A general framework to quantify the event importance in multi-event contests*

Daniel Goller[1,2], Sandro Heiniger[2]

[1]Centre for Research in Economics of Education, University of Bern

[2]Swiss Institute for Empirical Economic Research, University of St.Gallen

July 7, 2022

Abstract

We propose a statistical framework for quantifying the importance of single events that do not provide intermediate rewards but offer implicit incentives through the reward structure at the end of a multi-event contest. Applying the framework to primary elections in the US, where earlier elections have greater importance and influence, we show that schedule variations can mitigate the problem of front-loading elections. When applied to European football, we demonstrate the utility and meaningfulness of quantified event importance in relation to the in-match performance of contestants, to improve outcome prediction and to provide an early indication of public interest.

Keywords: Incentives; Event importance; Multi-event Contest; Front-loading; European Football;

*We are grateful for helpful comments and feedback from Nora Bearth, Enzo Brox, Hannah Busshoff, Christian Deutscher, Riccardo Di Francesco, Michael Knaus, Alex Krumer, Michael Lechner, Fabian Muny, Gabriel Okasa, and Carina Steckenleiter. The usual disclaimer applies.

1 Introduction

Incentives are an important tool for motivating people to exert effort. In many environments, from the workplace to sporting contests, incentives are put in place to ensure that invested efforts are optimised to achieve predefined goals (Lazear, 2000; Laffont & Martimort, 2002). Changing incentives directly translates into altered performance or success probability (Rosen, 1985; Prendergast, 1999). In contests where the reward depends solely on the outcome of a single event, incentives are provided directly through the potential rewards. In multi-event contests, the translation of contest rewards into incentives for single events is not directly observable. Moreover, single events may be unequally important for obtaining the final rewards, e.g. performance in an interview is rated higher than the previous assessment centre test, or the results of previous events lead to momentum for subsequent events.

As this transmission of multi-event contest rewards into incentives reflects the (personal) expected rewards, incentives vary not only between events but also between participants. Understanding these disparate and possibly asymmetric incentives in multi-event contests is essential as it could lead participants to strategically allocate their efforts (Preston & Szymanski, 2003). Asymmetric individual incentives may have spillover effects on the outcome probabilities of all other participants, which in turn could lead to potentially unbalanced or unfair contests.

In this work, we propose a general statistical framework to quantify the importance, or implicit incentive, of single events in complex multi-event contests for each participant individually – the *event importance* (EI). The EI measures the capability of an event to change a participant's (expected) reward for a contest. Our approach is based on two steps. First, we calculate the probability distribution for each contestant to reach certain end-of-contest rewards based on an outcome model determining the outcome probabilities for every single event. Second, the importance of a single event is determined through the changes in the end-of-contest reward probabilities with respect to the possible outcomes of this particular event. If the probabilities to reach the final rewards are changed substantially, the importance of this event for the participant is high and our methodology returns a high event importance measure.

The proposed framework generalises previous approaches (Schilling, 1994; Scarf & Shi, 2008; Buraimo, Forrest, McHale, & Tena, 2022) in that it is suitable for any contest design and any number of participants and is not specific to any particular contest environment. Moreover, it allows for participant-specific reward structures and both the reward structure and the schedule can change dynamically during the contest. Crucial for practical usage is that the proposed statistical procedure can also be used in situations in which the importance of the single event potentially plays a role in the determination of the event's outcome – this can be accounted for by calculating the specific event importance in an iterative procedure.

Our statistical framework can be applied in a variety of practical use cases, like competing pharmaceutical companies developing a drug for the same medical indication, presidential elections which are held in a series of local elections, a job or promotion contest among applicants or workers, or sports tournaments.

To showcase our methodology, we use our framework in two applications. In the first, we apply the framework to the US presidential primaries to examine the problem of front-loading: Earlier elections are known to have a greater impact on the outcome of the nomination process, which is why several states are pushing for earlier election dates. We analyse the Democrats' electoral schedule for the 2020 primaries and compare it with two alternative hypothetical schedules, one sorted by the number of delegates and one randomised. In this analysis, we find that the positioning of a state's election in the schedule substantially affects its impact on the outcome of the nomination – indicated by higher event importance measures. A comparison of the different schedules shows that the problem of front-loading can be mitigated by arranging the schedule according to the number of delegates in the states and completely eliminated by a random scheduling.

In a second application of our framework to the double round-robin tournament structure in football leagues we provide explicit measures of the EI that express implicit incentives for teams. In this setting, the relevance of a particular match with respect to the team's expected rewards varies substantially, even though every match is actually awarded the same number of points. Hence, the importance of a match varies over the season and between teams. This leads to pairings between teams with potentially very different incentives that change the presumed probabilities of winning.

We demonstrate the meaningfulness of the derived values by analysing their relationship to various observable characteristics of the matches. The integration of the EI information into a prediction model improves the accuracy of match outcome forecasts. We show that bookmakers do not fully take into account the team-specific importance of events in their prediction model. Furthermore, a positive interrelation can be drawn between the estimated importance of the match and the public's interest in certain matches in the form of larger stadium attendance and social media engagement for more important matches. For the in-match activity of the players and the outcomes of the match, we observe a comprehensive pattern suggesting that teams approach more important matches with a more aggressive, direct, and successful playing style.

Both the event importance values and replication code for the applications are publicly available on Harvard Dataverse (Goller & Heiniger, 2022).[1] The rest of the paper is structured as follows. Section 2 discusses related literature. Section 3 explains the proposed statistical method. Section 4 applies the framework to the front-loading in US primaries and the application to double round-robin tournaments appears in Section 5. Section 6 concludes.

[1]For the US presidential primaries, the EI estimates are published for all states and territories, for the three different scheduling scenarios. The published EI estimates for the European football leagues cover the seasons 2009/10 through 2018/19 and all seven leagues, i.e., the German 'Bundesliga 1', the Dutch 'Eredivisie', the Spanish 'La Liga', the French 'Ligue 1', the English 'Premier League', the Portuguese 'Primeira Liga', and the Italian 'Serie A'.

2 Related literature

This work mainly refers to two types of literature, (a) the importance of specific events and attempts to quantify them and (b) the literature on incentives in contests. The literature investigating the role of (explicit) incentives on performance generally finds that higher incentives increase performance (Ehrenberg & Bognanno, 1990; Prendergast, 1999; Lazear, 2000). However, it is important to distinguish between effort- and skill-based tasks; in the latter, strong incentives can lead to performance decrements, a phenomenon known as choking-under-pressure (Ariely, Gneezy, Loewenstein, & Mazar, 2009; Cohen-Zada, Krumer, Rosenboim, & Shapir, 2017; Goller, 2022). While most studies, and all of the studies mentioned above, consider individual incentives, the role of team-based incentives on performance may be different (Alchian & Demsetz, 1972).

The importance of specific events in multi-event contests can be found in several fields of literature. The order of action literature finds that the outcome of the contest is influenced by the order of the events which has been shown in musical contests (Ginsburgh & van Ours, 2003), song contests (de Bruin, 2005), or judicial decisions (Danziger, Levav, & Avnaim-Pesso, 2011). More specifically, several works focus on potential advantages in the first event in (usually sequential) contests, like in R&D (Harris & Vickers, 1987), sports (Apesteguia & Palacios-Huerta, 2010), or elections (Klumpp & Polborn, 2006).

Research documents the differential importance of sequential elections in US presidential primaries. Surprising wins in early states led to momentum effects in the 2004 primaries (Knight & Schiff, 2010). In their work they find an unbalanced influence on the final result for voters in early and late elections. Klumpp and Polborn (2006) model this first-winner advantage for primaries – known as the New Hampshire effect – giving an explanation for the more intense campaigning in early elections. With more influence in the nomination process in the early events front-loading, i. e. states moving their elections to earlier dates, is well documented (Mayer & Busch, 2003).

Ridout and Rottinghaus (2008) find more attention of the candidates to the states the earlier their elections are. Moreover, they find that scheduling is more important than the delegates' count.

The first approaches to determining an event's importance use rather simplistic measures (Jennett, 1984) and basic contest structures (Schilling, 1994). Most influential was Schilling (1994)'s general idea of defining and calculating event importance in terms of how the probability to reach a final goal varies for different event outcomes. Recently, more sophisticated approaches have emerged, for instance, Scarf and Shi (2008) simulating probabilities of final contest rewards conditional on two different event outcomes. Lahvička (2015) and Buraimo et al. (2022) build on the ideas of Jennett (1984) and Schilling (1994) but simulate final standings in the ranking in a Monte Carlo simulation to estimate the importance of single events. A different objective is followed in Geenens (2014): The importance of a match with regard to its influence on the final contest outcome is investigated. This has an interesting use case to investigate contests from the neutral spectator's perspective but is conceptually very different from the importance of a match for the specific contestant.

The drawback of all the discussed approaches is that they are specific to a certain type of contest that is prevalent in sports, i. e. a fixed number of event outcomes and one specific reward (e.g., winner-takes-it-all contests). This does not encompass more complex or dynamic contest designs and reward structures, which are common in society. The approach we propose in the following section provides the flexibility to handle a variety of practical applications with a variety of contest and reward structures.

3 The event importance

3.1 Introduction to the general framework

The event importance measures the difference between the contest reward probability distributions induced by the possible outcomes of a single event. If the probabilities for the final rewards vary substantially with the differential outcomes of the examined event, its impact on the tournament reward is large and a high event importance measure is attributed.

To quantify the importance of a particular event we hence require the probability distribution of the contest rewards conditional on each possible outcome of the investigated event. To determine the probability distributions, our framework sets the outcome of the examined event accordingly and solves the remainder of the contest by successive evaluation of the outcome model. Subsequent to the examined event, whose outcome is set by the framework, all entities (outcomes, covariates, schedule) are subject to the probabilistic outcome model. The successive application of the outcome model until the end of the contest generates the probability distribution for the contest reward conditional on the initial outcome.

There are six valuable attributes of our approach: First, by evaluating the reward probability from the perspective of every contestant individually, the event importance measure is specific to every participant and not the event itself. Second, we do not impose narrow restrictions on the contest setup. Since the contest reward probability distribution is conditional on each possible event outcome, we only need to assume a finite number of contestants, a finite number of possible outcomes for an event, and a finite schedule for the contest. Moreover, the tournament rewards have to be measurable based on the outcomes of all single events.

Third, the reward structure can be any function of all single event outcomes or a final contest ranking if such exists, e. g. close-by ranks can be grouped together if valued equally. Fourth, the framework is not restricted to a specific schedule. As the probability distributions are calculated through a successive evaluation of all events in the contest, the reward probabilities encompass all the essential features of the schedule as well, e. g. early elimination of participants in the contest or differences in information sets induced by events held in parallel or sequentially. Fifth, the framework is not tied to a specific distance metric to calculate the difference in the reward probability distributions. Sixth, if the outcome model is not known, it can be estimated on training data using any well-suited statistical method.

In the following, the details of how the described framework can be implemented to determine the event importance values in a general case are outlined.

3.2 Technical implementation

3.2.1 Notation

This section defines the notation used to describe the general framework. Table 1 summarises the notation as a quick reference. Upper case letters denote random variables, lower case letters denote their realisations or other exogenous variables, and calligraphic letters are sets. Multi-character names, such as *EI* or function names, are evident choices.

Table 1. Notation

$t \in \mathcal{T}$	Time t in contest schedule $\mathcal{T}$
$\mathcal{T}_{t^-} = \bigcup_{t' \leq t} t'$; $\mathcal{T}_{t^+} = \bigcup_{t' > t} t'$	Sub-schedule up to ($^-$) or after ($^+$) time t
$e_{t,i} \in e_t$	An event $e_{t,i}$ held at time t
$k \in \mathcal{K}_e \subseteq \mathcal{K}$	Contestant k participates in event e
$x_e = \bigcup_{k \in \mathcal{K}_e} x_{e,k}$	Information on all contestants in event e
$\mathcal{Y}_e$	Set of all possible outcomes for event e
$y_e \in \mathcal{Y}_e$	Realised outcome of event e
$\text{out}(x_e) = \bigcup_{\mathcal{Y}_e} P[Y_e = y_e \mid X_e = x_e]$	Probabilistic outcome model
$\left\{\mathcal{T}_{t^+}, x_{t^+}\right\} = \text{gen}\left(\mathcal{T}_{t^-}, x_{t^-}, y_{t^-}\right)$	Outcome-dependent schedule and covariates
$r_k = \text{rew}_k\left(\bigcup_{\mathcal{T}} y_e\right)$	Reward for contestant k after contest $\mathcal{T}$
$\text{EI}_{e,k} = \text{dist}\left(\bigcup_{\mathcal{Y}_e} r_{k,y_e}, \text{out}(x_e)\right)$	Event importance for contestant k in event e

Note: For better readability, we omit the subscripts of $e_{t,i}$ if a notation is not tied to a particular event but holds for any arbitrary event e.

The contest is held along a finite schedule $\mathcal{T}$. Because of the implicit chronological ordering of $\mathcal{T}$ we can define the notation $\mathcal{T}_{t-} = \bigcup_{t' \leq t} t'$ and $\mathcal{T}_{t+} = \bigcup_{t' > t} t'$, denoting the sub-schedule up to and after time t. Multiple events $e_{t,i}$ can be held simultaneously at time t, in this case $e_t = \bigcup_i e_{t,i}$. A finite set of contestants $\mathcal{K}$ participate in the contest of which a subset $\mathcal{K}_e \subseteq \mathcal{K}$ participate in event e. For each event, a set of covariates $x_e = \bigcup_{k \in \mathcal{K}_e} x_{e,k}$ and its outcome $y_e = \bigcup_{k \in \mathcal{K}_e} y_{e,k}$ is observed.[2] The outcome Y_e of event e is a random variable that follows a conditional probabilistic outcome model $\text{out}(x_e) = \bigcup_{\mathcal{Y}_e} P[Y_e = y_e | X_e = x_e]$ which, in case out() is not known, is approximated by $\widehat{\text{out}}()$. In the description of the general framework we assume w.l.o.g. that $\text{out}_e()$ is known and uniform. The cases of an approximated outcome model $\widehat{\text{out}_e}()$ or event specific outcome models $\text{out}_e()$ can both be handled in the general framework.

The chronological feature of the events further allows to define the sets of covariates $x_{t-} = \bigcup_{e,t' \leq t} x_{e_{t'}}$ and $x_{t+} = \bigcup_{e,t' > t} x_{e_{t'}}$ which combine all information on events and participants taking place either up to or after time t. The analogous operation on the outcomes defines y_{t-} and y_{t+}. In settings where parts of the covariates x and/or the schedule $\mathcal{T}$ depend on past outcomes, they are generated at run time based on the previous outcomes by $\left\{\mathcal{T}_{t+}, x_{t+}\right\} = \text{gen}\left(\mathcal{T}_{t-}, x_{t-}, y_{t-}\right)$. After the full contest $\mathcal{T}$, the probability distribution of the final rewards is determined according to the valuation function $r_{k,y_e} = \text{rew}_k\left(\bigcup_{\mathcal{T}} y_e\right)$ which can be individually specific for every contestant k. The event importance $\text{EI}_{e,k} = \text{dist}\left(\bigcup_{\mathcal{Y}_e} r_{k,y_e}, \text{out}(x_e)\right)$ for contestant k in event e is the difference between the multiple probability distributions of the final rewards measured by any distance measure dist(). The distance function can incorporate the outcome probabilities $\text{out}_e(x_e)$ of the starting event as weights.

[2]In this context, observed refers to recording an outcome within the framework and not to the actual outcome of an event that may have been observed.

3.2.2 Algorithm

Algorithm 1 describes the computation of the event importance value for a competitor k in event $e_{t,j}$. Readers that are less familiar with the pseudo-code notation can consult the literal translation of the algorithm in Appendix A.

Algorithm 1: Event Importance

Data: $e_{t,i}, k, t, \mathcal{T}, x, \mathcal{Y}_{e_{t,i}}$

Result: Event importance for competitor k in event $e_{t,j}$

begin

  forall $\boldsymbol{y}_{e_{t,i}}$ *in* $\mathcal{Y}_{e_{t,i}}$ **do** ▷ We denote all conditional dependence on $\boldsymbol{y}_{e_{t,i}}$ by *

    $y^*_{e_{t,i}} \leftarrow \boldsymbol{y}_{e_{t,i}}$

    forall $e_{t,j}$ *in* e_t *with* $j \neq i$ **do**

      $y^*_{e_{t,j}} \leftarrow \text{out}\left(x_{e_{t,j}}\right)$

    $\left\{\mathcal{T}^*_{t+}, x^*_{t+}\right\} \leftarrow \text{gen}\left(\mathcal{T}_{t-}, x_{t-}, y^*_{t-}\right)$

    forall t' *in* $\mathcal{T}^*_{t+}$ **do**

      $y^*_{e_{t'}} \leftarrow \text{out}(x^*_{e_{t'}})$

      $\left\{\mathcal{T}^*_{t'+}, x^*_{t'+}\right\} \leftarrow \text{gen}\left(\mathcal{T}^*_{t'-}, x^*_{t'-}, y^*_{t'-}\right)$

    $r_{k,\boldsymbol{y}_{e_{t,i}}} \leftarrow \text{rew}_k\left(\bigcup_{\mathcal{T}} y^*_e\right)$

  $\text{EI}_{e_{t,i},k} \leftarrow \text{dist}\left(\bigcup_{\mathcal{Y}_{e_{t,i}}} r_{k,\boldsymbol{y}_{e_{t,i}}}, \text{out}(x_{e_{t,i}})\right)$

  return $EI_{e_{t,i},k}$

3.2.3 Approximation of the probability distributions

By the subsequent evaluation of the outcome model, the probability distribution of the rewards at the end of the season can be determined, independent of the contest design. However, a large amount of sequential events opens an extremely large number of possible outcome paths which causes numerical problems if their probability would be evaluated exactly. For an outcome model that depends on past outcomes, the outcome paths can additionally become very complex. For this reason, it is often appropriate to perform a Monte Carlo simulation to approximate the probability distribution of the final rewards. Each run simulates one path for the remaining contest according to the outcome model and the generated covariates/schedule information at run time. With an adequate number of N_{MC} Monte Carlo runs, the estimated probability distribution and hence the event importance values become sufficiently close to the exact values.

In a simulation that estimates the EI values for all events consecutively, a chronologically backward iteration over the events allows for the reuse of already evolved paths as they can be merged with respective previous outcome to longer paths and thus reduce the computational complexity. In this case, N_{MC} is an upper bound for the number of actually performed runs in each step and at the same time, a lower bound for the number of runs the event importance estimate is premised on.

3.2.4 Iterative approximation of event importance

In many applications, the event importance is an integral part of the outcome model, e. g. when the importance can be interpreted as an incentive for the contestants to provide effort that in turn influences the outcome of the event. Independent of whether the outcome model is known or not, it encompasses the event importance values which are not available beforehand.

To determine the unknown event importance values, Algorithm 1 is at first executed with an approximate outcome model that does not feature the event importance in the variable set. This returns an initial approximation of the desired EI values. A subsequent iterative application of Algorithm 1 with the full covariate set including the preliminary EI variables updates the event importance estimates accounting for their own impact through the outcome model. This iterative procedure can be continued until a predefined stopping criterion is reached. The application in Section 5 is an example of an outcome model which includes the event importance in the covariates. Algorithm 2 in Appendix C.3 illustrates how the iterative procedure is implemented in the context of the application.

3.2.5 Distance functions

To measure the difference between the probability distributions on the contest rewards, an appropriate distance function needs to be chosen. In simple applications with only binary event outcomes and a binary reward scheme, the difference between the contest-winning probabilities conditional on the event outcome is a straightforward choice as the distance function.

More complex cases which feature either multiple possible event outcomes or multiple rewards require a statistical distance function. For most settings, the Jensen-Shannon divergence (JSD) is an appropriate distance function to cope with multiple discrete probability distributions (Lin, 1991). It is a common distance measure (Nielsen, 2020) with desirable properties, such as allowing for a weighting of the probability distributions as well as being bounded and symmetrical. The JSD measures the difference in the Shannon entropy between the probability distributions which implies that it does not have an intuitive linear interpretation. If such an interpretation is of relevance, other candidates such as the total variation distance can be applied.

4 Application: Front-loading in US primaries

4.1 Introduction

Presidential primary elections in the United States are held by the Democratic and the Republican party to determine the presidential election nominees. Both parties follow a similar procedure where each state, every permanently inhabited US territory, and party members living abroad[3] are attributed a certain number of votes (pledged delegates). In addition to the pledged delegates, selected party officials' have additional votes (unpledged delegates) that are not tied to states' election results. In order to be nominated for the presidential elections, the candidates in the

[3]For the ease of readability, all entities are henceforth labelled as states.

primaries must receive the majority of the delegate votes.

Each state holds its election or caucus on an individually chosen date, and the election results determine how its delegates vote. Several states can vote on the same day, e.g. on "Super Tuesday" about one third of all delegate votes are determined. Due to the partially sequential nature of the primaries, it may happen that later elections become irrelevant to the outcome of the nomination if a candidate has already received more than half of the total delegate votes. Moreover, the first elections are of greater importance as they reveal voters' preferences and influence later elections through their results. These two features of the electoral process lead to a long-known and unresolved problem of front-loading (Mayer & Busch, 2003; Ridout & Rottinghaus, 2008).

From a state's perspective, an early election date can increase its influence in the primaries. If all states are considering moving their elections to earlier dates, a solution must be found to regulate the timing of state elections that takes into account the different importance of the dates. Currently, additional delegates are granted for late election dates, but these do not provide sufficient incentive for states to resolve front-loading. In the following analysis, we compare the US Democrats' 2020 election schedule with two proposed election schedules, namely randomising the election dates and arranging the states according to their delegate count. The aim of the analysis is to find out whether this leads to a more balanced distribution of the importance of elections for the individual states that is less driven by the timing of the elections.

4.2 Setup

We utilise the actual schedule and reward structure of the 2020 democratic party presidential primary elections. The ordering of the elections and the number of delegates rewarded by the election are displayed in Table 2 in Appendix B. For ease of exposition, we simplify the model of the election process by discarding unpledged delegates, implementing only winner-takes-it-all elections, and engaging only two candidates $i \in \{0, 1\}$.

The reward function of the contest is given by winning the primaries, i. e. obtaining the majority of the delegates' votes. We model the state's election as a representative voter facing a binary choice model with random utility. The utility function (1) is composed of four components: a) the fixed reputation $\{\eta_0, \eta_1\} = \{0.5, 0\}$ of the candidates, b) the match between state preferences $\rho_s \sim \mathcal{N}(0, 1)$ and the candidates' positions $\{\rho_1, \rho_2\} = \{-1, 1\}$, c) spillover effects $\zeta_{i,s}$ of previous elections. d) a standard Type I extreme value error term $\epsilon_{i,s}$. Under the assumption, that the representative voter always chooses the candidate with maximum utility, the setting describes a conditional logit model (McFadden, 1974) with outcome probability $\pi_{i,s}$ that candidate i wins the election in state s as described by Equation (2).

Spillover effects from the results of early states on future elections occur if the share of obtained delegates' votes in prior states differs from the expected share solely based on the candidates' reputation. The spillover effect as defined in (3) is an example of dynamic covariates in the outcome model that have to be re-evaluated when new election results are determined.

$$\psi_{i,s} = \eta_i + \zeta_{i,s} - \frac{|\rho_i - \rho_s|^2}{2} + \epsilon_{i,s} \tag{1}$$

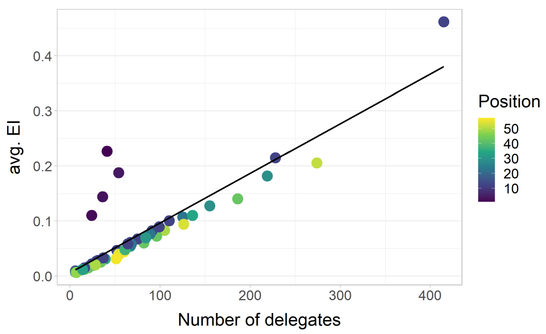
$$\pi_{i,s} = \frac{\exp(\psi_{i,s})}{\exp(\psi_{i,s}) + \exp(\psi_{1-i,s})} \tag{2}$$

$$\zeta_{i,s} = \frac{\sum_{s'<s} y_{i,s'}}{\sum_{s'<s}(y_{i,s'} + y_{1-i,s'})} - \frac{\exp(\eta_i)}{\exp(\eta_i) + \exp(\eta_{1-i})} \tag{3}$$

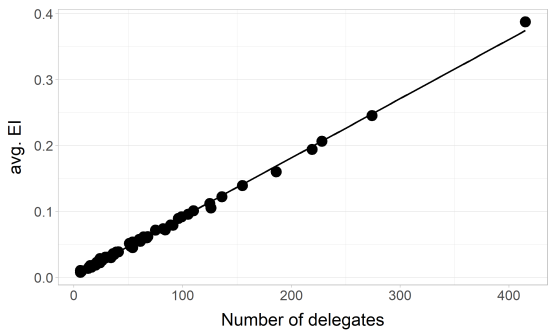
Based on the model for the election process given by equations (1)-(3), the probability for each candidate to win the primaries conditional on the outcome of a single state's result is determined. Because the number of states and territories is too large to allow an exact numerical calculation, the winning probabilities are approximated by a Monte Carlo simulation using 5'000 simulation runs. As suggested in Section 3.2.5, we choose the difference in the winning probabilities as the distance function for this contest with binary reward structure, i. e. to be nominated as a presidential candidate or not. This distance function leads to symmetric EI estimates for both candidates. To eliminate the dependency on a particular set of states' preferences, we simulate 1'000 random draws of state preferences.

The three schedules we evaluate are defined as follows: The *regular* schedule is according to the actual election dates[4] and the allocated number of delegates by state. In the *random* schedule, we randomly permute the ordering of the states keeping the framework of the schedule fixed. The *rank increase* scheme ranks the states increasing by their number of delegate votes and applies the ordering to the actual schedule framework.

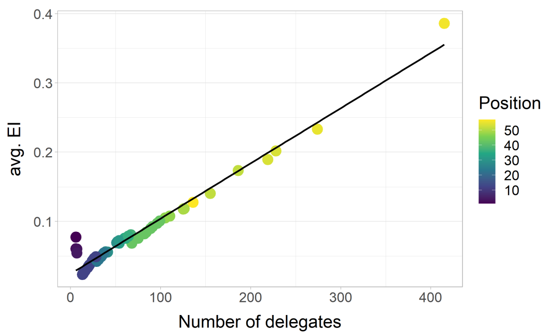
4.3 Results



(a) Regular



(b) Random



(c) Rank increase

Fig. 1. Average event importance estimates over 1'000 states' (preferences) samples with linear fit.

[4]Because of the Covid-19 pandemic, several states have postponed their election such that the actual dates can differ from the initially planned schedule.

Figure 1 shows the average EI estimates over all samples for the three schedule types. The regular schedule of the 2020 democratic party primaries (1a) displays the increased importance of the early elections, as the respective states have a higher average EI estimate than the number of delegates allocated to them would suggest. The randomised schedule (1b) reveals a linear relationship between the ability of a state to change the outcome of the primaries and its number of delegates. Since all states will eventually benefit from an early position in the schedule, the positive spillover effects are spread across all states and territories and balance each other out.

Ordering the states by their number of delegates (1c) cannot entirely eliminate the first-winner effect but substantially alleviates it. The increased importance of states due to the early election date can be compensated by a smaller number of delegates, which is an option already considered in the allocation of delegates. Because an ex-ante randomisation requires many repetitions to balance the positional effects, the ordered schedule seems to be a good compromise between practicability and fairness.

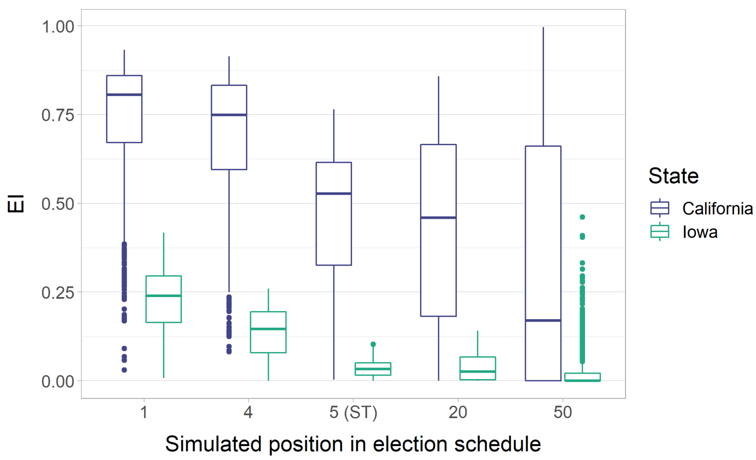


Fig. 2. Estimated event importance values for Iowa and California at different hypothetical positions in the election schedule for 400 random samples of states' preferences.

To illustrate the importance of both the first-winner effect and the state's size we show in Figure 2 the estimated EI values for two states, Iowa (41 delegates) and California (415 delegates), at different hypothetical positions in the election schedule. For the very early positions (1 and 4), both states are of considerable importance for the final nomination of the presidential candidates. From position 5, the "Super Tuesday" on which numerous states hold their elections, the capability of the elections to influence the final nomination decreases considerably.

The importance of elections in small states thrives on the fact that there are spillover effects through the first-winner effect. Because of the substantial amount of delegates, California remains considerably important for the nomination result in late stages of the schedule, while the low number of Iowa's delegates become irrelevant in many realisations. For Iowa in particular, if the election would be held in the middle (20) or at the end (50) of the schedule, the importance of the election would be determined only by the possibility that the state's delegates could act as tiebreakers in the nomination if the election race is close.

5 Application: European football leagues

5.1 Introduction

As with many analyses of contests, sports data provide a suitable and well-structured framework for applications because it features accurate observational data (Kahn, 2000; Bar-Eli, Krumer, & Morgulev, 2020). We apply the EI framework to the seven major European football leagues. Those contests have a non-trivial schedule of multiple events, held sequentially or in parallel, between pairings of the participants and a non-linear reward structure that can vary individually or change throughout the season – all of which can be handled naturally with the proposed framework. With, for example, postponed games leading to changes in the schedule, or supplementary rewards achieved by national cup tournaments changing the reward structure individually, this application is a good showcase to demonstrate the flexibility of the framework.

Quantifying EI in this context is interesting for several reasons. Contest designers should avoid match-ups that pit contestants with unequal incentive levels against each other. Such matches are potentially more susceptible to bribery and a lower engagement of certain participants could give an unfair (dis)advantage to participants not even involved in the event itself. Other valuable use cases of the EI in football tournaments include (a) selecting intense or interesting matches for prime-time broadcast, (b) improving the prediction of winning probabilities, and (c) detecting or avoiding unfair match schedules.[5]

5.2 Setup

5.2.1 Data

We analyse data from the 2006/07 through 2018/19 seasons of seven major European football leagues, namely the German 'Bundesliga 1', the Dutch 'Eredivisie', the Spanish 'La Liga', the French 'Ligue 1', the English 'Premier League', the Portuguese 'Primeira Liga', and the Italian 'Serie A'. These leagues were the major leagues in Europe in terms of sporting and financial success throughout the studied period. All leagues are designed as double round-robin tournaments, which means that each team plays each other twice - once at each home venue. The rewards are distributed after the season. With the seven analysed leagues, we cover a variety of different reward structures. A detailed description of the tournament design, league format, and reward structure of the considered European football leagues can be found in Appendix C.1.

For each individual match, we record a long list of characteristics: describing the match setting, such as the time or day of the week, characterising the participating teams, as their success in recent matches, whether they play in international competitions and metrics of the squad players, e. g. age, height, estimated market value, and preferred foot. The full set of all 133 covariates is listed in Appendix C.2.

[5] E. g. pairings with unequal EI levels may result in unfair schedules for involved or non-involved teams.

5.2.2 Specific application framework

In this section, we describe how we implement the general framework from Section 3 and elaborate on all generic functions outlined in Algorithm 1. Algorithm 2 in Appendix C.3 presents the pseudo-code of the specific framework tailored to this application. At the end of a football season, rewards are allocated to the teams based on their final rank. The areas in the ranking which denote the championship title, qualification for international competitions, and relegation are stated by strict thresholds. We use those boundaries to group all the ranks between two thresholds as a single reward.[6] More detailed information on the reward structures per league and season appear in Appendix C.3.2. Individual updating patterns of the reward scheme, e. g. because the UEFA Europa League place allocated to the national cup winner is transferred and included in the league's season rewards, are explained in Appendix C.3.3.

In Section 3, we have outlined how the general framework can be employed for applications with unknown outcome functions and those incorporating the EI itself. Outcomes of football matches do not follow a deterministic rule and can only approximately be described by a probabilistic model. We follow the approach of Goller, Knaus, Lechner, and Okasa (2021), using an ordered choice model with three outcome probabilities estimated by the Ordered Forest (Lechner & Okasa, 2019), hereafter abbreviated as *ORF*.[7] To restrict the number of covariates in the ORF model we perform a LASSO-based model selection step. Starting from the second iteration, this set of covariates is extended with the previously estimated EI values (as

[6]Financial rewards, i. e. money from broadcasting rights, are determined by the final league table. However, this gradation is less relevant for the team, the coaches, and the players. Becoming champions, qualifying for next year's European Cup or not being relegated to a lower league is what we assume is more in the focus of the involved entities. Strict thresholds implicitly assume that it's not particularly relevant for teams to finish 10^{th} or 11^{th}, but that the potential implications around crucial positions in the ranking – which guarantee participation in next year's Champions League, for example – are considerably greater. For a discussion on the financial dimensions consult Goller and Krumer (2020). Although the framework allows for a weighting scheme, for the sake of simplicity, we consider all thresholds to be equally relevant to each team.

[7]The general framework is not restricted to this specific method and the choice of the underlying outcome model is of second-order (see Appendix C.4.4).

outlined in Section 3.2.4). In addition, we also simulate the exact score of the match, drawn from two independent Poisson distributions, as the goal difference often serves as a tie-breaker in determining the final ranking.

The choice of the Jensen-Shannon divergence as the distance function, specified in equation (4), follows the argumentation in Section 3.2.5 for settings with multiple event outcomes and rewards. We use a scaling factor of $\ln(3)^{-1}$ to constrain the EI to the [0,1] interval and weight the probability distributions P_i by the match outcome probabilities $\{\pi_H, \pi_D, \pi_A\}$ to account for the likelihood of the three outcome scenarios.

$$\mathrm{JSD}_{\pi_H,\pi_D,\pi_A}(P_H, P_D, P_A)) = \frac{1}{\ln(3)}\left(\mathbf{H}\left(\sum_{i\in\{H,D,A\}} \pi_i P_i\right) - \sum_{i\in\{H,D,A\}} \pi_i \mathbf{H}(P_i)\right) \quad (4)$$

$$\textit{where } \mathbf{H}(P) = -\sum_{j=1}^{m} P(x_j)\ln(x_j)$$

5.3 Results

5.3.1 Distribution of the estimated values

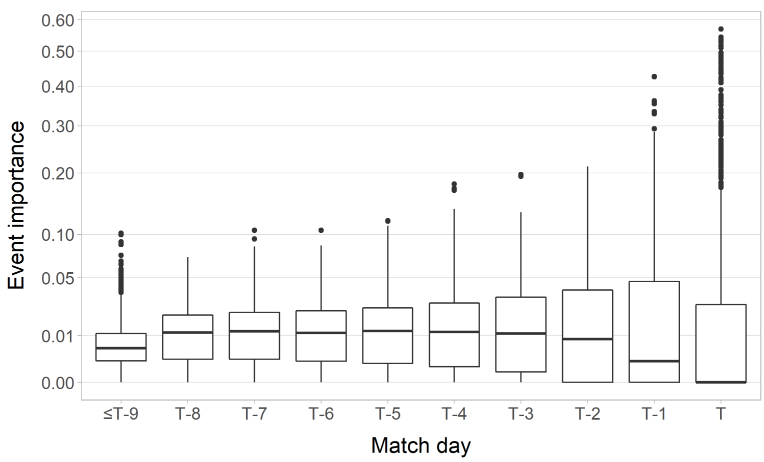


Fig. 3. Home and away team's event importance estimates grouped by match day enumerated in relation to the last match day, all seasons, and all leagues. Square-root transformation to y-axis applied

Figure 3 shows the distribution of estimated EI values by match days. For the majority of the season, the estimated EI values are concentrated around a value of about 0.01. In other words, most matches are similarly (un)important for the first parts of the season. Deviations can be observed in pairings between teams that are expected to be close in the final end-of-season standings, as in these matches a positive result implies a negative result for the opponent. This behaviour changes towards the end of the season, with non-relevant matches and frequent outliers of particularly important matches. The uncertainty about the outcome of the rest of the season is reduced with fewer unknown future results, and the results of individual matches can become more decisive for the end-of-season rewards. This results in more pronounced values of the EI towards the end of the season.

As an illustrative example, we show the estimated EI values for the last match day in the 2017/18 German Bundesliga 1 season in Appendix C.4.1.

5.3.2 Predicting match outcomes

To shed light on whether the quantified EI has an impact on outcome prediction, we compare the estimates of a 'baseline' ordered forest model that does not use the EI information with a 'richer' ORF model that includes the estimated EI of both, the home and away team, as additional input.

We fit both ORF models on half of the data and predict the outcome probabilities with the two models on the other half. Based on the outcome probabilities we construct the expected points (ExpP) measure by awarding points to the outcomes according to modern football rules – 3 points for a win, 1 for a draw, and 0 for a loss. This procedure is repeated with swapped training and prediction samples.

Figure 4 displays the difference in expected points between the rich and the baseline model by the difference in the EI values between the two competing teams. The generalised additive model (GAM) fit on the data confirms that teams with a higher absolute difference in EI are attributed a higher absolute prediction of ExpP with the richer model verifying that the inclusion of the EI variable is relevant for outcome

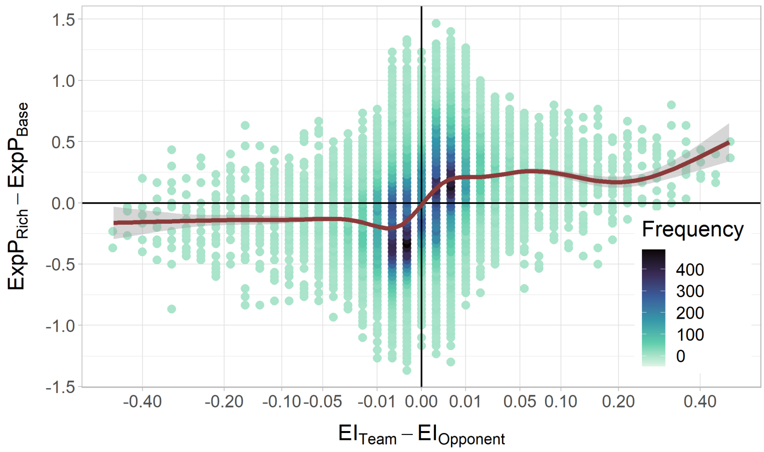


Fig. 4. The difference in expected points (ExpP) between the model including EI variables (Rich) and the baseline model (Base) by the difference in event importance (EI) between the team and its opponent. Values are rounded to the nearest grid point. Frequency indicates the number of points on a grid point. The red line denotes a GAM with a 95% confidence interval. Expected points are averaged over 100 estimates with different sample splits. Square-root transformation to x-axis applied.

prediction. The variable importance measures of the EI variables in the rich model are shown in Appendix C.4.3 and provide evidence for the notable contribution of the EI in the outcome model.

5.3.3 Prediction power improvement

In Section 5.3.2 we have shown, that the estimated EI values are picked up by an enriched outcome model. This raises the question of whether using EI values in an outcome model improves predictive performance.

We compare seven different prediction models to margin-free betting odds of the online betting platform B365.[8] The baseline ORF model as described in Section 5.3.2 (*ORF*), the richer model including the EI values (*ORF+EI*), and additionally including the difference of EI estimates (*ORF+EI+diff*), an ORF model with a binary importance measure[9] (*ORF+naive*), an ORF model (*ORF+add3*) that adds three

[8]The betting odds serve as a benchmark and are collected from the website `www.football-data.co.uk`. To ensure comparability with the model predictions, we linearly scale the odds to remove the bookmaker's margin.

[9]A binary variable indicating if the match is still relevant for attaining a better/worse reward or if the match cannot change the reward anymore. Several empirical works use such indicators (Fornwagner, 2019; Feddersen, Humphreys, & Soebbing, 2021)

covariates,[10] an ordered logit model with the baseline variables (*Logit*), and an ordered logit including the EI estimates (*Logit+EI*).

To evaluate the out-of-sample prediction accuracy we randomly split the data into two samples.[11] On one-half of the seasons, the models are fitted, on the other half, the prediction accuracy of the models is measured by the log-likelihood and the Brier score (results for the Brier score can be found in Appendix C.4.6). In each repetition, we index the accuracy measures by the results of the benchmark betting-odds model to balance any particular characteristics of the chosen sample.

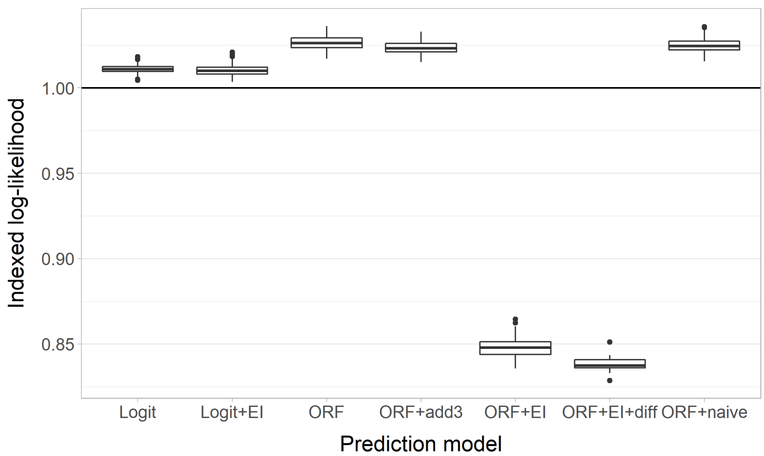


Fig. 5. Out-of-sample prediction accuracy of different models over 1000 repetitions. Values are quantified in log-Likelihood and indexed in each repetition by the performance of betting odds.

Figure 5 shows the out-of-sample prediction accuracy. For the logit model, the addition of the EI results in only a slight improvement, which is probably due to the linearity constraint. Including EI values in the ORF model substantially increases the predictive power, as the additional information contained in the EI variables can be fully utilised, resulting in better performance than the margin-free betting odds. Recording the event importance in a binary variable does not improve the accuracy of the prediction. The model with three added covariates indicates that the increase

[10]To investigate if a potential improvement is just induced mechanically by the larger set of covariates. We include travel distance, days since the last match of the home team & days to the next match of the away team.

[11]The split is performed on the full-season level to not give the proposed models an unfair advantage over the betting odds.

in prediction power is not just induced by the larger set of covariates.

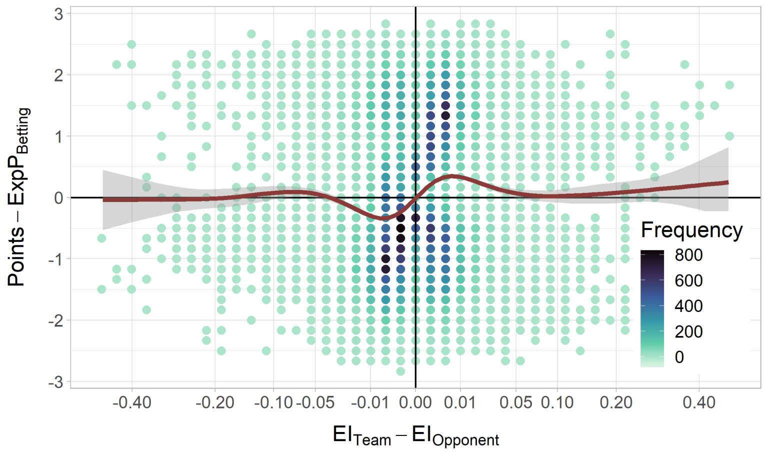


Fig. 6. The difference in realised and expected points according to betting odds by the difference in event importance between the team and its opponent. Values are rounded to the nearest grid point. Frequency indicates the number of points on a grid point. The red line denotes GAM with a 95% confidence interval. Square-root transformation to x-axis applied.

To break down the improvement by the EI information on the betting odds, we present in Figure 6 the difference between the achieved points and the expected points according to the betting odds in relation to the difference in the EI values between the teams and their opponents. A GAM fit on all data points indicates, that in particular across matches where the differences are small, the EI can partly explain the mismatch in the betting odds. For larger EI differences, the EI does not provide additional information to the bookmaker's model. We deduce that the betting odds already cover the unequal incentives when they are particularly pronounced but do not fully account for more subtle disparities in the importance of a match to the competing teams. This is generally in line with and extends the results of Feddersen et al. (2021), which show that bookmakers are aware of the impact of different incentives on the outcome of matches on the final match days. We provide additional evidence on the complimentary informational content of the EI measure to the betting odds in Appendix C.4.5.

5.3.4 Team performance

Besides the usefulness of the EI measure in predictions, we investigate whether the differences in incentives for teams is reflected in-match statistics that record a team's on-field behaviour and performance. For this, we investigate in-match statistics (data source: Opta[12]) for the 2010/11 through 2018/19 German Bundesliga 1 seasons with regards to our EI estimates. The team performance data is collected individually for both teams and pooled for the home and away teams. Outcome variables are totals per match, except for 'Duel win' and 'Tackles win' which are shares. For ease of interpretation, the EI estimates for the home and away teams are each divided into three groups - 'zero' (EI = 0), 'low', and 'high' (above 0.035) EI.[13]

Figure 7 shows the results for four outcomes: duels per game, number of completed passes, number of goals scored, and number of goals conceded.[14] Complementary results using other outcomes are shown in Appendix C.4.7. The results can be summarised as follows: Teams for which a match is particularly important (i) play more aggressive entering more duels on the pitch, (ii) play more directly towards goal with fewer passes, fewer touches, and more entries into the final third and penalty area, (iii) score more goals. In contrast, teams with zero importance exhibit a more passive style of play and concede more goals.

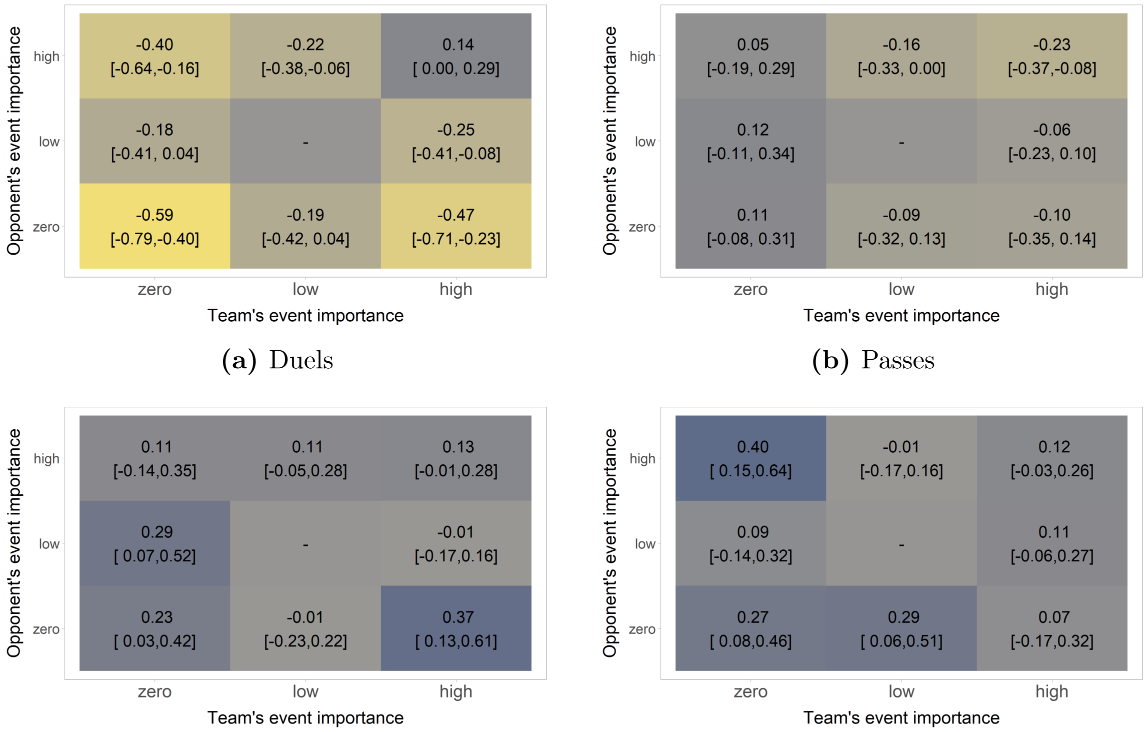
5.3.5 Public perception

Sport is entertainment and thrives on public perception. If the calculated EI can represent the (later realised) public interest in a specific match, it could be useful for several purposes – marketing, ticket pricing, or prime-time broadcasting selection. With this in mind, we relate the EI to the stadium attendance turnout, as well as

[12] `https://www.statsperform.com/opta/`

[13] 5% of the EI values are zero. We therefore choose the *high EI* threshold at the 95 % quantile of the EI values to obtain balanced groups.

[14] In a first step, we run a fixed effect regression for every outcome individually using the combinations 'Team x home/away x season', as well as 'Opponent x home/away x season' fixed effects'. The resulting residuals are centred and standardized by 'Team x home/away x season'. On those scaled residuals we run a regression using again the grid on the EI categories 'zero', 'low', and 'high' for both competing teams.



(c) Goals scored **(d)** Goals conceded

Fig. 7. Linear regression estimates of the centred and standardized residuals of different outcomes on the Event importance categories. 95% confidence intervals are in parentheses. The baseline is low by low category.

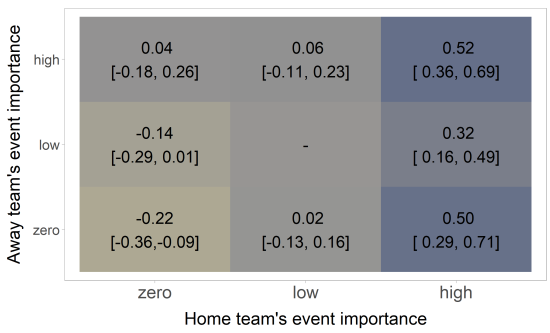
social media attention. While attendances are officially reported by clubs, social media attention is captured through club mentions and match hashtags within the 24 hours before kickoff on Twitter.[15]

As different clubs have put different emphasis on social media and this has changed over time, we control for the team and season-specific usage of social media.[16] In

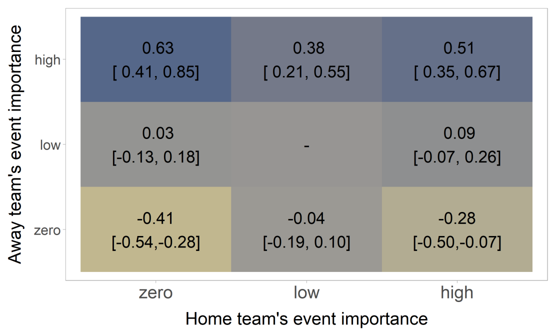
[15] Gathered using Twitter API v2. The analysis yields almost identical results when the time period is extended to 48h before the match. Match hashtags are a short abbreviation of the names of the two teams which are regularly used on Twitter, i. e. #BVBS04 relates to the match of the teams Borussia Dortmund against Schalke 04. Mentions of club accounts are Twitter tweets containing the account name of a team, e. g. @LFC - the official Twitter account of the English team Liverpool FC. Due to the inconsistency and lack of use of the aforementioned proxies in the early years and across the leagues, we can only perform this analysis beginning with the 2014/15 season and must exclude the Spanish and Portuguese leagues.

[16] The procedure of the analysis is similar to the residual analysis in Section 5.3.4. In a first step we calculate a linear fixed-effect model containing a FE for every 'team x home/away x season' interaction and for every calendar month and using the Twitter or attendance data, both in logs, as the outcome. In a second step, we centre and standardize the residuals from the linear fixed-effect models by 'league x season' pairs. This is the most natural procedure in our view. By taking nominal outcomes, other FE variants, or standardising over different groups interpretation of results does not change.

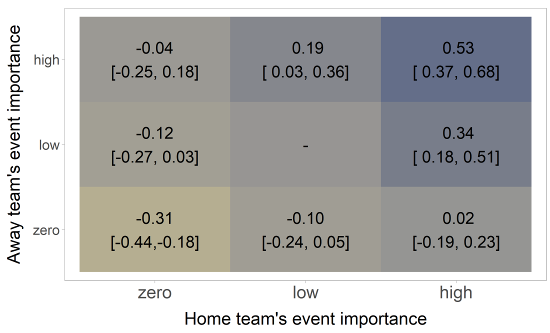
the analysis, we explain the standardised and centred attention measures by linear regression on the EI of the two competing teams. Figure 8 presents the point estimates and 95 % confidence intervals of the linear regressions. Stadium attendance is modestly associated with the home team's importance in the match. Here, restrictions on stadium capacity and (pre-sold) season tickets could mitigate the effect. Thus, social media attention might give a more clear picture of realised interest. We find team account mentions are strongly associated with the respective team's EI measure. Similarly, the match-tag mentions increase with both teams' EI. This is consistent with and complementary to Dobson and Goddard (1992) and Lei and Humphreys (2013) reporting higher stadium attendance for more important sporting events, and recent findings by Buraimo et al. (2022) that Premier League television audiences are larger for more important matches.



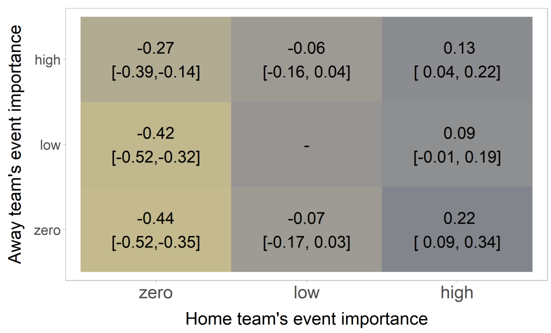
(a) Home team's Twitter account mentions



(b) Away team's Twitter account mentions



(c) Match-tag Twitter mentions



(d) Stadium attendance

Fig. 8. Linear regression estimates of the centred and standardized residuals of different outcomes on the event importance (EI) categories. 95% confidence intervals are in parentheses. The baseline is low by low category.

6 Conclusion

Public perception and academic research analyse incentives in simple situations where there is a direct link between performance and reward. More complex situations with indirect rewards and therefore unclear implicit incentive structures have received little attention.

In this article, we propose a statistical method to quantify the importance of single events in multi-event contests with end-of-contest reward structures. Thanks to its flexibility and generality the procedure covers a multitude of potential applications and can be valuable for various fields, including sales and marketing, human resources, or operations management. Our event importance framework can be adapted to different contest structures seen in society and opens a variety of potential research topics such as behavioural responses involving implicit incentives or operational fairness concerns in contest and reward structures. These include, for example, different valuations due to the order of actions or asymmetric incentives that lead to distorted probabilities of winning in a contest.

In an application to European football leagues, we show the association of the quantified importance of a match to in-match behaviour and the performance of the teams. As discrepancies in the EI can lead to altered outcome probabilities, the quantification of the event importance can help to ensure fair tournaments. The event importance measure also addresses other stakeholders in the football industry. As we show that the EI measure is consistent with the public interest in terms of social media and stadium attendance it can be useful for dynamic ticket pricing or TV stations that want to broadcast the most attractive match. Lastly, we illustrate the value of the EI measure for predicting match outcomes and point out under which circumstances the bookmakers do not yet account for the event's importance.

For the application to the US presidential primaries, we quantify the higher relevance of early election dates induced by the first-winner effect. For small states with a low number of delegates, this can substantially boost their influence on the nomination outcome as otherwise, their votes become irrelevant in many of the primaries. We show that the two investigated hypothetical schedules lead to more equitable distribution in the ratio of event importance values to the number of delegates rewarded by the election.

References

Alchian, A., & Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *The American Economic Review*, *62*(5), 777–795.

Apesteguia, J., & Palacios-Huerta, I. (2010). Psychological pressure in competitive environments: Evidence from a randomized natural experiment. *American Economic Review*, *100*(5), 2548–64.

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *Review of Economic Studies*, *76*(2), 451–469.

Bar-Eli, M., Krumer, A., & Morgulev, E. (2020). Ask not what economics can do for sports-ask what sports can do for economics. *Journal of Behavioral and Experimental Economics*, *89*, 101597.

Borup, D., Christensen, B. J., Mühlbach, N. S., & Nielsen, M. S. (2022). Targeting predictors in random forest regression. *International Journal of Forecasting*.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Buraimo, B., Forrest, D., McHale, I. G., & Tena, J. d. D. (2022). Armchair fans: Modelling audience size for televised football matches. *European Journal of Operational Research*, *298*(2), 644–655.

Cohen-Zada, D., Krumer, A., Rosenboim, M., & Shapir, O. M. (2017). Choking under pressure and gender: Evidence from professional tennis. *Journal of Economic Psychology*, *61*, 176–190.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, *108*(17), 6889–6892.

de Bruin, W. B. (2005). Save the last dance for me: Unwanted serial position effects in jury evaluations. *Acta psychologica*, *118*(3), 245–260.

Dobson, S. M., & Goddard, J. A. (1992). The demand for standing and seated viewing accommodation in the English Football League. *Applied Economics*, *24*(10), 1155–1163.

Ehrenberg, R., & Bognanno, M. (1990). Do tournaments have incentive effects. *Journal of Political Economy*, *98*(6), 1307–1324.

Feddersen, A., Humphreys, B. R., & Soebbing, B. P. (2021). Contest incentives, team effort, and betting market outcomes in European football. *European Sport Management Quarterly*(forthcom.), 1–17.

Fornwagner, H. (2019). Incentives to lose revisited: The NHL and its tournament incentives. *Journal of Economic Psychology*, *75*, 1–12.

Geenens, G. (2014). On the decisiveness of a game in a tournament. *European Journal of Operational Research*, *232*(1), 156–168.

Ginsburgh, V., & van Ours, J. (2003). Expert Opinion and Compensation : Evidence from a Musical Competition. *The American Economic Review*, *93*(1), 289–296.

Goller, D. (2022). Analysing a built-in advantage in asymmetric darts contests using causal machine learning. *Annals of Operations Research*, *forthcom.*, 1–31. doi: 10.1007/s10479-022-04563-0

Goller, D., & Heiniger, S. (2022). *Replication code and results for: 'A general framework to quantify the event importance in multi-event contests'.* Harvard Dataverse. doi: https://doi.org/10.7910/DVN/F3QA9N

Goller, D., Knaus, M. C., Lechner, M., & Okasa, G. (2021). Predicting Match Outcomes in Football by an Ordered Forest Estimator. In R. H. Koning & S. Kesenne (Eds.), *A modern guide to sports economics* (pp. 335–355). Edward Elgar Publishing.

Goller, D., & Krumer, A. (2020). Let's meet as usual: Do games played on non-frequent days differ? evidence from top european soccer leagues. *European Journal of Operational Research*, *286*(2), 740-754.

Harris, C., & Vickers, J. (1987). Racing with uncertainty. *The Review of Economic Studies*, *54*(1), 1–21.

Jennett, N. (1984). Attendances, Uncertainty of Outcome and Policy in Scottish League Football. *Scottish Journal of Political Economy*, *31*(1), 176–198.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of economic perspectives*, *14*(3), 75–94.

Klumpp, T., & Polborn, M. K. (2006). Primaries and the new hampshire effect. *Journal of Public Economics*, *90*(6-7), 1073–1114.

Knight, B., & Schiff, N. (2010). Momentum and social learning in presidential primaries. *Journal of political economy*, *118*(6), 1110–1150.

Laffont, J.-J., & Martimort, D. (2002). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Lahvička, J. (2015). Using Monte Carlo Simulation to Calculate Match Importance: The Case of English Premier League. *Journal of Sports Economics*, *16*(4), 390–409.

Lazear, E. (2000). The Power of Incentives. *The American Economic Review, Papers and Proceedings*, *90*(2), 410–414.

Lechner, M., & Okasa, G. (2019). Random Forest Estimation of the Ordered Choice Model. *arXiv:1907.02436*.

Lei, X., & Humphreys, B. R. (2013). Game importance as a dimension of uncertainty of outcome. *Journal of Quantitative Analysis in Sports*, *9*(1), 25–36.

Lin, J. (1991). Divergence measures based on the shannon entropy. *IEEE Transactions on Information theory*, *37*(1), 145–151.

Mayer, W. G., & Busch, A. E. (2003). *The front-loading problem in presidential nominations*. Brookings Institution Press.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York: Academic press.

Nielsen, F. (2020). On a generalization of the jensen–shannon divergence and the jensen–shannon centroid. *Entropy*, *22*(2).

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, *37*(1), 7–63.

Preston, I., & Szymanski, S. (2003). Cheating in contests. *Oxford review of economic policy*, *19*(4), 612–624.

Ridout, T. N., & Rottinghaus, B. (2008). The importance of being early: Presidential primary front-loading and the impact of the proposed western regional primary. *PS: Political Science & Politics*, *41*(1), 123–128.

Rosen, S. (1985). Prizes and Incentives in Elimination Tournaments. *National Bureau of Economic Research Working Paper Series*(1668).

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

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

A Details to Algorithm 1

The Algorithm 1 works as follows: Line 2 loops over all possible outcomes $\mathcal{Y}_{e_{t,i}}$ for the particular event $e_{t,i}$. This loop calculates the reward distribution at the end of the contest once for each possible outcome. Note that all variables in this loop are depending on the value of the loop iterator $\boldsymbol{y}_{e_{t,i}}$. To improve readability, we denote this by the superscript * instead of the more intuitive subscript $\boldsymbol{y}_{e_{t,i}}$ which prevents an additional subscript level. All variables are subject to a probability distribution which is formed by the continuous evaluations of the outcome model out().

At the beginning of each iteration of the loop, the outcome for event $e^*_{t,j}$ is set to the iterator outcome $\boldsymbol{y}_{e_{t,i}}$. In case there are other simultaneous events $e_{t,j}$ at time t, the loop in line 4 evaluates their outcome probabilities according to the outcome model $\mathrm{out}(x_{e_{t,j}})$. Next, the outcome-dependent elements of the future schedule and covariates are generated according to the outcome probabilities of y^*_{t-}, as well as $\mathcal{T}_{t-}$ and x_{t-}.

Line 7 loops over the future times t' until the end of the contest $\mathcal{T}_{t^+}{}^*$. Note that this set can be adapted or extended at run time by the gen() steps in case of an outcome-dependent schedule. In each iteration step, first, the outcome probabilities $y^*_{e_{t'}}$ for all events at this particular time t' are determined according to the outcome model, and then the contest schedule and the covariates are generated, in case they depend on the past outcomes. When the end of the contest $\mathcal{T}^*$ has been reached, line 10 determines the reward distribution $r_{k,\boldsymbol{y}_{e_{t,i}}}$ for contestant k based on the outcomes of all single events in the contest.

The final step, after the outer loop over all possible outcomes has finished, calculates the distance between the reward distributions which quantifies the event importance $\mathrm{EI}_{e_{t,i},k}$.

B 2020 Democratic Party presidential primaries

Table 2. Data and results summary on 2020 Democratic Party presidential primaries.

			avg. EI		
State/Territory	Election Date	Delegates	Actual	Rank increase	Random
Alabama	March 3, 2020	52	0.046	0.071	0.048
Alaska	April 10, 2020	15	0.012	0.027	0.018
American Samoa	March 3, 2020	6	0.009	0.060	0.010
Arizona	March 17, 2020	67	0.054	0.081	0.060
Arkansas	March 3, 2020	31	0.028	0.045	0.031
California	March 3, 2020	415	0.462	0.386	0.388
Colorado	March 3, 2020	67	0.061	0.079	0.060
Connecticut	August 11, 2020	60	0.044	0.076	0.058
Delaware	July 7, 2020	21	0.017	0.037	0.023
Democrats Abroad	March 10, 2020	13	0.011	0.023	0.014
District of Columbia	June 2, 2020	20	0.015	0.036	0.019
Florida	March 17, 2020	219	0.182	0.189	0.194
Georgia	June 9, 2020	105	0.083	0.105	0.096
Guam	June 6, 2020	7	0.006	0.060	0.010
Hawaii	May 22, 2020	24	0.019	0.043	0.029
Idaho	March 10, 2020	20	0.017	0.036	0.020
Illinois	March 17, 2020	155	0.127	0.140	0.139
Indiana	June 2, 2020	82	0.060	0.082	0.074
Iowa	February 3, 2020	41	0.227	0.056	0.039
Kansas	May 2, 2020	39	0.031	0.056	0.039
Kentucky	June 23, 2020	54	0.040	0.070	0.054
Louisiana	July 11, 2020	54	0.038	0.069	0.045
Maine	March 3, 2020	24	0.022	0.042	0.022
Maryland	June 2, 2020	96	0.072	0.097	0.089
Massachusetts	March 3, 2020	91	0.082	0.092	0.079
Michigan	March 10, 2020	125	0.107	0.118	0.112
Minnesota	March 3, 2020	75	0.067	0.076	0.072
Mississippi	March 10, 2020	36	0.030	0.053	0.035
Missouri	March 10, 2020	68	0.057	0.069	0.062
Montana	June 2, 2020	19	0.014	0.033	0.020
Nebraska	May 12, 2020	29	0.023	0.042	0.031
Nevada	February 22, 2020	36	0.144	0.053	0.036
New Hampshire	February 11, 2020	24	0.110	0.043	0.025
New Jersey	July 7, 2020	126	0.094	0.119	0.105
New Mexico	June 2, 2020	34	0.025	0.049	0.030
New York	June 23, 2020	274	0.206	0.233	0.246
North Carolina	March 3, 2020	110	0.100	0.108	0.101
North Dakota	March 10, 2020	14	0.012	0.025	0.016
Northern Marianas	March 14, 2020	6	0.007	0.077	0.008
Ohio	April 28, 2020	136	0.110	0.128	0.123
Oklahoma	March 3, 2020	37	0.033	0.055	0.035
Oregon	May 19, 2020	61	0.048	0.075	0.055
Pennsylvania	June 2, 2020	186	0.140	0.173	0.160
Puerto Rico	July 12, 2020	51	0.031	0.069	0.052
Rhode Island	June 2, 2020	26	0.019	0.047	0.026
South Carolina	February 29, 2020	54	0.188	0.072	0.053
South Dakota	June 2, 2020	16	0.012	0.029	0.017
Tennessee	March 3, 2020	64	0.058	0.079	0.062
Texas	March 3, 2020	228	0.215	0.202	0.207
Utah	March 3, 2020	29	0.027	0.044	0.030
Vermont	March 3, 2020	16	0.015	0.028	0.015
Virgin Islands	June 6, 2020	7	0.006	0.054	0.010
Virginia	March 3, 2020	99	0.089	0.100	0.092
Washington	March 10, 2020	89	0.076	0.090	0.080
West Virginia	June 9, 2020	28	0.020	0.049	0.029
Wisconsin	April 7, 2020	84	0.069	0.085	0.072
Wyoming	April 17, 2020	14	0.012	0.024	0.016

C European football leagues

C.1 League structures

We focus on the seven major European football leagues: The Dutch 'Eredivisie', the English 'Premier League', the French 'Ligue 1', the German 'Bundesliga 1', the Italian 'Serie A', the Portuguese 'Primeira Liga', and the Spanish 'Primera Division'. These leagues are leaders in terms of success in international competitions, financial capabilities, competitiveness and stadium attendance.

The investigated European football leagues are organised in a double round-robin tournament, in which clubs face each other twice during the season, once at each home ground. Thus, a season consists of $2(n-1)$ match days, where n is the number of competing teams in the league - that is 16, 18, or 20. Table 3 summarises the number of teams by the league, the number of observed distinct teams, and the total number of matches in the covered time period.

Table 3. Data summary on leagues.

League	Teams per season	Total matches	Observed teams
Bundesliga 1	18	3'060	28
Eredivisie	18	3'060	26
La Liga	20	3'800	35
Ligue 1	20	3'800	37
Premier League	20	3'800	36
Primeira Liga	16 or 18	2'730	30
Serie A	20	3'800	34

To level out the encounters over the whole season, the season is split into two halves in which each pairing is played once. With the exception of the English Premier League, all of the covered leagues feature a mid-season break around December/January – a season usually starts in August/September and finishes in May. On a match day, or round, $\frac{n}{2}$ matches are played with every team competing in one pairing.

Winning a match is rewarded with 3 points, draws are valued with 1 point, and losses with 0 points.. At the end of the season, when all matches have been played, the points of all match days are added up and the teams are ranked in descending order in the total sum of points. In the event of a tie in the ranking, the number of goals scored/achieved or the result of the match pairings between the teams concerned shall be decisive, depending on the league-specific tie-breaker rules.

C.2 List of covariates

Table 4 lists all covariates used in the dataset and whether they have been selected by the model selection procedure.

Table 4. List of all covariates in the data.

	Description	Type	Selected
One variable per game			
	Season	numeric	-
	Bundesliga	binary	-
	Eredivisie	binary	-
	La Liga	binary	-
	Ligue 1	binary	-
	Premier League	binary	yes
	Primeira Liga	binary	-
	Serie A	binary	-
	Kick-off [hour]	numeric	-
	Day of the week	numeric	yes
	Weekend match (Friday to Sunday)	binary	-
	Match during a public holiday	binary	-
	Match during Christmas holidays	binary	-
	Match before an international competition break	binary	-
	Match after an international competition break	binary	-
	Match before a national team break	binary	-
	Match after a national team break	binary	-
	Match during Africa Cup	binary	-
	Match after Asian Nations Cup	binary	-
	Stadium capacity	binary	-
	Travel distance between home and away [km]	numeric	-
	Travel distance between home and away [min]	numeric	-
	Season before World Cup	binary	-
	Season after World Cup	binary	yes
	Season before European Championships	binary	-
	Season after European Championships	binary	-
	Home total market value (MV) - Away total MV	numeric	yes
	Home average MV - Away average MV	numeric	yes
	Home standardized MV - Away standardized MV	numeric	yes
	Home total MV / Away total MV	numeric	-

Continued on next page

Table 4 – continued from previous page

Description	Type	Selected
Home standardized MV / Away standardized MV	numeric	-
Home average height - Away average height	numeric	-
Home avg. height Top-11 - Away avg. height Top-11	numeric	yes
Variables once for each team		
Team plays Champions League this season	binary	-
Team plays Europa League this season	binary	-
Team plays Champions or Europa League this season	binary	home
Size of squad	numeric	home
Number of new players this season	numeric	-
Days since last match in any competition	numeric	-
Days until next match in any competition	numeric	-
Points in last league match	numeric	home
Average points in last 2 league matches	numeric	home
Average points in last 3 league matches	numeric	away
Average points in last 4 league matches	numeric	-
Average points in all previous league matches	numeric	home
Total market value (MV) of all players	numeric	home
Average MV	numeric	-
Median MV	numeric	away
Standard deviation of MV	numeric	-
Skewness of MV	numeric	-
Total MV standardized by league and season	numeric	-
Herfindahl-Hirschman-Index of MV	numeric	-
Normalized Herfindahl-Hirschman-Index of MV	numeric	-
Top 1-3 MV / Top 9-11 MV	numeric	away
Top 1-3 MV / Top 12-14 MV	numeric	-
Top 1-3 MV / Top 12-14 MV	numeric	-
Standard deviation of MV / Average MV	numeric	-
Share all players right-footed	numeric	-
Share all players left-footed	numeric	-
Share all players both feet	numeric	away
Share Top-3 players right-footed	numeric	away
Share Top-3 players left-footed	numeric	-
Share Top-3 players both feet	numeric	-
Share Top-11 players right-footed	numeric	-
Share Top-11 players left-footed	numeric	-
Share Top-11 players both feet	numeric	home
Average height all players	numeric	-
Min height all players	numeric	-
Max height all players	numeric	-
Standard deviation of height all players	numeric	-
Average height Top-11	numeric	-
Min height Top-11	numeric	-
Max height Top-11	numeric	-
Standard deviation of height Top-11	numeric	-
Min age	numeric	home & away
Max age	numeric	-
Average age	numeric	-

Continued on next page

Table 4 – continued from previous page

Description	Type	Selected
Median age	numeric	home & away
Standard deviation age	numeric	-
Standard deviation age / Average age	numeric	-
Average age of Top-11	numeric	home
Average age of 1-11 / Average age of 12-21	numeric	home
Average age of players above 20y	numeric	away

C.3 Specific framework algorithm

C.3.1 Application-specific algorithm

Algorithm 2: AllEventImportances

Data: $\mathcal{T}, x_{\text{VarSel}}, y_{\text{VarSel}}, x, y$

Result: Event importance for all leagues and seasons

begin
- $x \leftarrow \text{VarSel}(x, x_{\text{VarSel}}, y_{\text{VarSel}})$
- iter $\leftarrow 1$
- **repeat**
  - **if** *iter=1* **then**
    - $x_{\text{train}} \leftarrow x$
  - **else**
    - $x_{\text{train}} \leftarrow \{x, \text{EI}\}$
  - $\text{ORF} \leftarrow \text{trainORF}(x_{\text{train}}, y)$ ▷ Alternatively an ordered logit/probit
  - **forall** *LS in {Leagues × Seasons}* **do**
    - **forall** t *in* $\mathcal{T}^{LS}$*, backwards* **do**
      - **forall** $e_{t,i}$ *in* e_t **do**
        - **forall** $\boldsymbol{y}$ *in {H,D,A}* **do**
          - **if** $(t \neq t_{\max}) \wedge (y_{e_{t,i}} = \boldsymbol{y}) \wedge (|e_t| = 1)$ **then**
            - $r_{k,\boldsymbol{y}} \leftarrow r_{k,e_{t+1}}$
          - **else**
            - $y^*_{e_{t,i}} \leftarrow \boldsymbol{y}$
            - **forall** $e_{t,j}$ *in* e_t *with* $j \neq i$ **do**
              - $y^*_{e_{t,j}} \leftarrow \text{drawOutcome}(\text{ORF}(x_{e_{t,j}}))$
            - $x^*_{t+} \leftarrow \text{gen}(\mathcal{T}_{t-}, x_{t-}, y^*_{t-})$
            - **forall** t' *in* $\mathcal{T}^{LS}_{t+}$ **do**
              - $y^*_{e_{t'}} \leftarrow \text{drawOutcome}(\text{ORF}(x^*_{e_{t'}}))$
              - $x^*_{t'+} \leftarrow \text{gen}(\mathcal{T}^{\text{LS}}, x^*_{t'-}, y^*_{t'-})$
            - $r_{\boldsymbol{y}} \leftarrow \text{rew} \bigcup_{\mathcal{T}^{\text{LS}}} y^*_e)$
        - $r_{e_{t,i}} \leftarrow \{z_H, z_D, z_A\}$
        - **forall** k *in {Home team, Away team}* **do**
          - $\text{EI}_{e_{t,i},k} \leftarrow \text{JSD}(r_{k,e_{t,i}}, \text{ORF}(x_{e_{t,i}}))$
      - $\text{EI}_t \leftarrow \bigcup_{e_t} \text{EI}_{e_{t,i},k}$
    - $\text{EI}_{\mathcal{T}^{\text{LS}}} \leftarrow \bigcup_{\mathcal{T}^{\text{LS}}} \text{EI}_{e_t}$
  - $\text{EI} \leftarrow \bigcup_{LS} \text{EI}_{\mathcal{T}^{\text{LS}}}$
  - iter $\leftarrow$ iter+1
- **until** *iter* > 3
- **return** *EI*

Algorithm 2 describes the determination of all the event importance values in the specific framework of the presented application.

We estimate the EI in an iterative approach as described in Section 3.2.4 as the EI values are an important feature in the outcome model. The comparison of estimated EI values after different numbers of iterations in Appendix C.4.2 indicates that three iteration steps are already enough to obtain a sufficient convergence of the EI estimation. The approximation of the final ranking distribution with a Monte Carlo simulation as described in Section 3.2.3 is performed with $N_{MC} = 7500$ runs.

C.3.2 End-of-season rewards

The ultimate goal of each team is to become champions of the respective season. However, only few teams are usually capable of participating in the race for the first place. Yet, the end-of-season ranking is relevant for various matters. First, the winner is crowned champion of the respective season. Second, the participants of the international competitions for the next season will be determined. There are multiple international competitions in European club football. The UEFA Champions League is the highest-rated competition, where the best-ranked teams in the leagues participate. The next best-positioned teams participate in the UEFA Europa League. The distribution of the starting places for the upcoming season granted by UEFA by the league associations is determined by hard thresholds in the ranking. The number of allocated spots is dependent on the previous success of the association's teams and is defined in the UEFA Access List and determined by the UEFA 5-years ranking. The official access list for all years can be found in the UEFA document library: `https://www.uefa.com/insideuefa/documentlibrary/`.

Third, relegation to lower leagues is decided in the fixed breakdown of the available spots. Depending on the league, relegation can be direct or decided by a play-off match-up between a candidate for relegation and a candidate for promotion, while direct relegation always applies to the lowest-placed candidates, if both variants exist.

Fourth, financial rewards, i. e. money from broadcasting rights, are determined by the final league table. However, this gradation is less relevant for the team, the coaches, and the players. Becoming champions, qualifying for next year's European Cup or not being relegated to a lower league is what we assume is more in the focus of the involved entities. While relegation denotes a massive cut in financial benefits, reputation, and attractiveness for a club, participation in international tournaments, or becoming champion of the league tremendously boosts them. For a discussion on the financial dimensions consult Goller and Krumer (2020).

Table 5 shows the different number of international and relegation ranks observed in the investigated leagues during the considered time period. Which reward structure has been applied in which season across the leagues is shown in Table 6. In the Dutch league, the last qualification spot for the Europa League is determined in a play-off format including four teams. These ranking positions therefore are aggregated to one reward.

Table 5. Code for every reward structure in the data.

Code	Champions league	Europa league	Relegation play-off	Direct relegation
4/3/DDD	4	3	0	3
4/3/PDD	4	3	1	2
3/3/DDD	3	3	0	3
3/3/DD	3	3	0	2
3/3/PDD	3	3	1	2
3/3/PD	3	3	1	1
2/4/DD	2	4	0	2
2/3/DDD	2	3	0	3
2/3/DD	2	3	0	2
2/3/PPD	2	3	2	1
2/2/DD	2	2	0	2
2/2/PPD	2	2	2	1

Table 6. Implemented rewards structure for each league and season.

League	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	2016/017	2017/18	2018/19
Bundesliga	3/3/PDD	3/3/PDD	4/3/PDD	4/3/PDD	4/3/PDD	4/3/PDD	4/3/PDD	4/3/PDD	4/3/PDD	4/3/PDD
Eredivisie	2/3/PPD	2/2/PPD	2/3/PPD	2/3/PPD	2/3/PPD	2/2/PPD	2/2/PPD	2/2/PPD	2/2/PPD	2/2/PPD
La Liga	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD
Ligue 1	3/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	3/3/PDD	3/3/PDD	3/3/PDD
Premier League	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD	4/3/DDD
Primeira Liga	2/3/DD	2/4/DD	3/3/DD	3/3/DD	3/3/PD	3/3/DD	3/3/DD	3/3/DD	2/3/DD	2/3/DDD
Serie A	4/3/DDD	4/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	3/3/DDD	4/3/DDD	4/3/DDD

Note: Codes are explained in Table 5. The 2006/07 through 2008/09 seasons are excluded as they are only used for model selection.

C.3.3 Individually different end-of-season rewards

Among the strengths of the proposed algorithm is that individually different and asymmetric reward schemes can be incorporated. In the application of European football leagues, differential reward schemes are not uncommon. All national league associations in the investigated European football leagues award a starting place for the UEFA Europa League to the winners of national cup tournaments. In the case of two national cup competitions, one spot is reserved for each winner. If a cup winner would also qualify for an international tournament by finishing high enough in the league (or winning the second cup title), the allocation of European starting places is regulated by the league associations. Typically, the place reserved for the cup winner is transferred to an additional place determined by the league table. This has consequences for this particular team, as its reward scheme distributed via the league table has changed, as well as for all other teams such that a lower position in the league table is already sufficient for them to qualify for a European starting place.

We consider the possibility of this special case from the moment the final pairing of the cup tournament is fixed. In every simulation run, we evaluate at the end of the season, whether all teams that are still in the hunt for the cup win at the moment of the initial event meet the required configuration in the ranking. If this condition is met, the threshold in the rankings for the Europa League is lowered by one rank. Depending on how likely it is that the qualifying place is transferred to the league ranking based on the initial situation, the probability distribution of the final rewards is the weighted average of the two possibilities. If there are two cup tournaments in one country, this procedure is evaluated for both tournaments, taking into consideration that a team possibly plays in both finals.

Other reasons for individually different reward schemes for clubs are e. g. the exclusion of teams from international competitions, winning the UEFA Champions/Europa League, or legal issues of teams. All of them are taken into account as individually different reward schemes but are not addressed in detail.

C.3.4 Details of the outcome model

We estimate the outcomes of football matches following the approach of Goller et al. (2021), which has proven successful in the outcome prediction of football matches. Our initial data set contains 133 variables. We perform a LASSO-based model selection step on the initial set of variables using the subset of the 2006/07 to 2008/09 seasons across all leagues. The variable selection with a linear LASSO model uses the points won by the home team as the outcome variable, 10 folds cross-validation, and the optimal λ at minimum MSE. The 2006/07 to 2008/09 seasons are excluded from the remainder of the analysis. After this data-driven variable selection, 26 ($\sim$ 20%) variables remain in the model - the full list of variables, as well as those selected, can be found in Appendix C.2. This model selection procedure is consistent with Borup, Christensen, Mühlbach, and Nielsen (2022), who find that for predictive models with many covariates, variable selection benefits prediction accuracy and usually 10-30% of the original set of covariates turns out to be optimal.

The ORF predicts probabilities of ordered outcomes in a flexible way, building on Random Forests developed in Breiman (2001). This machine learning method is tailored for the predictions of outcomes that appear in a natural order. The ORF model is estimated using the R-package `orf` with 1000 trees, without honesty option, minimal node size is 5, 5 randomly selected variables considered at each split and a sub-sampling rate of 2/3. The general framework is not restricted to this specific method and the choice of the underlying outcome model is of second-order (see Appendix C.4.4). Other well-suited methods to predict match outcomes could be employed as well, like ordered logit, ordered probit, Poisson distribution-based models, or others. The exact score of the match is drawn from two independent Poisson distributions, for which the λ parameters of the Poisson distributions of the scored goals are estimated on the full data set. The obtained values are 1.55 for the home team and 1.16 for the away team.

C.4 Additional results

C.4.1 An illustrative example of EI estimates

To illustrate the EI estimates, we show in Table 7a the ranking of the 2017/18 Bundesliga 1 season before the last match day and in Table 7b the fixtures of the final match day together with the estimated EI values. Several teams have already settled in their reward area, e. g. teams ranked 10^{th} to 14^{th}, for which the EI estimates are zero. The teams ranked 4^{th} to 8^{th} are in strong competition for the qualification for the European club tournaments, while Freiburg, Wolfsburg, and Hamburg fight against relegation to Bundesliga 2. In consequence, those teams have a particularly high EI estimate. Currently ranked 16^{th} Wolfsburg has the highest EI estimate of all teams, as a win in their last match can lift them out of the relegation zone and losing could result in direct relegation. For Dortmund and M.gladbach, it is highly unlikely that they climb either up or down to a different reward, as this requires not only a particular outcome of their own match but also of other matches (including a sufficiently large shift in the goal difference on top) – this is reflected in EI estimates very close to 0.

The results of the fixtures in Table 7b suggest, that incentives, as measured by the EI, might have an impact on the performance of the teams. Teams with a large EI estimate succeed over teams with a low or zero EI value, even though their opponents are much stronger on paper, i. e. Stuttgart beating the champions Bayern M. or Hamburg winning against M.gladbach.

C.4.2 Compare EI by iteration

Figure 9 shows the convergence of the estimated EI values over successive iterations. The second iteration of Algorithm 2 incorporates the EI estimates of the first iteration into the outcome model which has a notable effect on the results. Integrating

Table 7. The final match day of the 2017/18 German Bundesliga 1 season

(a) Ranking prior to match day 34

Rank	Team	Played	GD	Points
1	Bayern M.	33	67	84
2	Schalke 04	33	15	60
3	Dortmund	33	19	55
4	Hoffenheim	33	16	52
5	Leverkusen	33	13	52
6	RB Leipzig	33	0	50
7	Frankfurt	33	1	49
8	Stuttgart	33	-3	48
9	M.gladbach	33	-4	47
10	Hertha BSC	33	1	43
11	Augsburg	33	-1	41
12	Bremen	33	-4	39
13	Hannover	33	-9	39
14	Mainz 05	33	-13	36
15	Freiburg	33	-26	33
16	Wolfsburg	33	-15	30
17	Hamburg	33	-25	28
18	Köln	33	-32	22

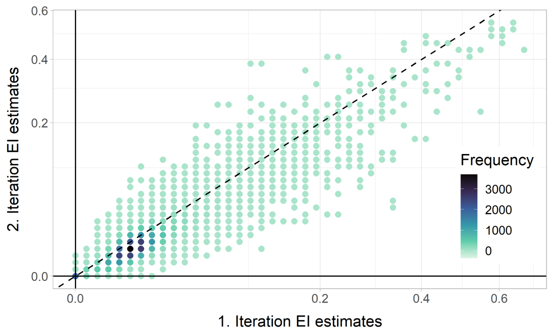
(b) Results of match day 34 and EI estimates.

Match result			EI	
			H	A
Bayern M.	1:4	Stuttgart	0.00	0.18
Hoffenheim	3:1	Dortmund	0.21	0.00
Hertha BSC	2:6	Leipzig	0.00	0.10
Freiburg	2:0	Augsburg	0.16	0.00
Schalke 04	1:0	Frankfurt	0.00	0.11
Leverkusen	3:2	Hannover	0.08	0.00
Hamburg	2:1	M.gladbach	0.09	0.02
Mainz 05	1:2	Bremen	0.00	0.00
Wolfsburg	4:1	Köln	0.24	0.00

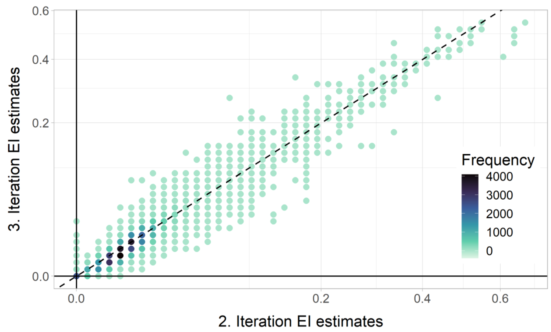
Note: In (a): GD refers to the difference between scored and conceded goals. Lines split the different reward areas in decreasing ordering (Championship, Champions League, Europa League, none, Relegation Playoffs, Relegation). In (b): EI provides the estimated event importance for the home (H) and away (A) team.

the refined estimates adds only a little information to the outcome model which is affirmed by the very high correlation of 0.96 for the estimated values of iterations 2 and 3. The fourth iteration increases the correlation to the previous estimates to 0.98, which cannot be substantially increased anymore.

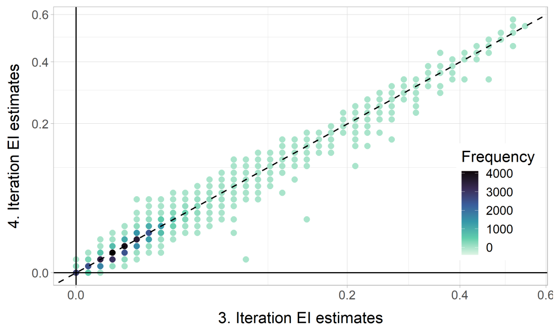
Because of the random component in Monte Carlo simulation, two consecutive iterations will never return the exact same values. More iterations cannot alleviate this source of randomness and exhibit similar correlations. Convergence can be improved further by increasing the number of runs in the Monte Carlo simulation.



(a) Iterations 1 and 2



(b) Iterations 2 and 3



(c) Iterations 3 and 4

Fig. 9. Comparison of EI estimates for different iterations

C.4.3 Variable importance in the outcome model

Table 8 displays the permutation-based variable importance of the rich ORF model including the EI estimates as input variables. The EI variables are among the most important variables which supports the evidence of the relevance of the EI in the prediction of the match outcomes. If the difference between the home and away EI is also included in the model, this third variable also has considerable variable importance, which is, however, lower than that of the home or away EI.

Table 8. Variable importance in the ORF model including the event importance estimates.

Variable	Importance
Away EI	0.052
Home average MV - Away average MV	0.036
Home EI	0.036
Home standardized MV - Away standardized MV	0.033
Home total market value (MV) - Away total MV	0.028
Home total market value (MV) of all players	0.007
Away median MV	0.005
Home average points in all previous league matches	0.005
Home plays Champions or Europa League	0.002
Home average age of Top-11	0.001
All other variables	≤ 0.001

Note: Permutation-based variable importance in rich ORF model (Base covariates + Home EI + Away EI). Only the top 10 variables are shown.

C.4.4 Alternative outcome model

As described in Section 5.2.2 we use an ordered choice model with outcome probabilities estimated with an Ordered Forest (ORF). In Figure 10 we compare the estimates between an ORF and an ordered logit outcome model. The estimated EI values are highly correlated (0.97) which suggests that the framework does not rely on a specific outcome model.

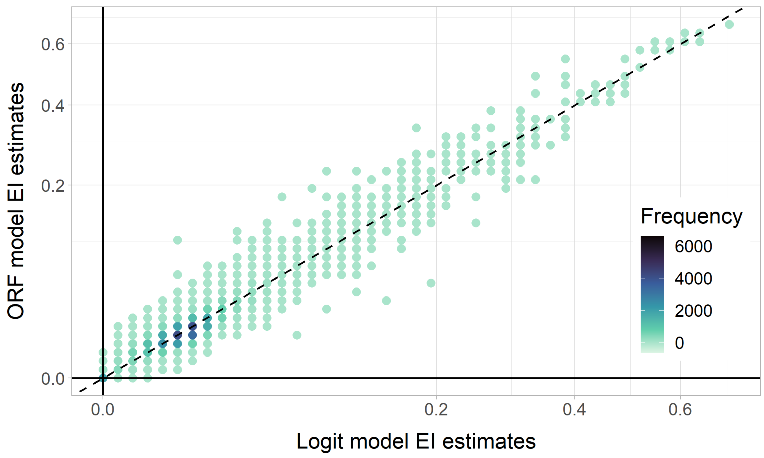
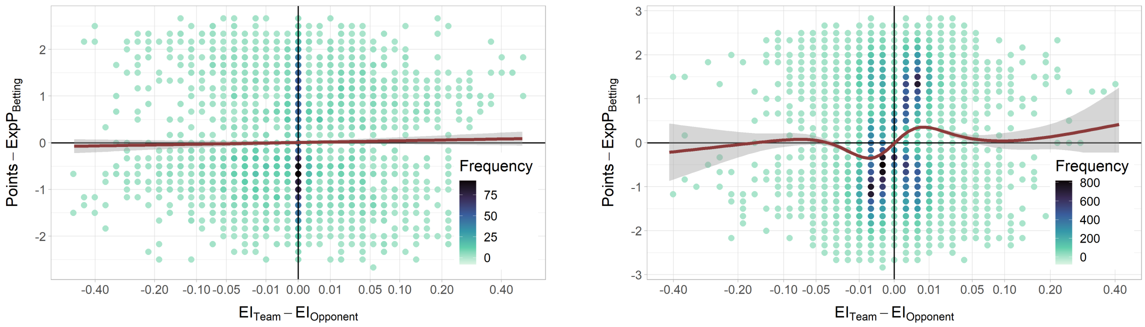


Fig. 10. Event importance (EI) estimates for the ORF and logit outcome model. Values are rounded to the nearest grid point. Frequency indicates the number of points on a grid point. Square-root transformation to x and y-axis applied.

C.4.5 Improvement on betting odds

In Section 5.3.3 we provide evidence that bookmakers are successful in integrating imbalances in incentives induced by large differences in the EI measure but fail to reflect the subtle differences in the EI values in the betting odds. To support this hypothesis we split the data into cases where at least one of the teams exhibits a zero EI (Figure 11a) and all other matches where both competing teams have a positive estimated EI (Figure 11b). The easily grasped case – a team is settled in its final reward area – is well accounted for in the betting odds as no difference between the expected and realised points can be explained by the EI measure. In matches between teams that both have a positive EI, their EI difference has still valuable information content which is not entirely integrated into the betting odds.



(a) At least one contestant with EI=0 **(b)** Both contestants with EI>0

Fig. 11. The difference in realised and expected points by the event importance for the team and its opponent, in samples with at least one contestant with null EI (a) and no contestant with null EI (b). Values are rounded to the nearest grid point. Frequency indicates the number of points on a grid point. Square-root transformation to x-axis applied.

To complete this analysis, we split the prediction sample into three groups indicating whether the matches are played in the first, second, or third chronological third of a season in Figure 12. This split points out, that the improvement relative to the betting odds arises from the first and second third of the season where differences in the EI are still very small and hence are not reflected in the betting odds.[17]

[17] In general, the matches in the last tier of the season are less accurately predicted by all models. This can be explained by the decreasing capacity of the available covariates to comprehensively characterise the teams with the duration of the season and by the fewer available matches with highly disparate EI values in the training data, which are frequent in the last third of the season.

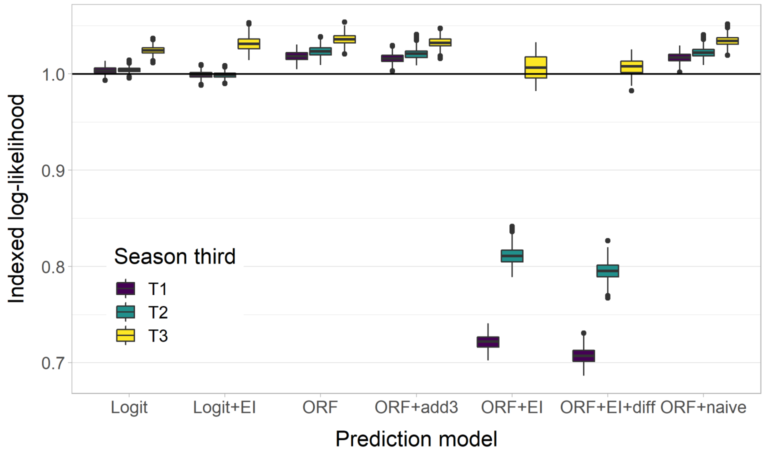


Fig. 12. Indexed log-likelihood of out-of-sample predictions of different models split in the first (T1), second (T2), and last (T3) third of the season over 1000 replications.

C.4.6 Alternative predictive power measure

To highlight the independence of the out-of-sample prediction power analysis on the measure, we provide the results of the identical analysis procedure measuring the prediction accuracy with the Brier score. Figures 13 & 14 are the equivalent plots to Figures 5 & 12 using the Brier score instead of the log-likelihood and exhibit the equivalent patterns independent of the measure.

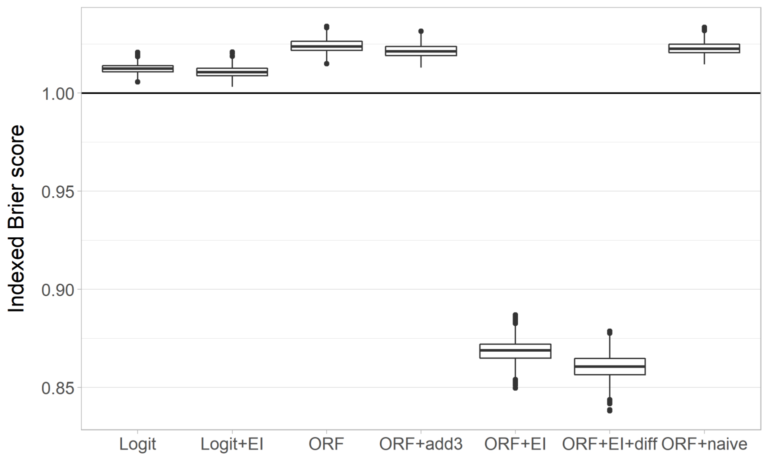


Fig. 13. Brier score out-of-sample prediction accuracy of different models, indexed by the performance of betting odds.

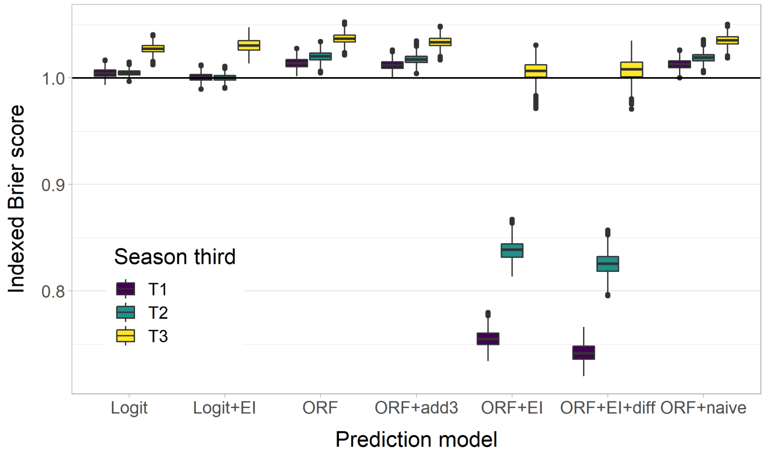
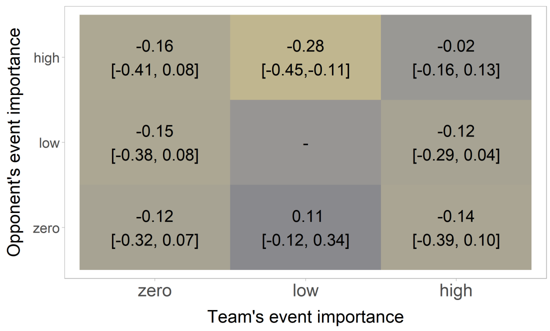
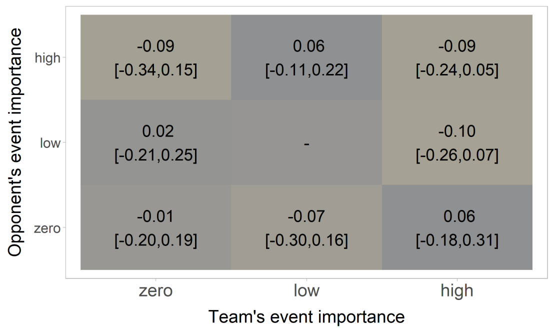


Fig. 14. Brier score of the out-of-sample prediction accuracy of different models split in the first (T1), second (T2), and last (T3) third of the season

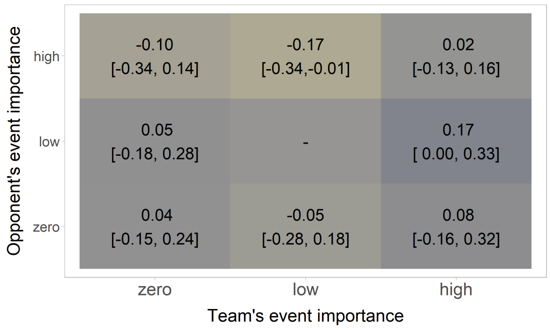
C.4.7 Team performance: Additional outcomes



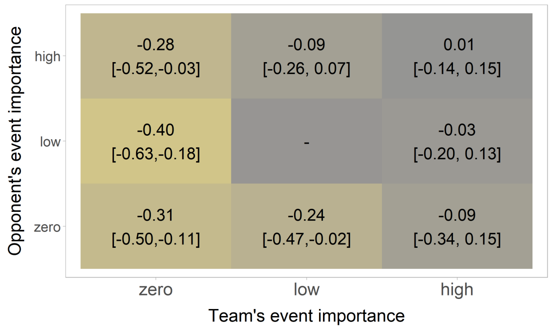
(a) Tackles



(b) Tackles win share

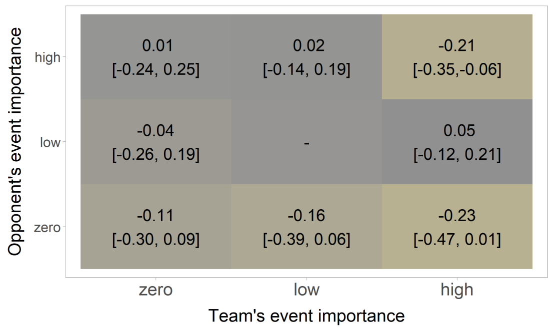


(c) Duels win share

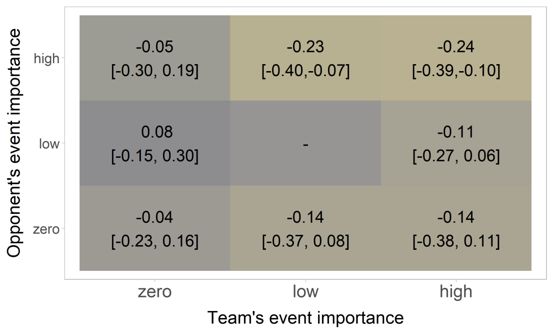


(d) Fouls

(e) Yellow cards

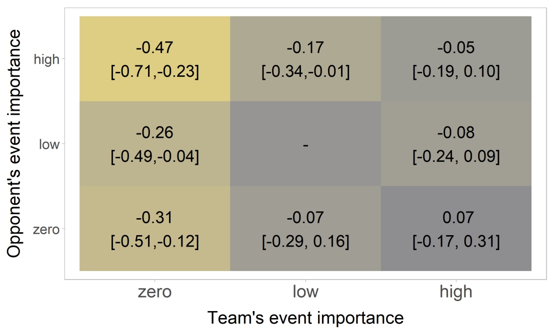


(f) Sent off

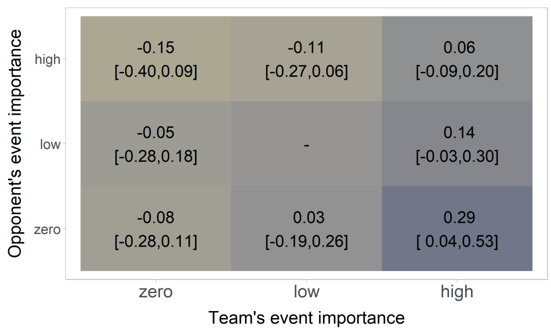


(g) Touches

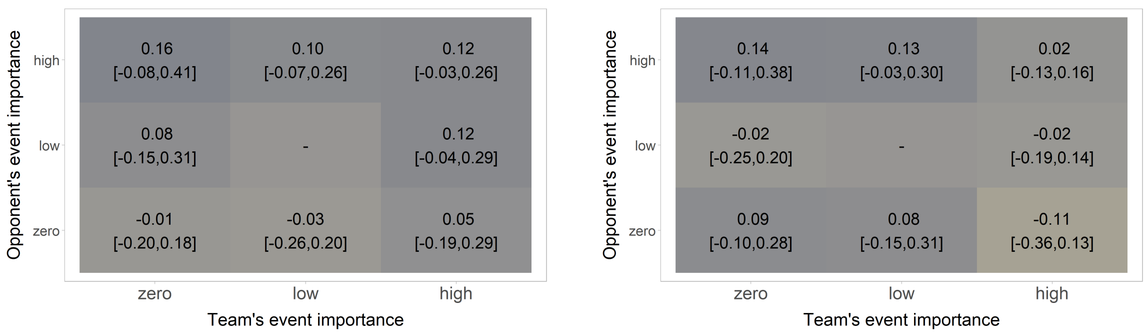
(h) Dribbling attempts



(i) Final third entries



(j) Penalty area entries



(k) Errors leading to shot

(l) Errors leading to goal

Fig. 15. Estimates of the FE regression residuals of different outcomes on the event importance categories. 95% confidence intervals in parentheses. The baseline is the low by low category.