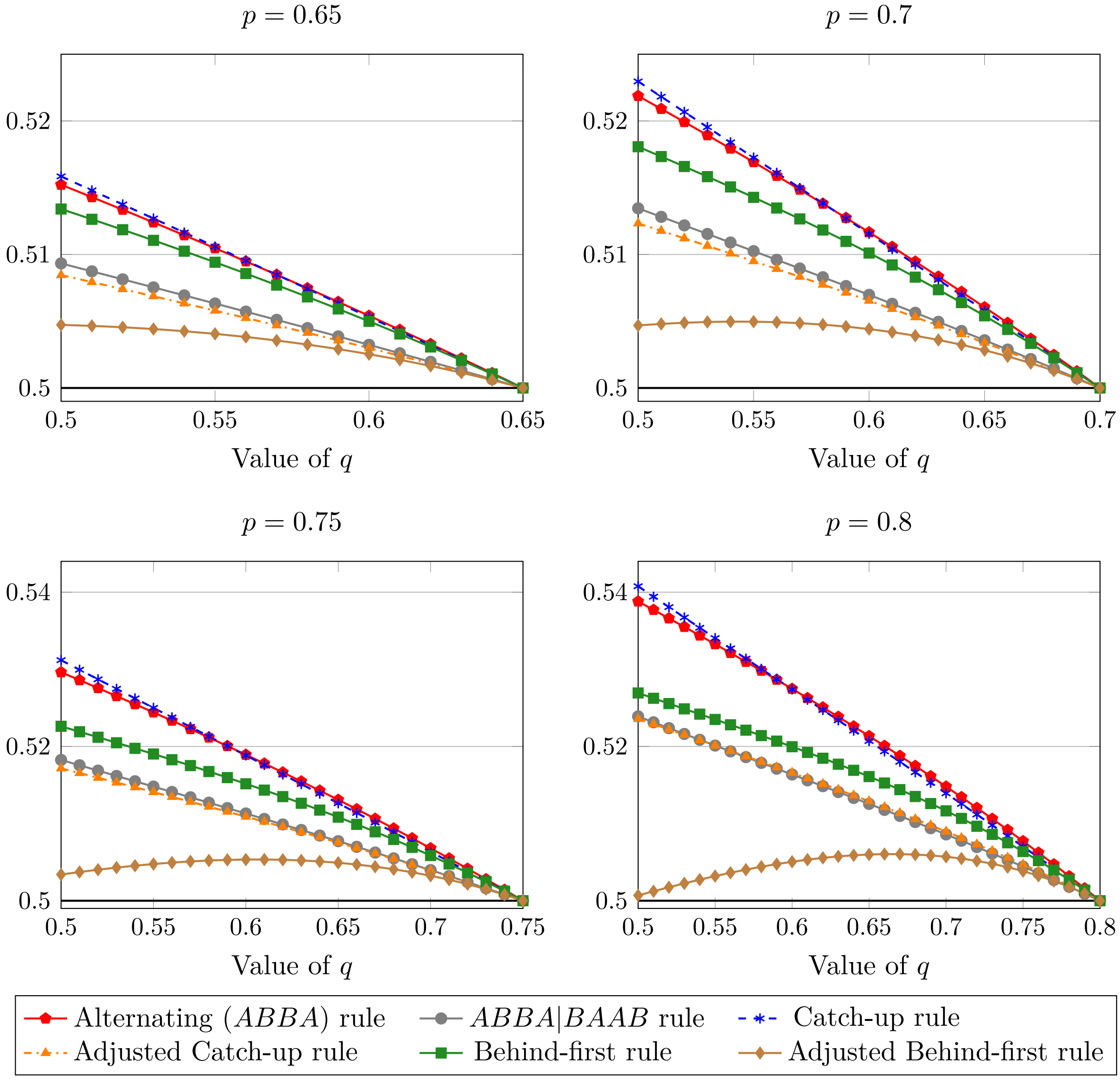
Figure 1: The probability that team A wins a penalty shootout over five rounds including sudden death, model M1



values of p, the scoring probability of the advantaged team. The Standard ($ABAB$) rule is not depicted due to its high level of unfairness, which makes it impossible to visualise this mechanism together with the other six. The lessons from these graphs can be summarised as follows:

- All mechanisms are closer to fairness in model M2 than in model M1 because the former punishes the second player only if the first player scores.
- Fairness is the most difficult to achieve in model M3, where the disadvantage of team B cannot be balanced by providing it with a higher scoring probability as the first kicker in a round. On the other hand, the Standard ($ABAB$) rule is the least unfair in model M3.
- The straightforward Alternating ($ABBA$) rule is not worse than the stochastic Catch-up rule, the use of the latter cannot be justified by the need to improve

Figure 2: The probability that team A wins a penalty shootout over five rounds including sudden death, model M2



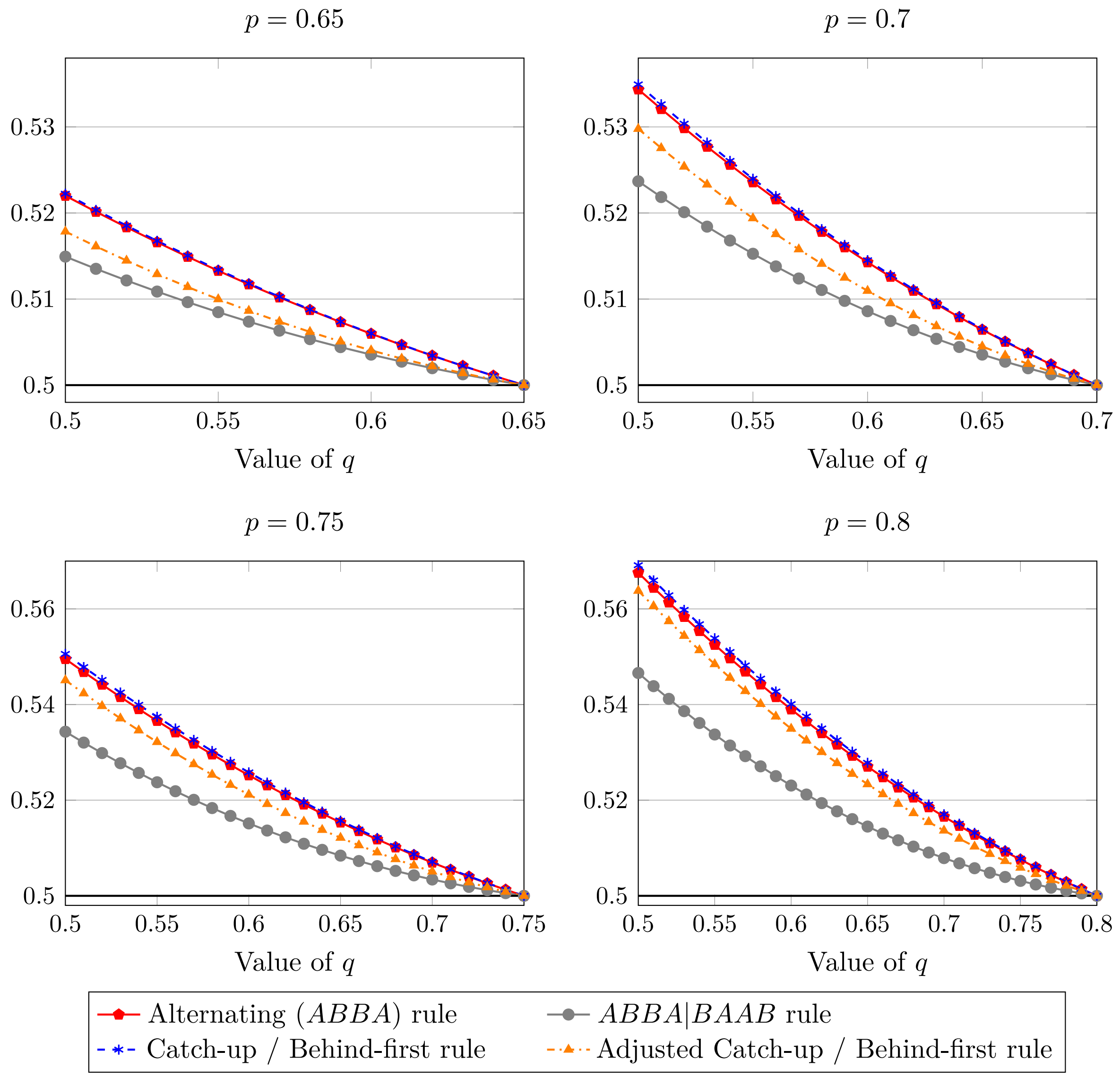
fairness. This finding considerably decreases the value of the contribution by Brams and Ismail (2018).

- The Behind-first mechanism outperforms the Alternating ($ABBA$) and the Catch-up designs, except for model M3, where it is equivalent to the Catch-up rule according to Proposition 1.

- The adjustment of the stochastic designs, suggested by Csató (2020), robustly improves fairness. It is worth guaranteeing the first penalty of the sudden death stage for team B.

- The $ABBA|BAAB$ rule remains competitive with the Adjusted Catch-up rule in models M1 and M2, while it is the closest to fairness among all designs considered here in model M3. If the Alternating ($ABBA$) rule is judged inadequate, and the use of stochastic mechanisms should be avoided, further steps towards the Prohuet-True-Morse sequence can be effective to increase fairness.

Figure 3: The probability that team A wins a penalty shootout over five rounds including sudden death, model M3



- The Adjusted Behind-first rule outperforms all other mechanisms if the first-mover advantage originates exclusively from the shooting order, that is, model M1 or model M2 is valid. The compensation of team B in the sudden death is so effective that the winning probability of team A becomes non-monotonic as the difference between p and q grows. However, the psychological pressure is unlikely to reach this level in practice, the observation remains a theoretical curiosity.

3.4 Further issues

In the following, we discuss four other topics: the expected length of the sudden death stage and the probability of reaching it, as well as the complexity and the strategy-proofness of the mechanisms.

Denote the expected length of the sudden death by ε, and the probability that it finishes in a given round by R. Then

$$\varepsilon = R + (1 - R)(1 + \varepsilon)$$

because the expected length of the sudden death is $1 + \varepsilon$ if it is not decided in the first round. Consequently, $\varepsilon = 1/R$, which does not depend on the penalty shootout mechanism.

In model M1, $R_1 = p(1-q) + (1-p)q = p + q - 2pq$, which has already been calculated in Brams and Ismail (2018, p. 193). In models M2 and M3, $R_2 = p(1-q) + (1-p)p = 2p - pq - p^2$, thus the sudden death is expected to be shorter here compared to the previous model.

Table 3: The probability that a penalty shootout over five rounds is continued with sudden death

Mechanism	Model M1	Model M2	Model M3
$ABAB$	0.263	0.260	0.215
$ABBA$	0.275	0.266	0.215
$ABBA\|BAAB$	0.275	0.266	0.215
(Adjusted) Catch-up	0.284	0.274	0.215
(Adjusted) Behind-first	0.319	0.299	0.215

On the other hand, the probability that the sudden death stage is reached can be influenced by the penalty shootout design. Table 3 reports these values. Note that the adjustment does not affect the stochastic mechanisms in the regular phase. Penalty shootouts taken by the Alternating ($ABBA$) and the $ABBA|BAAB$ designs have the same probability to continue with sudden death in models M1 and M2, where the scoring probabilities depend only on the shooting order, and both rules provide three penalties for team A and two penalties for team B as the first kicker. The sudden death can be avoided with the highest probability in model M3, which punishes the team already lagging behind.

Since the sudden death is perhaps the most dramatic part of a penalty shootout, media attention can be maximised by the (Adjusted) Behind-first rule, followed by the (Adjusted) Catch-up. Deterministic mechanisms seem to be unfavourable from this point of view.

The last column of Table 3 refers to a possible analytical result in the case of model M3. This is verified below.

Proposition 2. *In model M3, the probability of reaching the sudden death is not influenced by the penalty shootout mechanism, that is, by the shooting order.*

Proof. We focus on the standing of the penalty shootout after some rounds. The possible states can be distinguished by the difference between the number of goals scored by the two teams, denoted by k. The transition probabilities from any round to the next round can be computed easily.

- *The score difference is* $k = 0$
  The scores will remain tied if both teams fail in the next round with probability $(1-p)(1-p)$, or both teams succeed in the next round with probability pq. Otherwise, the shootout goes to state $k = 1$.

- *The score difference is* k
  The score difference will remain k if both teams fail in the next round with

probability $(1-p)p$, or both teams succeed in the next round with probability pq. The shootout goes to state $k+1$ if the team lagging behind misses its penalty, while the other team scores. This has the probability $p(1-q)$, independently of the shooting order. Otherwise, if the team lagging behind scores and the other team misses, the penalty shootout goes to state $k-1$ with probability $(1-p)q$.

Table 4: Transition probabilities between the states of a penalty shootout, model M3

To → From ↓	Possible states Score diff. $k=0$	Score diff. $k=1$	Score diff. $k=2$	$\cdots$
Score diff. $k=0$	$pq+(1-p)(1-p)$	$p(1-q)+(1-p)p$	0	$\cdots$
Score diff. $k=1$	$(1-p)q$	$pq+(1-p)(1-p)$	$p(1-q)$	$\cdots$
Score diff. $k=2$	0	$(1-p)q$	$pq+(1-p)(1-p)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Table 4 overviews the possible moves between these states. Since the transition probabilities are independent of the mechanism used to determine the shooting order, the probability that the penalty shootout finishes in the state $k=0$ at the end of the regular phase—and continues with the sudden death stage—depends only on the parameters p and q. □

Csató (2020) quantifies the complexity of a penalty shootout design by the minimal number of binary questions required to decide which team is the first-mover in a given round, without knowing the history of the penalty shootout.

Proposition 3. *The complexities of the seven penalty shootout designs are as follows:*

- *Standard ($ABAB$) rule:* 0*;*
- *Alternating ($ABBA$) rule:* 1*;*
- *$ABBA|BAAB$ rule:* 1*;*
- *Catch-Up rule:* 2*;*
- *Adjusted Catch-Up rule: between* 2 *and* 3*;*
- *Behind-first rule:* 2*;*
- *Adjusted Behind-first rule: between* 2 *and* 3*.*

Proof. See Csató (2020) for the Standard ($ABAB$), Alternating ($ABBA$), Catch-up, and Adjusted Catch-up mechanisms.

The $ABBA|BAAB$ rule can be followed if it is known that the remainder after dividing the number of the rounds by four is from the set $\{0;1\}$ or from the set $\{2;3\}$.

The Behind-first rule requires two questions. If the result is tied, then the first kicker in the previous round determines the shooting order in the current round. Otherwise, if the result is not tied, one should know which team is lagging behind.

The Adjusted Behind-first mechanism is composed of the Behind-first rule in the regular phase (two questions), and the Alternating ($ABBA$) rule in the sudden death (one question), thus the number of binary questions needed is either two or three, depending on the answer to the first question. □

Proposition 2 reveals that the complexity measure of Csató (2020) is imperfect for deterministic mechanisms as the $ABBA|BAAB$ rule is more complicated than the Alternating ($ABBA$) rule according to common intuition. Perhaps the set of binary questions allowed can be restricted to get a better quantification of simplicity.

In deterministic mechanisms, the shooting order cannot be controlled by the teams. However, under a stochastic rule, the team kicking the second penalty might gain from deliberately missing it if the rewards outweigh the loss of an uncertain goal.

Strategy-proofness does not mean a serious problem according to the following results.

Proposition 4. *The (Adjusted) Catch-up rule is strategy-proof in model M1 if $p - q \leq 0.5$. It cannot be manipulated in models M2 and M3.*

Proof. See Brams and Ismail (2018, Proposition 4.1) for model M1.

The idea behind the proof of Brams and Ismail (2018, Proposition 4.1) can be followed for model M2. The expected gain of the second shooter from trying and succeeding after the first kicker fails is $1 + pq + p - p^2$ as it scores with probability $pq + (1 - p)p$ in the next round. The average gain from not trying is p because it is certainly the first-mover in the next round. The net gain of the second shooter is $1 + pq - p^2$.

For the first shooter, the expected gain in the next round is p goals if the second shooter tries and scores, otherwise the average gain is the scoring probability as the second-mover, $(1 - p)p + pq$. The net gain is $p^2 - pq$.

Consequently, the net difference in expected goals between the two teams at the end of the next round is $1 + pq - p^2 - (p^2 - pq) = 1$, which is always positive. There would be no incentive for any player to deliberately miss a penalty.

Since the scoring probabilities are independent of the shooting position in model M3, a deliberate miss cannot yield any profit.

The Adjusted Catch-up rule offers fewer opportunities to change the shooting order, hence it is strategy-proof if the Catch-up rule satisfies this condition. □

Since the difference between the scoring probability of the first and the second shooter is unlikely to be higher than 50%, the Catch-up rule is essentially non-manipulable.

Proposition 5. *The Behind-first rule is always strategy-proof.*

Proof. A team can become first-mover from being a second-mover only at the price of having a lower score, which cannot be optimal.

The Adjusted Behind-first rule offers fewer opportunities to change the shooting order, hence it is strategy-proof if the Behind-first rule satisfies this condition. □

4 Conclusions

We have analysed seven soccer penalty shootout rules in three different mathematical models. The Standard ($ABBA$) mechanism is the simplest but a robustly unfair design, at least compared to the others. The Alternating ($ABBA$) mechanism remains a straightforward and relatively fair solution, which is already used in tennis tiebreaks. The $ABBA|BAAB$ mechanism is a deterministic rule approaching the Prohuet-Thue-Morse sequence that further improves fairness and outperforms all other designs in model M3 when the team lagging behind has a lower scoring probability.

Two stochastic mechanisms have also been evaluated together with their slightly modified variants. The Catch-up rule yields no gain in fairness compared to the Alternating

($ABBA$) rule, thus it remains a poor choice to try. Its only advantage resides in the higher probability of reaching the most exciting sudden death stage but the fairer Behind-first rule is even better from this aspect. Both the Catch-up and the Behind-first designs can be adjusted according to the proposal of Csató (2020), by compensating team B—the second-mover in the first round—in the sudden death phase. These versions do not affect strategy-proofness and the frequency of sudden death, however, they take a step towards fairness at the price of marginally increasing complexity.

In the comparison of the designs, only one particular aspect of soccer penalties has been used, namely, that the majority of them are successful. Therefore, most findings remain valid and applicable in other sports. For instance, the order of actions influences multi-game chess matches, too, where playing white pieces confers an advantage for winning that is similar to kicking the first penalty in a soccer penalty shootout (González-Díaz and Palacios-Huerta, 2016).

While the IFAB has recently agreed to stop the experiments with fairer penalty shootout systems, the debate will probably continue as advancing to the next round in a knockout tournament has huge financial implications. In addition, the use of penalty shootouts can spread further in the future. For example, Palacios-Huerta (2018) recommends to apply them in the group matches of the 2026 FIFA World Cup. The theoretical results above can help to decide what mechanisms are worthwhile to choose for implementing on the field. Hopefully, our paper might also become a starting point and a standard reference for later research on soccer penalty shootouts.

Acknowledgements

László Csató, the father of the first author has contributed to the paper by writing the key part of the code making the necessary computations in Python.
We are deeply indebted to *Steven J. Brams* and *Mehmet S. Ismail*, whose work was a great source of inspiration.
We are grateful to *Tamás Halm* for useful advice.
The research was supported by the MTA Premium Postdoctoral Research Program grant PPD2019-9/2019, the NKFIH grant K 128573, and the ÚNKP-19-3-III-BCE-97 New National Excellence Program of the Ministry for Innovation and Technology.

References

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. (2020). A comparison of penalty shootout designs in soccer. *4OR*, in press. DOI: 10.1007/s10288-020-00439-w.

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.

Del Giudice, P. E. S. (2019). Modeling football penalty shootouts: how improving individual performance affects team performance and the fairness of the ABAB sequence. *International Journal of Sport and Health Sciences*, 13(5):240–245.

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.

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.

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.

Palacios-Huerta, I. (2018). Penalties for fair play. *New Scientist*, 238(3184):24–25.

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

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.