arXiv:2605.30209v2 [econ.GN] 7 Oct 2026

Integrity at Stake: Statistical Detection of Anomalous In-Play Betting in Football

David Winkelmann[† ‡ § ⋆], Maya Vienken[† ‡], Christian Deutscher[†], and Roland Langrock[‡]

[†]Department of Psychology and Sports Science, Bielefeld University, Bielefeld, Germany

[‡]Department of Business Administration and Economics, Bielefeld University, Bielefeld, Germany

[§]Department of Business Decisions and Analytics, University of Vienna, Vienna, Austria

[⋆]Corresponding author: david.winkelmann@univie.ac.at

October 8, 2026

Abstract

Match-fixing undermines the integrity of sport by eroding public trust and threatening the financial sustainability of clubs and leagues. The expansion of sports betting markets has created new incentives and opportunities for manipulation, calling for rigorous, data-driven monitoring tools. Football, which accounts for the largest share of global betting turnover, remains particularly exposed: integrity reports continue to flag several suspicious matches, with past scandals in Italy and Turkey underlining the problem's persistence. This study uses high-frequency in-play (live) betting data from the Italian Serie B (seasons 2018/19–2020/21) to explore statistical approaches for detecting unusual betting behaviour. A state-space modelling framework is employed to describe standard betting market dynamics and to predict expected betting volumes conditional on match characteristics. Deviations from these expectations can then be analysed using

outlier detection techniques to identify potentially suspicious betting activities. The results demonstrate how statistical modelling can contribute to the early identification of unusual betting patterns, thereby supporting integrity assurance in in-play sports betting markets.

Keywords: OR in sports; betting markets; match-fixing; outlier detection; state-space modelling.

1 Introduction

Match-fixing poses a serious threat to the integrity of sports competitions. It undermines public trust and diminishes fan engagement, thereby threatening the commercial viability of clubs, leagues, and associated industries. It also compromises the role of sport as a vehicle for fair play, community cohesion, and youth development. Match-fixing becomes lucrative and more likely when fixers can profit financially. The proliferation of global sports betting markets is hence accompanied by an increase in opportunities for profitable match-fixing, due to the ease of access and high liquidity, which reduces the chances of detection. If match-fixing were to become a regular and undeniable occurrence, it would further affect public trust and discourage fair athletes (Breuer and Forrest, 2018). This underscores the need for sophisticated, data-driven systems to detect and prevent betting-related corruption, thereby securing the integrity of sports.

Football is the frontrunner in terms of betting turnover worldwide (Statista, 2025) and also the sport with the highest number of suspected match-fixing incidents reported by the industry: In their 2024 annual integrity report, Sportradar flagged 721 suspicious football matches (Sportradar, 2025). While match-fixing is often associated with individual matches, some scandals even span multiple matches. Notably, the Italian Serie B has been repeatedly implicated, with the 2011–12 scandal involving players such as Giuseppe Signori and Cristiano Doni, leading to sanctions for clubs and individuals. More recently, the 2015 "Dirty Soccer" investigation revealed that Catania Calcio's management orchestrated the fixing of five matches to avoid relegation, resulting in arrests and league penalties (Costa, 2018). In 2025, a betting scandal in Turkey involving illegal sports betting and referees, club presidents, and footballers was revealed (Kocsoy, 2026). These incidents highlight the ongoing challenges in safeguarding the integrity of football competitions.

A major difficulty in assessing fraud detection models lies in the limited availability of verified ground-truth data on match-fixing. However, given the incidents in the Italian Serie B, this league appears to be a natural platform for testing candidate models for fraud detection and for investigating potentially interesting anomaly patterns in betting behaviour. Although automated data-driven corruption-detection systems can clearly be valuable tools for identifying irregularities and hence protecting the sport, the corresponding literature to date is relatively scarce, given the market size. It involves data analyses and outlier detection methods applied to, e.g. basketball (Wolfers, 2006), sumo wrestling (Duggan and Levitt, 2002), and football (Reade and Akie, 2013; Ötting et al., 2018). However, the existing literature has mostly focused on pre-match markets, despite the fact that the in-play (live) betting market today accounts for about 40% of the total betting volume and is particularly relevant when examining suspicious matches (European Gaming & Betting Association, 2025). To comprehensively monitor irregular betting activity, there is thus a pressing need to develop fraud warning systems for the in-play betting market as well. Furthermore, the high resolution and, consequently, the granularity of betting activity surrounding major in-play events may offer additional insights and support fraud detection that would otherwise remain hidden. For example, unusually high betting volumes might be placed on relatively unlikely outcomes shortly before the end of matches – so as to maximise profits as the odds increase with time running out – which were manipulated to ensure that outcome.

We extend previous research by analysing in-play betting data to investigate potentially fraudulent betting behaviour in the in-play betting market. Specifically, this paper seeks to identify potentially suspicious deviations from typical patterns in betting market activity during football matches. This requires two steps: First, the dynamics of betting markets under regular conditions need to be understood. This provides the foundation for predicting expected betting volumes. Second, we compare these predictions with the actually observed stakes to detect anomalies. Outlier detection techniques are then applied to flag potentially suspicious matches or situations that may indicate manipulation. To this end, we consider high-resolution betting data from the 2018/19–2020/21 seasons of the Italian Serie B, provided by a major European bookmaker. The dataset contains second-by-second records of betting volume for potential match outcomes, along with the bookmaker's in-play odds. Following approaches used in general financial market analyses (see, e.g. Barra et al., 2017

(Warne et al., 2017)), we employ a state-space modelling framework to capture the underlying serial correlation in the market activity level. Controlling for key covariates capturing responses to major events (Ötting et al., 2024) and in-play dynamics (Michels et al., 2023) as well as the anticipation of news in betting markets (Winkelmann and Deutscher, 2026) as the expected key drivers of betting actions, we isolate unexplained large positive deviations from the model-implied expected betting volumes as the instances of potential interest. Our study can be seen as a proof-of-concept for the practical benefits of advanced statistical modelling as a tool for flagging suspicious betting volumes in in-play betting.

2 Characteristics of betting markets

Sports betting markets are of high economic relevance, given their gross revenue of approximately €21 billion in Europe (European Gaming & Betting Association, 2025). From an economic perspective, they share various similarities with financial markets. In particular, placing a bet is conceptually similar to buying a financial asset (Sauer, 1998). A crucial advantage of betting markets is that the outcome of a bet (the asset's uncertain value) becomes observable at a fixed deadline, typically the end of a match (Thaler and Ziemba, 1988). According to the efficient market hypothesis, betting odds offered by bookmakers must contain all available information (Fama, 1970), thereby eliminating any systematic strategies that could generate returns for bettors. However, bookmakers maximise their profits by "balancing their books", i.e. they offer odds in response to (expected) bettors' demand, such that they secure profitability. Thus, odds-implied probabilities (corrected for the bookmaker's overround, cf. Winkelmann et al., 2024) may deviate from *true* probabilities, challenging the efficient market hypothesis (Thaler and Ziemba, 1988).

Potential match-fixing can manifest itself in anomalies, i.e. large unexplained deviations from typical patterns in betting markets. Effectively identifying such deviations requires an accurate quantification of the market dynamics. To this end, this section reviews the literature on the structure of both pre-match and in-play betting markets. Furthermore, it provides an overview of recent approaches to detecting betting fraud.

2.1 Pre-match betting

Most research on betting markets has focused on pre-match betting (see, e.g. Thaler and Ziemba, 1988; Dixon and Pope, 2004; Feddersen et al., 2017; Deutscher et al., 2018), as in-play betting has only recently become more common. However, understanding pre-match market behaviour is also relevant for the in-play market, as in-play betting odds are built on these initial expectations. The majority of the corresponding studies do not focus on fraud detection, but rather on potential market inefficiencies, i.e. situations where the odds do not accurately reflect the true probability of an outcome. Nevertheless, these contributions help to understand the dynamics of the betting market and can thus support the development of a fraud detection system.

Market efficiency is often evaluated by testing simple strategies, such as betting on home teams or favourites, with mixed results (Hegarty and Whelan, 2024). Some studies find general inefficiencies (Constantinou and Fenton, 2013; Dixon and Pope, 2004; Angelini and De Angelis, 2019), while others highlight specific cases, such as recently promoted teams (Deutscher et al., 2018), end-of-season matches (Borghesi, 2007), or the reduced home advantage during COVID-19 (Winkelmann et al., 2021). Many of the observed inefficiencies appear to reflect irrational behaviour (biases) rather than fraudulent activity.

For example, the *favourite-longshot bias* (FLB) refers to some bettors' tendency to overvalue longshots and undervalue favourites (Ottaviani and Sørensen, 2008; Snowberg and Wolfers, 2010). Several studies confirm the FLB in European football betting markets (Cain et al., 2000; Constantinou and Fenton, 2013; Buhagiar et al., 2018; Angelini and De Angelis, 2019), while others find a reverse FLB, where favourites are overvalued (Dixon and Pope, 2004; Gil and Levitt, 2007; Woodland and Woodland, 1994). Newall and Cortis (2021) suggest that the direction of the bias depends on the number of outcomes and the favourite's implied probability; in addition, the FLB is linked to sentiment and risk preferences (Forrest and Simmons, 2008). Other studies provide evidence that bookmakers undervalue the advantage of playing at home, driven by venue familiarity, reduced travel fatigue, and crowd support (Carmichael and Thomas, 2005; Carron et al., 2005), thus offering higher odds for home teams, referred to as *home bias* (Forrest and Simmons, 2008; Angelini and De Angelis, 2017). The *sentiment bias* suggests that bettors favour well-known or popular teams, but the

evidence in the literature is again mixed (Feddersen et al., 2017; Forrest et al., 2008).

In a comprehensive study, Winkelmann et al. (2024) analyse common pre-match biases for the top five European football leagues and several seasons. They consider their empirical results in light of a simulation-based analysis, demonstrating that the frequency of inefficiencies is barely higher than what would be expected in a fully efficient market by chance alone. This leads to the conclusion that inefficiencies in pre-match betting markets are, contrary to the impression given by previous literature, neither persistent nor systematic across leagues.

2.2 In-play betting

While pre-match odds are relatively stable and mainly respond to injuries or line-up changes, in-play betting is highly dynamic: beyond pre-match expectations, odds and stakes continuously respond to evolving match dynamics and events, such as goals, red cards, or momentum shifts. While the share of revenue from in-match betting, relative to the total volume of sports betting, fluctuates around 40% (European Gaming & Betting Association, 2025), there is still very little publicly available academic research on in-play betting, with most of the existing literature considering the case of betting exchanges. For the latter, Choi and Hui (2014) find that bettors underreact to moderately surprising events but overreact to major ones, with the effects vanishing within minutes. Further studies provide mixed evidence regarding market efficiency (Croxson and Reade, 2014; Gil and Levitt, 2007). In their mispricing detection framework, Angelini et al. (2022) test efficiency and assess behavioural biases in in-play betting exchange data. Using first goals as events of major news, they identify pre-match and in-play inefficiencies to positively correlate with the level of surprise, driven by biases such as the reverse FLB.

The recent availability of in-match betting data from bookmakers, including both odds offered and stakes placed by bettors, allows for increasingly comprehensive analyses of in-play betting markets. Michels et al. (2023) find that betting decisions reflect both pre-match expectations and in-match team strength, with bettors placing more money on matches exhibiting higher outcome uncertainty, particularly in the final 30 minutes of a match (Ötting et al., 2024). Research on goals as a major in-match event with its decisive impact in the low-scoring sport of football suggests that neither bookmakers nor bettors anticipate their

occurrence, indicating that unusual odds movements or substantially elevated stakes placed shortly before a team scores may indicate fraudulent behaviour (Winkelmann and Deutscher, 2026). In contrast, market activity does increase after major news, especially after surprising goals (Ötting et al., 2024). This is consistent with the *cognitive bias* in psychology and has been observed across different sports (Durand et al., 2021).

2.3 Fraud detection in pre-match and in-play betting

The substantial growth of the online and, in particular, the in-play betting market has fundamentally changed the environment in which match-fixing can occur, offering increasingly many avenues for manipulation and complicating its detection, as betting activity is globally distributed and no unified global monitoring framework exists. Match-fixing occurs more frequently when corrupt behaviour appears justifiable or attractive to protagonists, for example, due to dissatisfaction with low wages or in matches with little at stake, typically towards the end of a season. As a consequence, second- or third-tier leagues, within which players and referees receive comparatively low compensation, while the betting markets are still highly liquid, are especially prone to corruption (for broader discussions of the motives behind match-fixing and documented cases in professional football, see Forrest et al., 2008; Forrest, 2012). On the referee side, anomalies such as systematically higher betting volumes for matches officiated by certain referees have been documented (Deutscher et al., 2025).

Despite these risk factors outlined above, detecting suspicious activity remains challenging. Unusual betting patterns or poor on-field performance may simply result from players conserving energy in low-stakes matches (Forrest et al., 2008). The detection of fraud in in-play betting is even more complicated due to short time windows within which actions occur, strongly fluctuating odds, and higher betting limits (Van Rompuy, 2015). Recent approaches to fraud detection increasingly rely on data analytics and artificial intelligence. Systems such as Sportradar's *Universal Fraud Detection System* (UFDS) compare in-play market odds with model-based forecasts (Sportradar, 2024). In the absence of manipulation, both should closely align as all information is publicly available (Reade and Akie, 2013). Forecasting methods employed include econometric models (Reade and Akie, 2013), Bayesian networks (Razali et al., 2017), Bayesian inference with rule-based reasoning and time-series inputs (Min et al.,

2008), and forensic statistics for in-play forecasting (Forrest and McHale, 2019). Other studies rely on machine learning approaches, such as deep neural networks (Anfilets et al., 2020), long short-term memory networks (LSTMs) models with attention mechanisms (Zhang et al., 2022), or a combination of different methods (Kim et al., 2024). Most of the existing literature focuses exclusively on betting odds when detecting market anomalies. However, Ötting et al. (2018) show that including betting volumes can enhance model accuracy, as high liquidity can stabilise odds and conceal manipulation. The authors combine odds and volumes and apply outlier detection using distributional regression on pre-match Serie B data, successfully identifying suspicious matches in seasons with known fixes. Their findings suggest that continuous monitoring of both odds and volumes can improve match-fixing detection, forming the basis for our study.

3 Data

Our dataset is provided by one of Europe's largest bookmakers and contains in-play betting records from three consecutive seasons of the Italian Serie B (2018/19–2020/21). The sample covers 1,097 matches involving 32 different teams. Four matches were excluded due to missing bookmaker data, and one further match was omitted because it was interrupted by fog and resumed on the following day. A football match consists of two halves of 45 minutes each, typically extended by a few minutes of injury time at the end of each half. For each match, our dataset includes information on match-specific characteristics, bookmaker odds, and betting activity. The original data were recorded at a very high resolution of 1 Hz. To reduce volatility in betting stakes, thereby facilitating the statistical modelling of the betting dynamics, we aggregate observations into one-minute intervals (cf. Michels et al., 2023; Winkelmann and Deutscher, 2026). To avoid the need to deal with noise and often extreme covariate values (e.g. for the Gini coefficient) during the very last minutes of a match, we exclude observations from the injury time after the second half.

We are specifically interested in bets placed on the home and away teams, as betting on draws is generally unpopular in football – in fact, in our dataset, more than 80% of the stakes are placed on one of the two teams – and, more importantly, draws can be expected to be less prone to fraud, as this would typically require the coordinated cooperation of

both teams. Therefore, we mirror the dataset and include each observation twice, once from the home team's perspective and once from the away team's perspective. This results in a total of 235,882 observations. To comply with confidentiality requirements, bettors' stakes are scaled by a fixed constant, ensuring that relative comparisons across observations remain valid. Stakes can only be placed when the betting market is open. We define an open market as the possibility to bet on all three outcomes. Bookmakers close the market right after major events, such as goals and red cards, and towards the end of a match if its outcome is effectively decided. We capture this information by the variable *open* taking the value 1 if the market is open. We observe a closed market ($open = 0$) for 9.8% of our observations. We are mainly interested in the stakes placed on a certain team within a given minute, denoted by the variable *stake_team*. In our sample, stakes range from 0 to 257.03 and are highly right-skewed (median: 0.37, mean: 1.19; cf. Table 1), with no stakes placed on a team despite an open market in 9.4% of observations. To facilitate the modelling in case of positive stakes, we therefore use the logarithm of these stakes in our models in Section 4 (see the boxplots in Figure A1 in Appendix A for a visualisation of stakes).

The dynamics of betting activity are affected by various drivers, which we quantify using both pre-match and in-match variables. Betting stakes vary strongly with the teams involved in a match. The average stakes per minute placed on the different teams in our dataset, denoted by *stake_avg_team*, are illustrated in Figure B1 in Appendix B (for the total stakes placed on each team over the whole observation period, see Figure C1 in Appendix C). The distribution is again right-skewed, with a mean of 1.49 and a median of 1.18. Hellas Verona and US Palermo appear as outliers with exceptionally high average stakes. In addition, stakes vary with the *minute* of the match: Figure 1 displays average stakes placed as a function of the time elapsed since kick-off. On average, the highest stakes are placed during the initial 10 minutes of a match. They again increase during halftime and late in the second half.

Generally, pre-match characteristics such as team strength and squad changes are captured by the implied winning probabilities, derived from pre-match odds offered by the bookmaker (cf. Winkelmann et al., 2024). We find that the pre-match implied probabilities of teams, denoted by *improbteam_start*, vary between 7.69% and 73.85% (mean 34.15%). These pre-match probabilities are higher for home teams (42.15%) than for away teams (28.46%), reflecting the well-known home advantage.

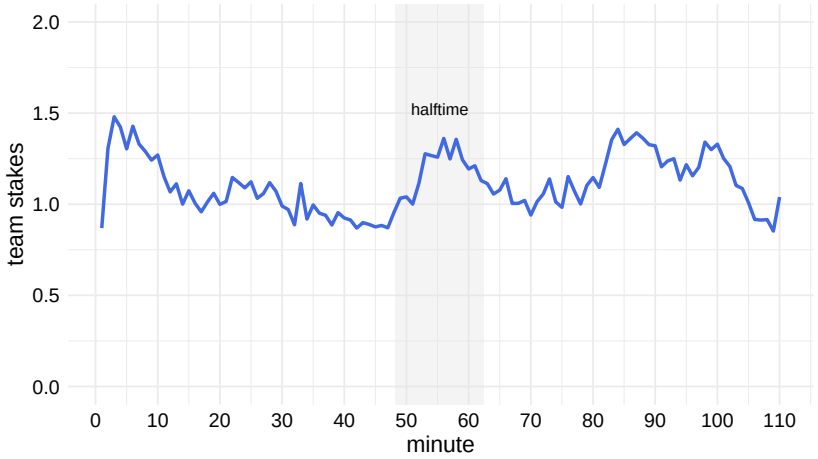


Figure 1: Average stakes placed per minute. The average halftime is highlighted in grey. To avoid misleading impressions induced by a small number of matches with very long injury times, we display average stakes only until 110 minutes after kick-off.

In-match implied probabilities, and likely also betting activity, are affected by the score difference, by major events such as goals and red cards, and by the time elapsed since kick-off (Winkelmann and Deutscher, 2026). The number of total goals in our dataset ranges from 0 to 8 (mean: 2.32). Whether a match is effectively decided is mainly determined by the difference in the score between both teams, denoted by *scorediff_team* from the perspective of the team under consideration. A positive score difference indicates that this team is in the lead. Figure D1 in Appendix D displays the distribution of stakes placed depending on the score difference. The magnitude of in-play shifts in winning probabilities depends on the level of surprise associated with a goal. We follow the definition of Ötting et al. (2024) and denote surprising goals as those scored by teams with a pre-match odds-implied winning probability of at most 0.25. Analogously, slightly surprising (unsurprising) goals are defined as those scored by teams with a pre-match implied winning probability between 0.25 and 0.5 (above 0.5).

Bettors may react not only to goals themselves but also to major goal-scoring opportunities that potentially shift the dynamics of the match (Winkelmann and Deutscher, 2026). We therefore consider the expected (number of) goals retrieved from `https://fbref.com` from the start of the match until a given minute under consideration as an additional covariate. We denote the difference in expected goals between the team under consideration and its opponent as *xg_diff*, taking a maximum value of 4.78. In addition to goals and goal-scoring

opportunities, we consider the variable *redcard_team*, which denotes the number of red cards issued to the team, whereas the variable *redcard_opponent* indicates the number of red cards issued to its opponent. Within our dataset, 231 red cards were issued, with at most two within a single match. To measure uncertainty about the match outcome, we calculate the (normalised) Gini coefficient across implied probabilities for the three potential outcomes. The variable *Gini* provides its current value in a given minute. The Gini coefficient ranges from 0 to 1, with lower values indicating greater uncertainty about the match outcome and potentially attracting bettors to place higher stakes.

Table 1: Summary statistics of the key variables considered in our model.

variable	mean	sd	min	max	median
stake_team	1.19	3.18	0.00	257.03	0.37
stake_avg_team	1.42	0.90	0.51	6.60	1.23
improbteam_start	0.35	0.12	0.08	0.74	0.34
redcard_team	0.03	0.17	0.00	2.00	0.00
redcard_opponent	0.03	0.17	0.00	2.00	0.00
scorediff_team	0.00	1.04	−7.00	7.00	0.00
xg_diff	0.00	0.71	−4.78	4.78	0.00
Gini	0.45	0.28	0.00	0.99	0.42

Table 1 presents summary statistics for the key variables to be included in our model. Additionally, we consider the situation of no stakes being placed on a team. We observe this situation more frequently when the Gini coefficient is higher (correlation coefficient: 0.41), i.e. when the outcome of the match is more certain, and later in the match (correlation coefficient with the minute of the match: 0.23).

4 Modelling betting market dynamics

The data to be modelled are the sequences of aggregated stakes placed by bettors per minute. These sequences exhibit fairly strong serial correlation, induced by the underlying market dynamics, with matches generally going through varying levels of excitement – and thus betting activity. To capture this dependence of the stakes on the not directly observable market activity level, we employ state-space models (SSMs) to model stakes placed on a team at a given point in a match, conditional on the underlying market activity level. This

approach is analogous to stochastic volatility modelling in financial markets, where the share returns are assumed to be driven by the underlying latent nervousness of the market (see, e.g. Taylor, 1994; Abanto-Valle et al., 2017), and has already successfully been transferred to the context of betting markets (Michels et al., 2023; Ötting et al., 2024). In the following, we outline the modelling framework and describe how to conduct statistical inference for this model. We then report results for models of increasing complexity and proceed to conduct outlier detection based on the final model.

4.1 Modelling framework and parameter estimation

An SSM comprises an observed state-dependent process $\{y_t\}$ – these will be the sequences of stakes placed – and an unobserved state process $\{s_t\}$, a proxy for the overall market activity level. We are interested in the absolute stakes y_{ijt} placed on team i in match j and minute t. For ease of notation, we will omit the subscripts i and j in the following. Given the strong skewness of the stakes, we consider the logarithm of these stakes to facilitate the modelling. We then use a hurdle model, with a point mass for the case where no stakes are placed in a given minute, i.e.

$$\Pr(y_t = 0) = \pi_t, \quad \text{hence} \quad \Pr(y_t > 0) = 1 - \pi_t,$$

fixing $\pi_t = 1$ for all minutes where the market is closed. For the log-transformed positive stakes, we assume

$$\log(y_t) \mid y_t > 0 \;=\; \mu_t + \sigma\sqrt{\tfrac{\nu-2}{\nu}}\,\xi_t, \qquad \xi_t \sim t_\nu,$$

i.e. a location-scale Student's t-distribution with $\nu > 2$ degrees of freedom, parameterised such that $\mathrm{Var}\big(\log(y_t) \mid y_t > 0\big) = \sigma^2$. In other words, the distribution assumed for y_t is a two-component mixture of a point mass on zero with a location-scale t-distribution, the latter on the log scale. Compared to the normal distribution, the t-distribution places more probability mass both close to its centre and in the tails, with the normal distribution as the limiting case for $\nu \to \infty$. An analysis of (pseudo-)residuals of an initial model specification with normally distributed errors indicated the necessity to account for such heavy tails (for details, see Section 4.5). The standardisation ensures that σ retains its interpretation as the standard deviation of the log-transformed stakes. The mean of the log-transformed stakes is modelled as $\mu_t = \eta_t + s_t$, with a linear predictor $\eta_t = \mathbf{x}_t^{\prime(1)}\boldsymbol{\beta}$ comprising the relevant covariates

and s_t the latent market activity level. To capture the time-varying levels of excitement during a match, the latter is assumed to be a covariate-driven AR(1) process,

$$s_t = \phi s_{t-1} + \mathbf{x}_t'^{(2)}\boldsymbol{\omega} + \sigma_s \epsilon_t,$$

where $|\phi| < 1$ is the persistence parameter, $\epsilon_t \sim \mathcal{N}(0,1)$, $\sigma_s > 0$, and covariates affecting the state process $\mathbf{x}_t'^{(2)}$. The dependence structure of the SSM is displayed in Figure 2.

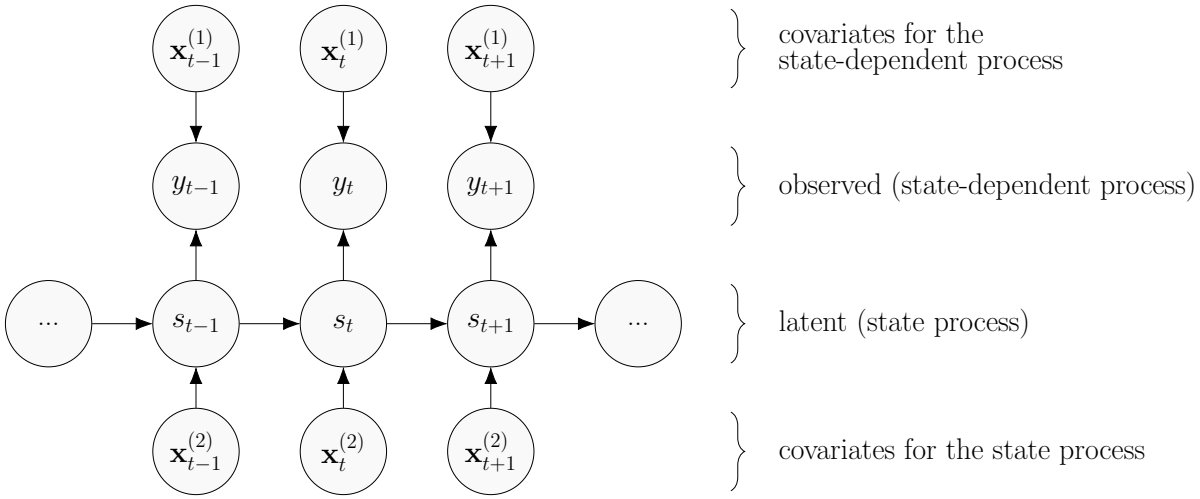


Figure 2: Dependence structure of the SSM with latent state s_t satisfying the Markov property, covariates $\mathbf{x}_t^{(1)}$ affecting the state-dependent process, covariates $\mathbf{x}_t^{(2)}$ affecting the state process, and response variable y_t giving the log-transformed stakes.

The model is estimated via maximum likelihood using the techniques first introduced by Kitagawa (1987), where the state space is very finely discretised to address the otherwise analytically intractable likelihood. Integrating over all possible state sequences and exploiting the dependence structure shown in Figure 2 (cf. Zucchini et al., 2017), the exact likelihood is obtained as follows:

$$\mathcal{L}_T(\boldsymbol{\Theta}) = \int \cdots \int f(s_1) f(y_1|s_1, \mathbf{x}_1^{(1)}) \prod_{t=2}^{T} f(s_t|s_{t-1}, \mathbf{x}_t^{(2)}) f(y_t|s_t, \mathbf{x}_t^{(1)}) ds_T \ldots ds_1, \qquad (1)$$

with the parameter vector Θ comprising all regression coefficients as well as the error variances, the persistence parameter of the AR(1) process, and the degrees of freedom ν. To very closely approximate this likelihood, we then apply numerical integration via a simple midpoint rule, finely discretising the state space into m intervals $B_i = (b_{i-1}, b_i)$ with length $h = (b_m - b_0)/m$

and midpoints $b_i^* = 0.5(b_i + b_{i-1})$, with a large m (typically around 100; cf. Langrock et al., 2012; Mews et al., 2024; Winkelmann and Michels, 2026). This yields the approximation

$$\mathcal{L}_T(\boldsymbol{\Theta}) \approx h^T \sum_{i_1=1}^{m} \cdots \sum_{i_T=1}^{m} f(b_{i_1}^*) f(y_1 | b_{i_1}^*, \mathbf{x}_1^{(1)}) \prod_{t=2}^{T} f(b_{i_t}^* | b_{i_{t-1}}^*, \mathbf{x}_t^{(2)}) f(y_t | b_{i_t}^*, \mathbf{x}_t^{(1)}). \tag{2}$$

The numerical integration corresponds to a close approximation of the SSM by a hidden Markov model (HMM) with a large number of states m and highly structured state transition dynamics. The computational cost of evaluating the likelihood of the approximating HMM can therefore be dramatically reduced, from the order $\mathcal{O}(Tm^T)$ to the order $\mathcal{O}(Tm^2)$, by making use of the recursive forward algorithm, a standard HMM tool (Zucchini et al., 2017, Chapter 11). For this, we summarise the approximate probabilities of switching from interval i to interval j, $\gamma_{ij} = hf(s_t = b_j^* | s_{t-1} = b_i^*, \mathbf{x}_t^{(2)})$, $i, j = 1, \ldots, m$, in the transition matrix $\boldsymbol{\Gamma}$. This matrix is fully determined by the parameters ϕ and σ_s as well as the regression coefficients $\boldsymbol{\omega}$. We further define the $m \times m$ diagonal matrix $\mathbf{P}(y_t)$ with the state-dependent densities $f(y_t | s_t = b_i, \mathbf{x}_t^{(1)})$, $i = 1, \ldots, m$, from the above hurdle model on the diagonal. Application of the forward algorithm then allows us to rewrite the (approximate) likelihood (2) as the matrix product

$$\mathcal{L}_T(\boldsymbol{\Theta}) \approx \boldsymbol{\delta} \mathbf{P}(y_1) \boldsymbol{\Gamma} \mathbf{P}(y_2) \ldots \boldsymbol{\Gamma} \mathbf{P}(y_T) \mathbf{1},$$

with $\mathbf{1} = (1, \ldots, 1)' \in \mathbb{R}^m$ and $\boldsymbol{\delta}$ the initial distribution of the AR(1) process.

For the given longitudinal dataset, we assume independence across matches and teams, considering each match from the perspective of each team as a single time series (match-team-series). Consequently, the total likelihood is calculated as the product of individual likelihoods of the above form. To obtain parameter estimates, we numerically maximise the likelihood in Python via the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, a quasi-Newton method, subject to the usual technical details arising in effectively every numerical optimisation (Zucchini et al., 2017).

4.2 A baseline model

As a benchmark for the more complex models below, we initially consider a model without any of the covariates described in Section 3. The parameter estimates with their 95% confidence

intervals for this baseline model are given in Table 2. The estimated persistence of $\hat{\phi} = 0.992$ in the latent AR(1) process indicates strong serial correlation in the market activity level. Further justification for the use of the SSM framework is provided by the Akaike Information Criterion (AIC), which shows a dramatic improvement of the fit compared to a simple hurdle model without an underlying state process ($\Delta\text{AIC} = 173,302$). Furthermore, the usage of the Student's t-distribution is verified by $\hat{\nu} = 4.123$, indicating substantially heavier tails than under normality, and an improved AIC ($\Delta\text{AIC} = 12,549$ compared with a normal distribution). The probability of no stakes being placed in a given minute with an open market is $\hat{\pi} = 0.094$ and thus identical to the empirical frequency of 0.094.

parameter	estimate	95% confidence interval
ϕ	0.992	[0.992; 0.992]
σ_s	0.216	[0.215; 0.217]
β_0	-0.740	[-0.747; -0.734]
σ	0.953	[0.950; 0.955]
ν	4.123	[4.109; 4.136]
π	0.094	[0.093; 0.094]

Table 2: Parameter estimates and 95% confidence intervals for the baseline model, where $\mu_t = \beta_0 + s_t$ and s_t is a simple AR(1) process.

4.3 Including covariates in the state-dependent process

Accurately identifying outliers for fraud detection requires the comparison of the actual stakes placed to their probability distribution given the current market situation. The latter needs to take into account not only the general dynamics of the in-play betting market as in the baseline model above, but also team- and match-specific variables and events that may elevate or reduce expected betting activity. We thus extend our model based on suggestions from the literature and insights from our descriptive analysis. Ötting et al. (2024) show that team strength affects bettors' tendency to place stakes. In our data, the implied winning probability correlates positively with both the expected goal difference (0.49) and the average stakes placed on the team (0.09). Consistent with this pattern, favourites attract larger stakes, with an average stake of 1.98 compared with 1.34 for underdogs. Since simple measures of team strength such as Elo ratings may not fully capture the overall sentiment towards a team – which however has been shown to have an effect on betting behaviour (Forrest and

Simmons, 2008) – we consider the average stakes per minute placed on the team across the whole observation period, denoted as *stake_avg_team*, as a measure capturing information on both team strength and sentiment. The inclusion of this covariate ensures that under our model, the actual stakes placed are compared to their overall level (and distribution) for the specific team under consideration. This is crucial, given the strong variation in the stakes placed across teams as shown in Figure B1 in Appendix B.

To further account for bettors' tendency to place higher stakes on teams designated as favourites, we include the pre-match implied probability for the team considered *improbteam_start*. In contrast to the variable *stake_avg_team*, this measure captures the team's relative strength by explicitly accounting for the strength of the opponent. According to Winkelmann and Deutscher (2026), a red card issued to the opponent (team) increases (decreases) average stakes placed on the team, leading us to include the variables *redcard_team* and *redcard_opponent*, denoting the number of red cards issued to the team and opponent, respectively. This is confirmed in our data by a positive correlation between red cards issued to the away team and stakes placed on the home team (0.17), and vice versa (0.15).

To also account for in-play dynamics, we further include the variable *scorediff_team* denoting the current score difference from the perspective of the team under consideration. Finally, the descriptive analysis shows that stakes vary with the minute. Specifically, Figure 1 in Section 3 shows a ∪-shaped relation between the minute and the stakes placed outside halftime (where betting activity is also elevated). We thus include a quadratic effect of the *minute* (in line with Winkelmann and Deutscher, 2026) as well as a binary variable *halftime*, taking the value 1 if the match is in its halftime. We standardise the variable *minute* by subtracting its mean and dividing by its standard deviation to ensure numerical stability. In preliminary checks, we also tested whether stakes on home teams are higher than those on away teams, beyond what is already captured in pre-match implied probabilities, and whether the average stake placed on the opponent affects stakes on the team under consideration. However, both possible effects were deemed to be insignificant. Overall, we arrive at the following time-varying linear predictor for the mean stakes, on the log-scale:

$$\mu_t = \beta_0 + \beta_1 \text{stake_avg_team}_t + \beta_2 \text{ improbteam_start}_t + \beta_3 \text{ redcard_team}_t + \beta_4 \text{ redcard_opponent}_t + \beta_5 \text{ scorediff_team}_t + \beta_6 \text{ minute}_t + \beta_7 \text{ minute}_t^2 + \beta_8 \text{ halftime}_t + s_t$$

Additionally, our descriptive statistics show that the case of no stakes being placed despite an open market arises more frequently when the match is already decided – as indicated by a high value of the Gini coefficient – in particular as the match progresses (cf. Ötting et al., 2024). We therefore model the probability that zero stakes are placed as

$$\pi_t = \text{logit}^{-1}\left(\alpha_0 + \alpha_1\ \text{Gini}_t + \alpha_2\ \text{Gini}_t^2 + \alpha_3\ \text{minute}_t + \alpha_4\ \text{minute}_t \cdot\ \text{Gini}_t + \alpha_5\ \text{minute}_t \cdot\ \text{Gini}_t^2\right)$$

variable	parameter	estimate	95% confidence interval
	ϕ	0.985	[0.985; 0.985]
	σ_s	0.206	[0.203; 0.208]
intercept	β_0	-2.310	[-2.377; -2.242]
stake_avg_team	β_1	0.249	[0.223; 0.276]
improbteam_start	β_2	2.582	[2.525; 2.638]
redcard_team	β_3	-0.369	[-0.454; -0.285]
redcard_opponent	β_4	1.208	[1.139; 1.277]
scorediff_team	β_5	0.382	[0.359; 0.406]
minute	β_6	0.043	[0.025; 0.061]
minute^2	β_7	0.028	[0.007; 0.048]
halftime	β_8	0.197	[0.178; 0.216]
	σ	0.953	[0.948; 0.958]
	ν	4.087	[4.049; 4.125]
intercept	α_0	-4.818	[-4.873; -4.763]
Gini	α_1	3.760	[3.674; 3.846]
Gini^2	α_2	0.965	[0.932; 0.998]
minute	α_3	-0.980	[-1.020; -0.940]
minute $\cdot$ Gini	α_4	1.764	[1.704; 1.825]
minute $\cdot$ Gini^2	α_5	-0.501	[-0.534; -0.467]

Table 3: Parameter estimates and 95% confidence intervals for the model including covariates in the state-dependent process.

Table 3 reports the estimated parameter values together with their 95% confidence intervals. The fitted state process confirms strong serial correlation in the placement of betting stakes. We further see confirmation of the bettors' tendency to place higher stakes on the pre-match favourite, as indicated by the positive value of $\hat{\beta}_2$. Consistent with the findings of Winkelmann and Deutscher (2026) for the German Bundesliga, bettors tend to place higher (lower) stakes when the opponent (team) has received a red card, with the effect of a red card issued to the opponent being stronger. In addition, it is confirmed that bettors place higher stakes when the team's in-match winning chances increase (cf. $\hat{\beta}_5$). Finally, with respect to the minute of the match considered, we observe the highest stakes at the beginning and the end of a match, with average stakes additionally elevated during the halftime break by approximately 21.8%.

4.4 Including covariates in the state process

Not only the observed betting stakes, but also their underlying market activity level may respond to covariates. Michels et al. (2023) already showed that betting activity is strongly affected by in-match dynamics. We thus extend the AR(1) process used for the market activity level to an ARX(1) process, such that the state process of the corresponding SSM is affected by covariates. As scoring opportunities may affect the underlying betting dynamics, we include the difference in expected goals *xg_diff* as a covariate in the state process. Furthermore, the persistence in betting activity may differ during halftime (see Figure 1 in Section 3), such that we include the variable *halftime* also in the state process. Additionally, Ötting et al. (2024) suggest that bettors increase their activity after a goal has been scored, with stronger effects following surprising goals. This effect is expected to fade within a few minutes after the goal. To capture this effect, we use indicator variables $I_w \in \{0, 1\}$, taking the value 1 if the most recent goal was surprising ($w = 1$), slightly surprising ($w = 2$), or unsurprising ($w = 3$) – and all $I_w = 0$ if no goal has been scored yet. The variable $\text{goal_team}_t^{(w)}$ measures the time in minutes since that last goal (set to the value 1 if no such goal has occurred yet). This yields the following specification of the state process:

$$s_t = \phi s_{t-1} + \sum_{w=1}^{3} \omega_w I_w / \text{goal_team}_t^{(w)} + \omega_4 \ \text{xg_diff}_t + \omega_5 \ \text{halftime}_t + \sigma_s \epsilon_t.$$

Table 4 provides the results for this full model. Most estimated coefficients associated with the covariates affecting the state-dependent stakes differ only slightly from those presented in Table 3. In contrast to the German Bundesliga, where *Valuing Actions by Estimating Probabilities* was found to affect the underlying (latent) betting activity (Michels et al., 2023), in the Serie B we do not find the match dynamics, measured by the difference in expected goals, to significantly affect the state of the market (cf. $\hat{\omega}_4$). Betting activity during halftime is slightly elevated. The results on the reactions to goals are similar to those reported for the German Bundesliga (cf. Ötting et al., 2024): While surprising goals, i.e. those scored by pre-match underdogs, significantly increase the state and hence the market activity level, unsurprising goals, i.e. those scored by pre-match favourites, significantly decrease the market activity.

variable	parameter	estimate	95% confidence interval
	ϕ	0.985	[0.984; 0.986]
	σ_s	0.197	[0.193; 0.201]
intercept	β_0	-2.572	[-2.636; -2.508]
stake_avg_team	β_1	0.226	[0.206; 0.245]
improbteam_start	β_2	3.420	[3.297; 3.542]
redcard_team	β_3	-0.363	[-0.451; -0.275]
redcard_opponent	β_4	1.178	[1.107; 1.248]
scorediff_team	β_5	0.420	[0.397; 0.443]
minute	β_6	0.052	[0.031; 0.073]
minute^2	β_7	0.030	[0.017; 0.043]
halftime	β_8	0.197	[0.179; 0.215]
surprising	ω_1	0.274	[0.245; 0.303]
slightly_surprising	ω_2	-0.037	[-0.053; -0.021]
unsurprising	ω_3	-0.394	[-0.422; -0.366]
xg_diff	ω_4	0.001	[-0.000; 0.003]
halftime	ω_5	0.004	[0.001; 0.007]
	σ	0.955	[0.948; 0.962]
	ν	4.103	[4.007; 4.203]
intercept	α_0	-4.819	[-4.864; -4.773]
Gini	α_1	3.761	[3.559; 3.963]
Gini^2	α_2	0.964	[0.745; 1.184]
minute	α_3	-0.980	[-1.030; -0.930]
minute $\cdot$ Gini	α_4	1.765	[1.492; 2.039]
minute $\cdot$ Gini^2	α_5	-0.502	[-0.776; -0.227]

Table 4: Parameter estimates and 95% confidence intervals for the model including covariates in the state process and in the state-dependent process.

4.5 Model diagnostics

Performing outlier detection to reliably flag suspicious observations requires an accurate predictive model of betting stakes. We therefore conduct several diagnostic checks to assess the model's ability to comprehensively capture the market's dynamics. We restrict the model diagnostics to observations for which the market was open. Our hurdle model consists of two components. First, it models the probability of observing a zero stake, π, as a function of the match minute and the Gini coefficient. Second, conditional on a positive stake, the model uses the SSM described above to explain log-transformed stake sizes. We begin by analysing the latter, since the case of positive stakes is more relevant to our study's aim.

To evaluate the SSM, we recursively compute the model- and data-implied one-step-ahead state probabilities, $p_{t|t-1}(i) = \Pr(s_t = i \mid y_{1:t-1}, x_{1:t})$, with the states i corresponding to the intervals B_i from the discretisation of the market activity level described in Section 4.1. From

this, we obtain the one-step-ahead predictive distribution as

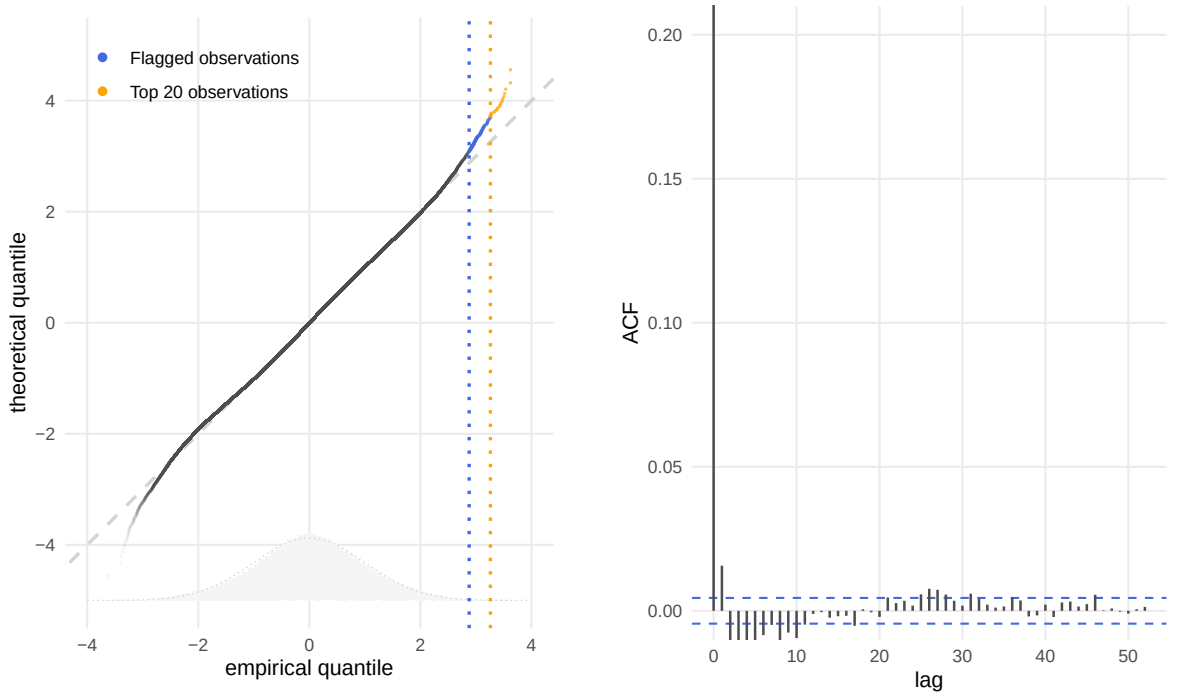
$$f\big(\log(y_t) \mid y_{1:t-1}, x_{1:t}\big) = \sum_{i=1}^{m} p_{t|t-1}(i)\, f\big(\log(y_t) \mid s_t = i, x_t\big).$$

Let $F_t\big((\log(y)\big) = \Pr\big(\log(Y_t) \leq y \mid y_{1:t-1}, x_{1:t}\big)$ denote the corresponding one-step-ahead predictive CDF (a mixture of Student-t distributions). We compute probability integral transform (PIT) values $u_t = F_t(\log(y_t))$ and form quantile pseudo-residuals $r_t = \Phi^{-1}(u_t)$. Under a correct specification – i.e. in case the statistical model closely mimics the real data-generating process – the pseudo-residuals should be approximately i.i.d. standard normally distributed. The most extreme values of r_t are the ones which are the least likely to have occurred under the given model. For the subsequent outlier detection (see Section 5), we are mainly interested in particularly large positive values, as these may signal suspiciously high betting activity.

Figure 3a shows a Q–Q plot comparing theoretical and empirical quantiles (pseudo-residuals), which indicates an overall reasonable model fit. Notable deviations from the standard normal distribution – which would be expected under a perfect model specification – arise only in the extreme tails, specifically concerning the smallest as well as the largest 0.5% of all observations. For our purposes, these deviations are relevant only in the upper tail, where observed betting activity exceeds what the model predicts. We therefore identify the 195 largest observations as those exceeding the 99.8% quantile of the predictive distribution, with particular focus on the top 20. We will examine these outliers more closely in the outlier analysis (see Section 5). Figure 3b indicates very weak serial dependence in the pseudo-residuals, suggesting that the SSM does a reasonable job at capturing the time dependence between consecutive stakes.

As a check of our model's predictive performance, we further simulate the one-step-ahead predictive distribution of (log-)stakes and compare it with the empirical series. For each time point t, we consider the one-step-ahead state probabilities $p_{t|t-1}(i)$ and sample 10,000 interval midpoints from this state distribution, conditional on which we compute the predicted time-varying mean $\hat{\mu}$ using the estimated model parameters (see Table 4). We then sample innovations $t_{\hat{\nu}}$ from the fitted Student-t distribution and obtain draws for predicted log-stakes as

$$\log(y_t^{\text{pred}}) = \hat{\mu}_t + \hat{\sigma}\sqrt{\frac{\hat{\nu}-2}{\hat{\nu}}}\,\xi_t.$$



(a) Q–Q plot of pseudo-residuals associated with positive betting stakes. The dashed grey 45-degree line indicates the expected alignment between empirical and theoretical quantiles if the pseudo-residuals were standard-normally distributed. The empirical distribution of the pseudo-residuals is shown along the horizontal axis, with the N(0,1) density overlaid as a dotted line. Points corresponding to stakes exceeding the 99.8% quantile (flagged observations) and the 99.94% quantile (top 20 observations) of their respective one-step-ahead predictive distributions are coloured.

(b) Autocorrelation function (ACF) of the pseudo-residuals for positive betting stakes in our dataset. Dashed lines indicate the 95% confidence interval under independence over time. The vertical axis is truncated at 0.2. The value at lag 0 is 1 by construction.

Figure 3: Model diagnostics for the state-space model (SSM) based on pseudo-residuals.

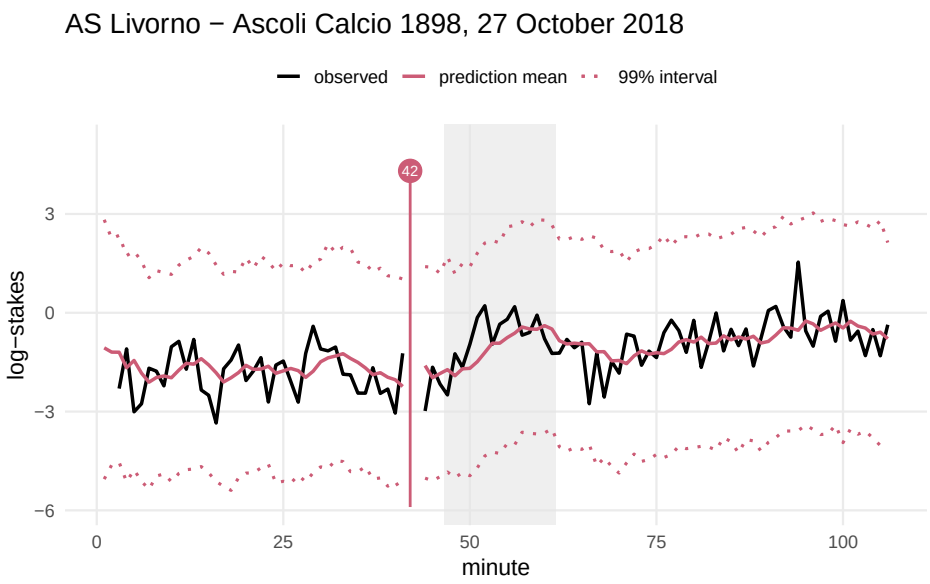


Figure 4: One-step-ahead predictions together with corresponding intervals covering the central 99% of all predictions for log-stakes placed on AS Livorno in their home match against Ascoli Calcio on 27 October 2018. The vertical line indicates the 1-0 goal scored by AS Livorno in minute 42. The grey-shaded area corresponds to halftime.

Figure 4 shows the predicted mean and the 99% predictive interval for the log-stakes placed on AS Livorno in their home match against Ascoli Calcio on 27 October 2018. During the first 40 minutes, average stakes were slightly below the predicted mean. After the opening goal in minute 42, the market was closed for a short period before observed stakes increased during halftime and remained elevated in the second half, closely aligning with the prediction, albeit with some fluctuation. Despite this volatility, all observations clearly lie within the 99% predictive interval, providing no evidence of suspicious behaviour in this example match.

In the second step, we consider the binary component used to predict the probability of observing zero stakes in an open market. Figure 5 shows a calibration plot comparing mean predicted probabilities (x-axis) with the empirical frequencies (y-axis), together with their 95% confidence intervals. Predicted probabilities are grouped into 2.5-percentage-point bins, with point sizes proportional to the number of observations per bin (for better readability, we exclude those 0.07% of observations where the predicted probability of observing zero stakes exceeds 0.5). Most observations fall into bins where both predicted and empirical zero frequency are close to zero. For empirical frequencies between 0.2 and 0.3, the model slightly underestimates the probability of zero stakes, whereas it tends to slightly overestimate it for frequencies above 0.4. Since our primary objective is to identify suspiciously large

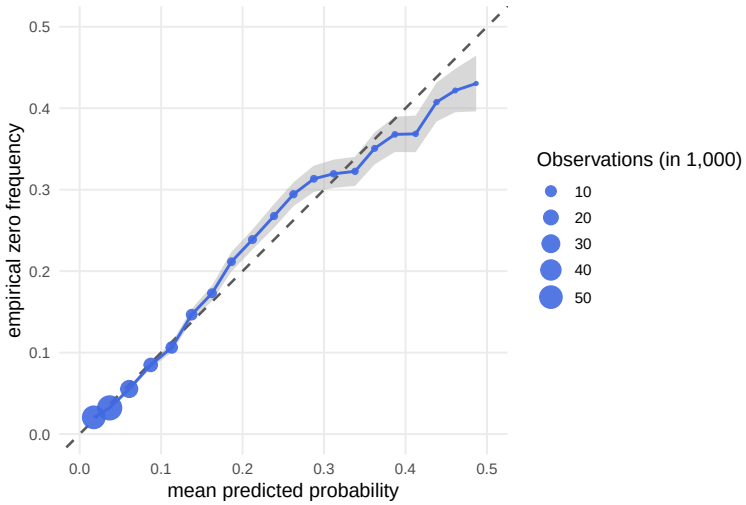


Figure 5: Calibration plot of theoretical and empirical probability to observe zero stakes despite an open market, including 95% confidence intervals. Point sizes indicate the number of underlying observations in the respective bin.

observations, the zero-stake component is of minor importance, and its fit is sufficiently precise for our purpose. This is also reflected in the scoring rules: the Brier score is 0.074 (an improvement over a constant-rate benchmark of approximately 12.5%) and the log loss is 0.260 (an improvement of approximately 16.4%).

5 Outlier analysis

To address the paper's central questions – namely, how to safeguard integrity in sports by detecting suspicious betting patterns that may be consistent with match-fixing – we closely examine the pseudo-residuals r_t introduced in Section 4.5. For each match, observation, and outcome, these pseudo-residuals capture the discrepancy between the stakes predicted by our model and the stakes observed in the market. This is equivalent to flagging observations in which realised stakes exceed a given threshold – effectively a quantile – in the model's predictive distribution, thereby enabling bookmakers to react in-play. Our primary focus is on extreme positive deviations, as corresponding outliers indicate unusually large volumes placed on a specific outcome relative to what would be expected given the available information, which may therefore warrant further investigations into potentially irregular activity. For

the following analysis, we consider all stakes that exceed the 99.8% quantile, yielding 195 observations (0.09% of all observations in which the market was open; for the relation between the chosen threshold and the number of flagged observations, see Figure E1 in Appendix E). These 195 observations are associated with 175 distinct time series of individual teams in single matches (match-team-series, cf. Section 4.1). The flagged series involve 174 matches; there is thus only a single match in which betting on both teams is flagged, and only 16 series feature two or more unusually large stakes (maximum: four).

To identify potentially recurring patterns, we compare the 195 flagged observations with all observations for which placing bets was possible (i.e. when the market was open). Since no confirmed cases of match-fixing are documented for the matches in our sample, this comparison serves to characterise the situations in which such large deviations occur, rather than to establish them as indicators of manipulation. Flagged observations disproportionately often involve bets on away teams (64.6%) and on teams that are one goal behind (25.1% compared with 18.7% in the full dataset), while teams one goal ahead are underrepresented (4.1% compared with 18.7%). Unsurprisingly, this coincides with a lower mean implied win probability (23.6% compared with 35.2% overall); in 76.4% of cases, the implied win probability is lower than at kick-off. The mean stake on the flagged outcome is 23.25, approximately 19 times the overall mean of 1.19.

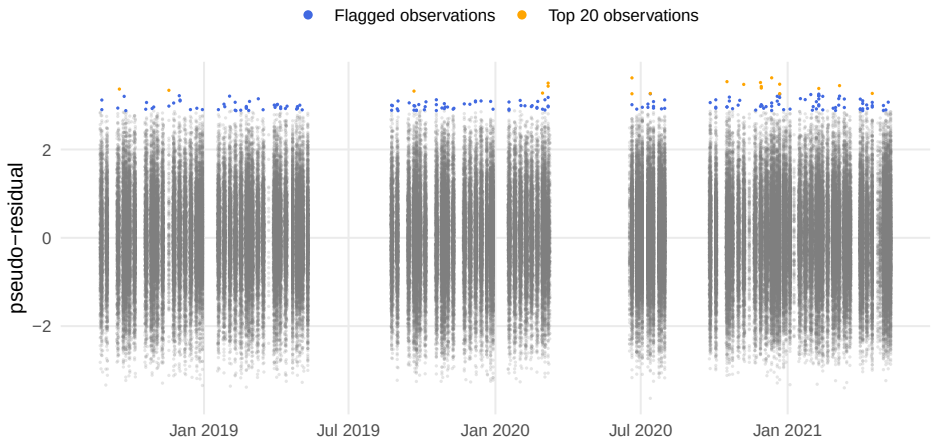


Figure 6: Pseudo-residuals over calendar time. Observations with stakes exceeding the 99.8% and 99.94% quantile of their respective one-step-ahead predictive distributions are coloured.

Figure 6 plots pseudo-residuals against calendar date. Flagged observations occur across

all three seasons in our sample and throughout all periods of the season. However, their frequency appears to increase slightly in the 2020/21 season (40% of the flagged observations stem from this season). With respect to the minute of the match, Figure 7 shows that flagged stakes are more common in the first half (55.4%) and up to the end of halftime (68.7%). This exceeds the corresponding share of overall betting activity (39.3% during the first half; 53.3% until the end of halftime), as reflected in the average team stakes in Figure 1. Relating flagged observations to the number of matches played, Juve Stabia (7 of 38 matches), Virtus Entella (13 of 76), and Perugia (11 of 74) are flagged most frequently. By contrast, Palermo and Hellas Verona, which stand out in terms of average stakes per minute (cf. Figure B1 in Appendix B), are flagged only once and not at all, respectively.

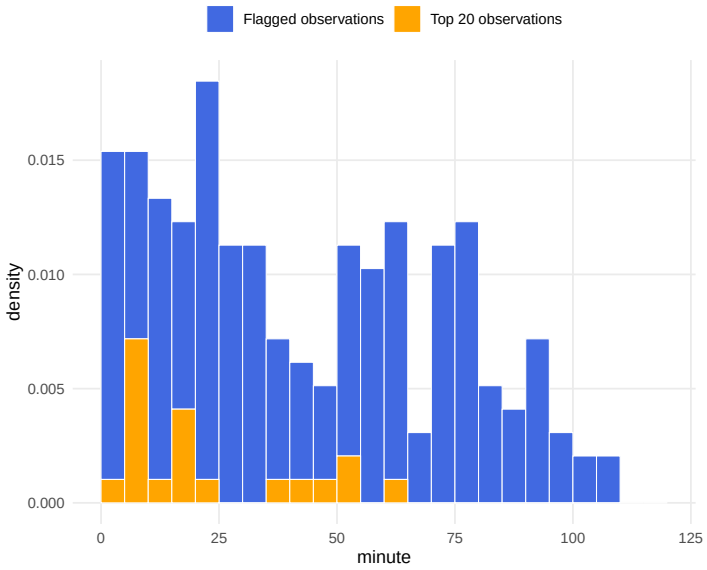


Figure 7: Distribution of flagged observations over the course of the match

Turning to outcomes, among the 175 flagged series, 69 matches ended in a home win (39.4%; 42.5% in the full dataset), 56 in a draw (32.0%; 31.3%), and 50 in an away win (28.6%; 26.2%). Accordingly, the empirical distributions of flagged match-team-series and of all matches are very similar. Within three minutes after a flagged observation, the backed team scores in 2.6% of cases, while the opposing team scores in 6.2% (across all observations, a goal occurs within the subsequent three minutes in 3.1% of the cases). At a five-minute horizon, the corresponding figures are 3.6% (backed team scores) and 8.2% (opponent scores), compared with 5.2% across all observations. The backed team wins in only 30 of the 175

flagged match-team-series (17.1%). Figure 8 reports the absolute return per € wagered in individual bets (right y-axis) and the average relative accumulated return on investment when placing the same stake on the bets with the top-x pseudo-residual values (left y-axis). The high odds on some bets lead to a positive average cumulative return at a few points among the first 25 observations. Nevertheless, with only a few winning bets among all flagged observations, it is unsurprising that the average return on investment of -40% is clearly negative and falls below the average bookmaker margin of approximately 8.7% across all observations where betting was possible.

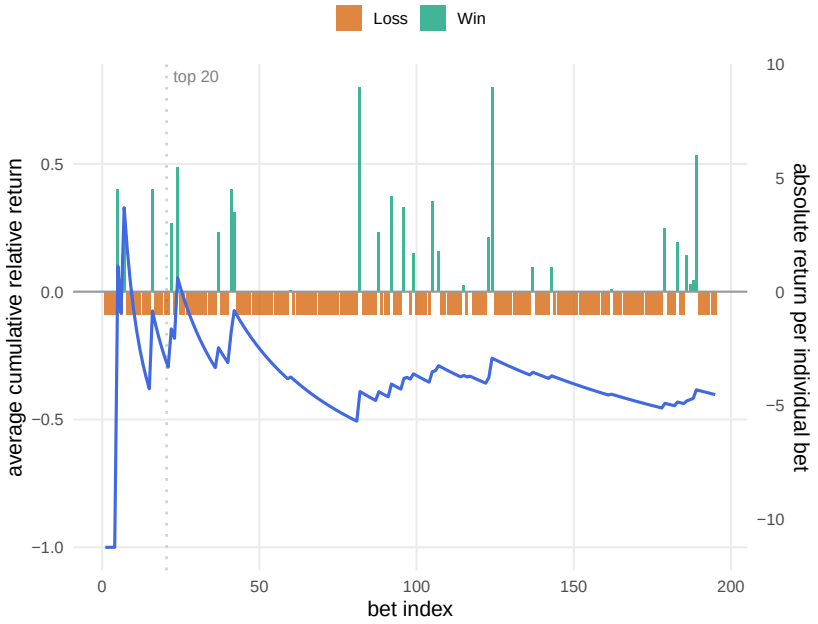


Figure 8: Absolute return per € wagered in individual bets (right y-axis) and average cumulative relative return on investment for the bets with the x highest pseudo-residual values in the dataset, when placing the same stakes on each bet (blue line; left y-axis). For example, the average cumulative relative return on investment for the 50 bets with the highest pseudo-residual values is -22.2%. The grey-dotted line indicates the top 20 flagged observations.

Finally, we examine the 20 observations with the highest pseudo-residual values (those exceeding the 99.94% quantile of the predictive distribution) in more detail. Table 5 reports the corresponding match and date, the predicted and actual stakes, and key match characteristics. Seventeen out of 20 bets are on away teams, and all are placed before or during halftime at a score of either 0-0 or 1-0. Fifteen out of 20 observations occur between 07 March 2020 and 07 March 2021 (cf. Figure 6). To check whether seasonal differences in betting activity explain this pattern, we repeat the analysis using season-specific rather than overall average

team stakes in Equation 4.3. However, the 20 identified matches remain the same. Moreover, all bets are on the underdog or very slight favourite: the implied win probability is below 25% in 15 out of 20 cases. All realised amounts clearly exceed the 99% quantile of the predicted distribution by a wide margin (up to a factor of more than 160). Three out of the 20 bets are won, implying a return on investment of -26% under constant wagering across matches. This is again considerably less than the average bookmaker's margin. Regarding match outcomes, nine out of the 20 matches with the most extreme observations ended in a draw.

home team	away team	date	minute	bet on team	current score	final score	predicted stake (99% quantile)	actual stake	odds	implied win probability (%)	bet won
AS Livorno	FC Crotone	17.09.2018	6	home	0-0	0-1	1.43	48.99	2.9	31.50	no
Spezia Calcio	**Benevento Calcio**	18.11.2018	21	away	1-0	3-1	1.82	51.88	7.0	13.04	no
AS Livorno	**Pordenone Calcio**	21.09.2019	54	away	1-0	2-1	0.54	12.95	13.0	7.13	no
AC Chievo Verona	**AS Livorno**	29.02.2020	36	away	0-0	0-1	2.34	51.52	5.5	16.64	yes
Frosinone Calcio	**US Cremonese**	07.03.2020	10	away	0-0	0-2	0.62	51.33	5.5	16.80	yes
FC Empoli	**Trapani Calcio**	07.03.2020	44	away	0-0	1-1	1.26	64.93	5.5	16.41	no
US Salernitana 1919	**Pisa SC**	20.06.2020	7	away	0-0	1-1	0.62	103.05	3.8	24.00	no
AS Livorno	AS Cittadella	20.06.2020	62	home	0-0	0-2	5.21	102.8	4.5	20.35	no
Cosenza Calcio	**Perugia Calcio**	13.07.2020	8	away	0-0	2-1	5.29	103.28	2.6	34.97	no
Virtus Entella	**US Reggina 1914**	17.10.2020	53	away	0-0	1-1	1.15	103.10	3.8	23.49	no
US Cremonese	**Vicenza Virtus**	07.11.2020	7	away	0-0	0-1	1.17	77.02	3.8	24.00	yes
AC Monza	**US Reggina 1914**	28.11.2020	50	away	0-0	1-0	1.13	77.55	7.5	12.24	no
Cosenza Calcio	US Salernitana 1919	29.11.2020	2	home	0-0	0-1	2.54	102.71	2.7	34.94	no
AC Reggiana	**US Cremonese**	29.11.2020	19	away	1-0	1-1	1.55	52.72	5.0	18.48	no
US Cremonese	**Ascoli Calcio 1898**	12.12.2020	16	away	0-0	3-3	0.33	51.41	5.0	18.18	no
FC Empoli	**AC Reggiana**	22.12.2020	18	away	0-0	0-0	0.75	51.35	6.5	14.07	no
Cosenza Calcio	**Venezia**	22.12.2020	6	away	0-0	0-0	3.75	77.21	2.7	34.15	no
US Lecce	**Brescia Calcio**	09.02.2021	13	away	0-0	2-2	0.47	20.53	5.0	18.14	no
FC Empoli	**AS Cittadella**	07.03.2021	9	away	0-0	1-1	0.38	20.23	4.3	21.18	no
AC Monza	**US Cremonese**	17.04.2021	16	away	0-0	2-1	0.72	15.49	3.4	27.18	no

Table 5: Matches, characteristics of the bet, and predicted as well as empirical stakes for the 20 observations with the highest pseudo-residual value. Teams backed are highlighted.

The teams AS Livorno and US Cremonese each appear in three cases as the team backed. However, only one bet on each was successful. In addition, the match between US Lecce and Brescia Calcio on 09 February 2021 is of particular interest: four pseudo-residual values for bets on Brescia Calcio exceed our 99.8% quantile threshold, and a fifth is close. In this match, US Lecce (the home team) scored in the 48th and 56th minutes of the match (note that these differ from the values of *minute* in our dataset, which also counts the time elapsed during half-time). The away team, Brescia Calcio, scored in the 75th minute of the match and equalised late in the injury time added to the second half. Our analysis of pseudo-residuals indicates extreme betting activity on Brescia Calcio in the 13th, 23rd, 28th, 39th, and 43rd minutes. However, less than 2% of the stakes on this team were placed after US Lecce scored the opening goal at the very beginning of the second half. Figure 9 shows the predicted stakes, including 99% confidence intervals, together with the actual stakes until the end of the

halftime. The outlying betting stakes are highlighted. Although Brescia Calcio eventually equalised, all bets on a win for this team were lost.

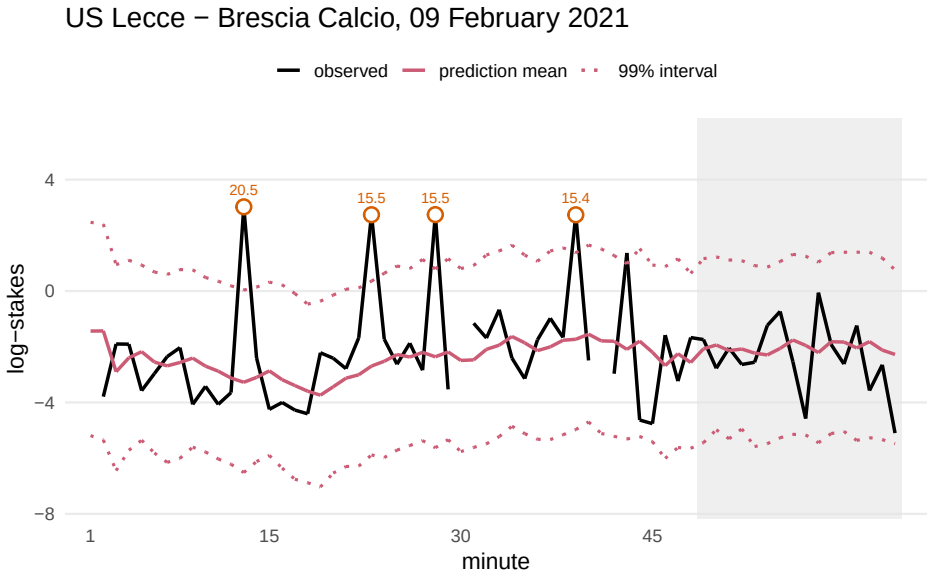


Figure 9: One-step-ahead predictions together with corresponding intervals covering the central 99% of all predictions for log-stakes placed on Brescia Calcio in their away match against US Lecce on 09 February 2021. The figure shows only the period until the end of the halftime, as only 2% of total stakes were placed during the second half, with numerous observations in which no stakes were placed. Unusual betting activity (stakes exceeding the 99.8% quantile of the predictive distribution) is highlighted, along with the actual betting stakes on the original scale. The grey-shaded area corresponds to halftime.

6 Conclusion

We developed a state-space modelling framework for high-frequency betting stakes in the Italian Serie B, using data from the 2018/19–2020/21 seasons provided by a leading European bookmaker. Incorporating covariates that capture team- and match-specific characteristics, motivated by existing literature and a descriptive analysis, the model provides predictions of expected betting volumes and their distributions for every minute of each match. We find that bettors tend to place higher stakes on pre-match favourites and on teams in the lead, whereas a red card is associated with lower average stakes on the affected team and higher stakes on the opponent. The underlying (latent) level of betting activity responds significantly to goals: surprising goals increase activity, while unsurprising goals decrease it. These predictions form the basis for comparing model-implied and empirical stakes through pseudo-residuals, thereby facilitating the identification of abnormal betting activity.

We classify an observation as potentially suspicious if the realised betting stakes exceed the 99.8% quantile of our model's predictive distribution. Among these flagged bets, no clear patterns could be identified. The relative frequency of flagged observations is somewhat higher before and up to the end of halftime, and 17 of the 20 largest pseudo-residuals concentrate between February 2020 and April 2021. While bets on away teams and on teams trailing by one goal, and therefore those with lower implied win probabilities, are overrepresented, the distribution of match outcomes among flagged bets closely matches that of the full sample. In particular, we do not find evidence that teams attracting unusually large stakes have a higher probability of scoring in the subsequent minutes; if anything, the opposite holds: in these situations, the opponent turned out to be more likely to score than in the full dataset. Consistent with this, the average return on investment from wagering the same amount on all flagged bets is clearly negative and below that of a typical benchmark strategy (although a few bets did yield high positive returns). Hence, while the Italian Serie B has historically been affected by match-fixing and fraudulent betting activity, our analysis does not provide evidence that the identified anomalies reflect match-fixing in our sample. Nevertheless, our proof-of-concept approach provides a promising tool for monitoring betting activity and identifying potentially suspicious behaviour in real time, thereby demonstrating how advanced statistical modelling can support the early detection of irregular betting patterns and strengthen integrity assurance in in-play (live) sports betting markets.

There are several natural ways to extend our study. First, our data covers only three seasons of Italy's Serie B. Although the league has experienced known match-fixing cases in the past, no fixed matches are known to have occurred in our sample. We therefore lack ground-truth data against which to assess whether the flagged observations correspond to actual manipulation. If suitable data become available, future research could evaluate the approach across seasons or leagues with confirmed cases, including earlier seasons of Serie B or recent cases in the Turkish league, and measure its ability to accurately identify such cases. Second, our predictions are generated in-sample. An out-of-sample evaluation, i.e. training the model on earlier seasons and assessing it on a subsequent season, would provide a more realistic test of its performance and better reflect how a bookmaker might use the framework to monitor matches in real time. Moreover, our outlier-detection procedure assesses observations individually rather than across consecutive periods. It may therefore flag individual spikes

in betting volume but fail to detect a sustained increase in stakes spread across multiple observations. Future work could develop methods that detect such temporal patterns or take into account the actions of individual bettors, if such data were available. Finally, given the relatively small share of stakes placed on draws and the potentially greater difficulty of fixing a match to produce that outcome, our analysis is limited to bets on home and away teams and excludes bets on draws. Extending the model to include draw stakes would likely require altering the covariates and modelling assumptions.

Data availability statement: The data used for this research comprises processed and analysed datasets derived from raw data provided by a major European bookmaker. The raw data itself is proprietary and cannot be shared due to contractual confidentiality agreements.

Generative AI statement: The code was written from scratch by the authors. ChatGPT and Claude were used to assist in improving and structuring the code. The authors carefully reviewed the code and take full responsibility for its final version. In addition, during the preparation of this work, the authors used ChatGPT-5 to assist in improving language and clarity. The authors subsequently reviewed and edited the content to meet the required standards and take full responsibility for the final content of this publication.

Funding information: This work was supported by the German Research Foundation (Grant 431536450), which is gratefully acknowledged.

References

Abanto-Valle, C. A., Langrock, R., Chen, M.-H., and Cardoso, M. V. (2017). Maximum likelihood estimation for stochastic volatility in mean models with heavy-tailed distributions. *Applied Stochastic Models in Business and Industry*, 33(4):394–408.

Anfilets, S., Bezobrazov, S., Golovko, V., Sachenko, A., Komar, M., Dolny, R., Kasyanik, V., Bykovyy, P., Mikhno, E., and Osolinskyi, O. (2020). Deep multilayer neural network for predicting the winner of football matches. *International Journal of Computing*, 19(1):70–77.

Angelini, G. and De Angelis, L. (2017). PARX model for football match predictions. *Journal of Forecasting*, 36(7):795–807.

Angelini, G. and De Angelis, L. (2019). Efficiency of online football betting markets. *International Journal of Forecasting*, 35(2):712–721.

Angelini, G., De Angelis, L., and Singleton, C. (2022). Informational efficiency and behaviour within in-play prediction markets. *International Journal of Forecasting*, 38(1):282–299.

Barra, I., Hoogerheide, L., Koopman, S. J., and Lucas, A. (2017). Joint Bayesian analysis of parameters and states in nonlinear non-Gaussian state space models. *Journal of Applied Econometrics*, 32(5):1003–1026.

Borghesi, R. (2007). The late-season bias: explaining the NFL's home-underdog effect. *Applied Economics*, 39(15):1889–1903.

Breuer, M. and Forrest, D. (2018). *The Palgrave Handbook on the Economics of Manipulation in Sport*, volume 10. Springer.

Buhagiar, R., Cortis, D., and Newall, P. W. (2018). Why do some soccer bettors lose more money than others? *Journal of Behavioral and Experimental Finance*, 18:85–93.

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

Carmichael, F. and Thomas, D. (2005). Home-field effect and team performance: evidence from English Premiership football. *Journal of Sports Economics*, 6(3):264–281.

Carron, A. V., Loughhead, T. M., and Bray, S. R. (2005). The home advantage in sport competitions: Courneya and Carron's (1992) conceptual framework a decade later. *Journal of Sports Sciences*, 23(4):395–407.

Choi, D. and Hui, S. K. (2014). The role of surprise: understanding overreaction and underreaction to unanticipated events using in-play soccer betting market. *Journal of Economic Behavior & Organization*, 107:614–629.

Constantinou, A. C. and Fenton, N. E. (2013). Profiting from arbitrage and odds biases of the European football gambling market. *The Journal of Gambling Business and Economics*, 7(2):41–70.

Costa, J. (2018). The globalised network of a dirty game: match-fixing, illegal betting and transnational organised crime in Italian football. *Global Crime*, 19(2):125–145.

Croxson, K. and Reade, J. J. (2014). Information and efficiency: goal arrival in soccer betting. *The Economic Journal*, 124(575):62–91.

Deutscher, C., Dimant, E., Humphreys, B. R., and Nowak, A. (2025). Match fixing and sports betting in football: Empirical evidence from the German Bundesliga. *Public Finance Review*, 53(5):600–627.

Deutscher, C., Frick, B., and Ötting, M. (2018). Betting market inefficiencies are short-lived in German professional football. *Applied Economics*, 50(30):3240–3246.

Dixon, M. J. and Pope, P. F. (2004). The value of statistical forecasts in the UK association football betting market. *International Journal of Forecasting*, 20(4):697–711.

Duggan, M. and Levitt, S. D. (2002). Winning isn't everything: corruption in Sumo Wrestling. *American Economic Review*, 92(5):1594–1605.

Durand, R. B., Patterson, F. M., and Shank, C. A. (2021). Behavioral biases in the NFL gambling market: overreaction to news and the recency bias. *Journal of Behavioral and Experimental Finance*, 31:100522.

European Gaming & Betting Association (2025). Annual Activity Report 2025. `https://www.egba.eu/uploads/2025/03/250331-EGBA-Annual-Activity-Report-2025.pdf`.

Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25:383–417.

Feddersen, A., Humphreys, B. R., and Soebbing, B. P. (2017). Sentiment bias and asset prices: evidence from sports betting markets and social media. *Economic Inquiry*, 55(2):1119–1129.

Forrest, D. (2012). The threat to football from betting-related corruption. *International Journal of Sport Finance*, 7(2).

Forrest, D. and McHale, I. G. (2019). Using statistics to detect match fixing in sport. *IMA Journal of Management Mathematics*, 30(4):431–449.

Forrest, D., McHale, I. G., and McAuley, K. (2008). "Say It Ain't So": Betting-related malpractice in sport. *International Journal of Sport Finance*, 3(3):156.

Forrest, D. and Simmons, R. (2008). Sentiment in the betting market on Spanish football. *Applied Economics*, 40(1):119–126.

Gil, R. G. R. and Levitt, S. D. (2007). Testing the efficiency of markets in the 2002 World Cup. *The Journal of Prediction Markets*, 1(3):255–270.

Hegarty, T. and Whelan, K. (2024). Comparing two methods for testing the efficiency of sports betting markets. *Sports Economics Review*, 8:100042.

Kim, C., Park, J.-H., and Lee, J.-Y. (2024). AI-based betting anomaly detection system to ensure fairness in sports and prevent illegal gambling. *Scientific Reports*, 14(1):6470.

Kitagawa, G. (1987). Non-Gaussian state-space modeling of nonstationary time series. *Journal of the American Statistical Association*, 82(400):1032–1041.

Kocsoy, A. (2026). Determinants of betting behaviour among professional football players: Evidence from administrative data. *Journal of Gambling Studies*. Advance online publication.

Langrock, R., MacDonald, I. L., and Zucchini, W. (2012). Some nonstandard stochastic volatility models and their estimation using structured hidden Markov models. *Journal of Empirical Finance*, 19(1):147–161.

Mews, S., Langrock, R., Ötting, M., Yaqine, H., and Reinecke, J. (2024). Maximum approximate likelihood estimation of general continuous-time state-space models. *Statistical Modelling*, 24(1):9–28.

Michels, R., Ötting, M., and Langrock, R. (2023). Bettors' reaction to match dynamics: evidence from in-game betting. *European Journal of Operational Research*, 310(3):1118–1127.

Min, B., Kim, J., Choe, C., Eom, H., and McKay, R. B. (2008). A compound framework for sports results prediction: a football case study. *Knowledge-Based Systems*, 21(7):551–562.

Newall, P. W. and Cortis, D. (2021). Are sports bettors biased toward longshots, favorites, or both? A literature review. *Risks*, 9(1):22.

Ottaviani, M. and Sørensen, P. N. (2008). The favorite-longshot bias: an overview of the main explanations. *Handbook of Sports and Lottery Markets*, pages 83–101.

Ötting, M., Langrock, R., and Deutscher, C. (2018). Integrating multiple data sources in match-fixing warning systems. *Statistical Modelling*, 18(5-6):483–504.

Ötting, M., Michels, R., Langrock, R., and Deutscher, C. (2024). Demand for live betting: an analysis using state-space models. *Applied Stochastic Models in Business and Industry*, 40(2):527–541.

Razali, N., Mustapha, A., Yatim, F. A., and Ab Aziz, R. (2017). Predicting football matches results using Bayesian networks for English Premier League (EPL). In *IOP Conference Series: Materials Science and Engineering*.

Reade, J. J. and Akie, S. (2013). Using forecasting to detect corruption in international football. In *Proceedings of the 4th International Conference on Mathematics in Sport*.

Sauer, R. D. (1998). The economics of wagering markets. *Journal of Economic Literature*, 36(4):2021–2064.

Snowberg, E. and Wolfers, J. (2010). Explaining the favorite–long shot bias: is it risk-love or misperceptions? *Journal of Political Economy*, 118(4):723–746.

Sportradar (2024). Integrity in Action: 2024 Global Analysis & Trends. https://sportradar.com/content-hub/report/integrity-in-action-2024-global-analysis-and-trends/.

Sportradar (2025). Integrity in Action: 2025 Global Analysis & Trends. https://sportradar.com/content-hub/report/integrity-in-action-2025-global-analysis-trends/.

Statista (2025). Amount Wagered on Sports Betting Worldwide in 2024, by Sport. https://www.statista.com/statistics/1534873/leading-sports-handle-betting-worldwide/.

Taylor, S. J. (1994). Modeling stochastic volatility: a review and comparative study. *Mathematical Finance*, 4(2):183–204.

Thaler, R. H. and Ziemba, W. T. (1988). Anomalies: parimutuel betting markets: racetracks and lotteries. *Journal of Economic Perspectives*, 2(2):161–174.

Van Rompuy, B. (2015). The Odds of Match Fixing—Facts & Figures on the Integrity Risk of Certain Sports Bets. *SSRN Working Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5413142.

Warne, A., Coenen, G., and Christoffel, K. (2017). Marginalized predictive likelihood comparisons of linear Gaussian state-space models with applications to DSGE, DSGE-VAR, and VAR models. *Journal of Applied Econometrics*, 32(1):103–119.

Winkelmann, D. and Deutscher, C. (2026). Do betting markets sense a goal coming? Evidence from the German Bundesliga. *International Journal of Sport Finance*, page 15586235261463775.

Winkelmann, D., Deutscher, C., and Ötting, M. (2021). Bookmakers' mispricing of the disappeared home advantage in the German Bundesliga after the COVID-19 break. *Applied Economics*, 53(26):3054–3064.

Winkelmann, D. and Michels, R. (2026). Momentum effects in team sports: analyzing the

interplay between offense and defense in the NBA. *The American Statistician*, 80(3):432–442.

Winkelmann, D., Ötting, M., Deutscher, C., and Makarewicz, T. (2024). Are betting markets inefficient? Evidence from simulations and real data. *Journal of Sports Economics*, 25(1):54–97.

Wolfers, J. (2006). Point shaving: corruption in NCAA basketball. *American Economic Review*, 96(2):279–283.

Woodland, L. M. and Woodland, B. M. (1994). Market efficiency and the favorite-longshot bias: the baseball betting market. *The Journal of Finance*, 49(1):269–279.

Zhang, Q., Zhang, X., Hu, H., Li, C., Lin, Y., and Ma, R. (2022). Sports match prediction model for training and exercise using attention-based LSTM network. *Digital Communications and Networks*, 8(4):508–515.

Zucchini, W., MacDonald, I. L., and Langrock, R. (2017). *Hidden Markov Models for Time Series: An Introduction Using R*. Chapman and Hall/CRC, New York.

A Stake volumes per minute

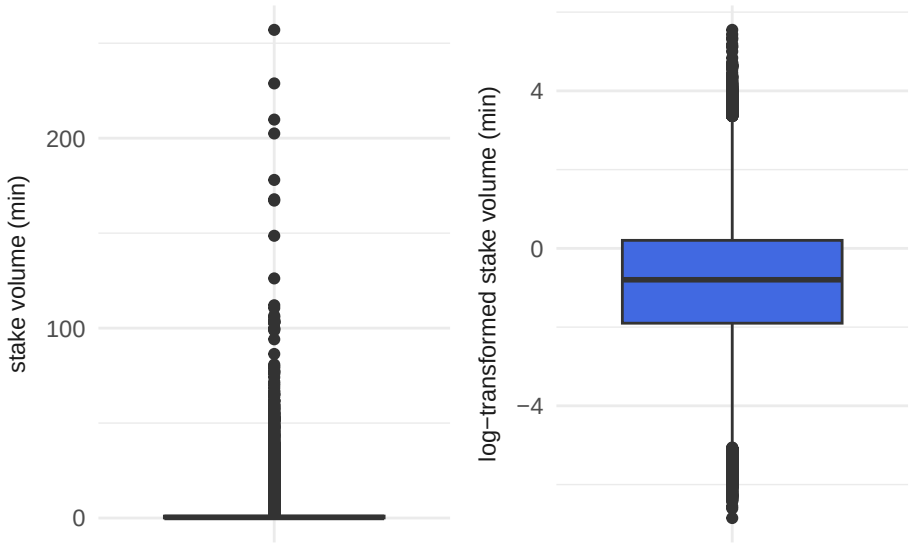


Figure A1: Distribution of stake volume per minute and its log-transformed version, shown as boxplots to compare the scale and skewness of the original and transformed data.

B Average team stakes

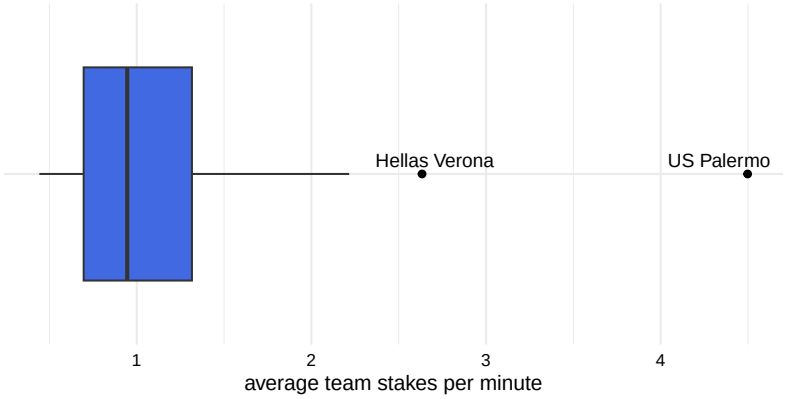


Figure B1: Boxplot of the average stakes placed per minute across all matches depending on the team.

C Total stakes per team

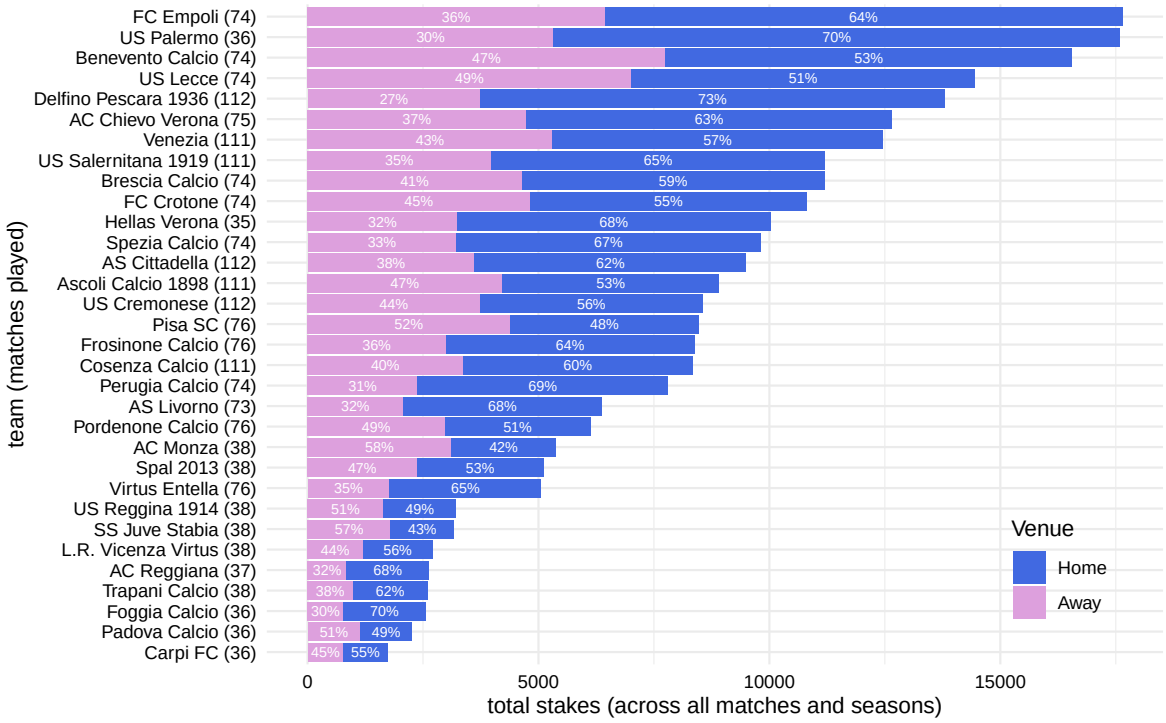


Figure C1: Total stakes placed on each team, split by whether the team played at home or away. Percentages give the home and away share of each team's total. The number of matches included differs across teams, since not all teams were in the league in all three seasons, and is given in the brackets after the team name.

D Stakes per scorediff

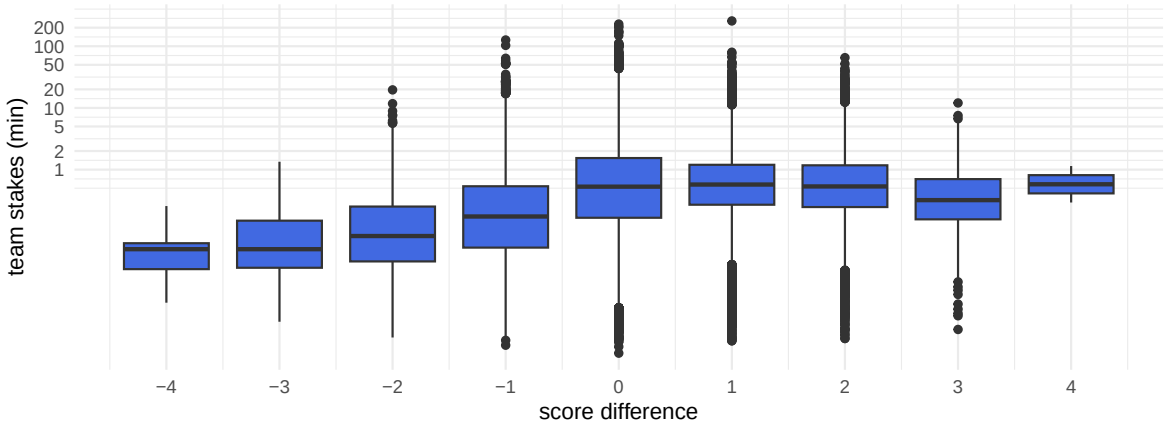


Figure D1: Boxplots of total stakes placed per minute by current score difference. Note that for visualisation purposes, we use a log scale on the y-axis.

E Percentage of flagged observations depending on quantile threshold

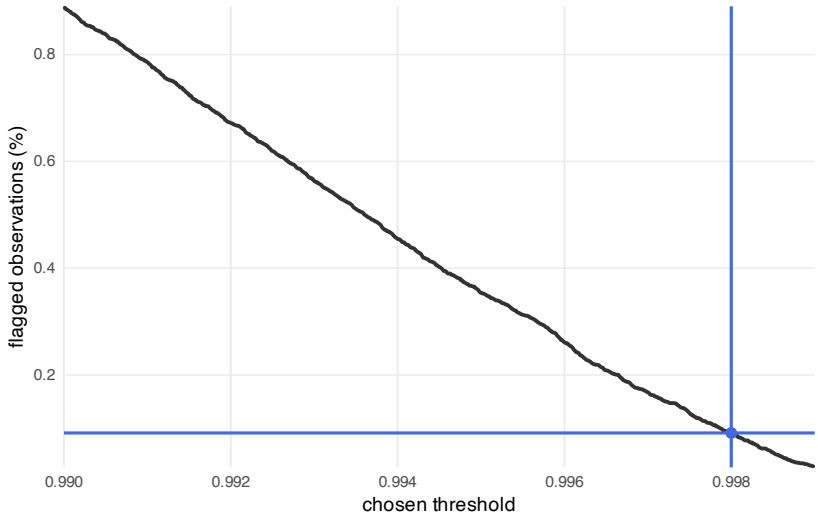


Figure E1: Percentage of flagged observations, depending on the chosen threshold of the model's predictive distribution of stakes placed.