Quantifying Behavioral Responses in Economics Through Data Science

Probabilistic Methods and Advanced Regularization for High-Dimensional Statistical Models

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von
Michael Nagel

Tübingen
2024

1. Betreuer:	Prof. Dr. Augustin Kelava
2. Betreuer:	Prof. Dr. Tim Pawlowski
Tag der mündlichen Prüfung:	30.04.2025
Dekanin:	Prof. Dr. Taiga Brahm
Dekan:	Prof. Dr. Dominik Papies
1. Gutachter:	Prof. Dr. Augustin Kelava
2. Gutachter:	Prof. Dr. Joachim Grammig

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Augustin Kelava and Prof. Tim Pawlowski, for their invaluable guidance, support, and mentorship throughout this journey. Their insightful feedback, encouragement, and expertise were crucial in shaping both my research and my personal growth during the course of this doctoral thesis. It has been a privilege to work under their supervision, and I am sincerely thankful for their unwavering belief in my abilities.

A special thank you goes to Prof. Thomas Dimpfl. His constant support and willingness to offer advice at every stage of my PhD and during my studies have been truly exceptional. Thomas was always there to listen, challenge my ideas, and offer thoughtful guidance, and his companionship throughout this process has made a significant impact on both my academic and personal development. I cannot express enough how much I appreciate his time, patience, and expertise during these years.

I would also like to extend my gratitude to Prof. Joachim Grammig from the University of Tübingen. He was one of the most influential professors during my studies, profoundly shaping my understanding of statistics and econometrics. His teachings have left a lasting impression on my academic journey.

Finally, my deepest thanks go to Lukas Fischer, who has been by my side from the start of my studies through the completion of this PhD. Our collaborations not only enriched the work but made the entire experience meaningful and filled with unforgettable moments. I am grateful for the countless experiences we shared, the insightful ideas he contributed, and the great time we spent together along the way.

Remarks

This doctoral thesis is part of a joint research project between the Methods Center (Prof. Augustin Kelava) and the Chair of Sport Economics, Sport Management, and Media Research (Prof. Tim Pawlowski) at the University of Tübingen. The project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2064/1, Project 390727645). Parts of the project were presented at different academic events, including the annual conferences of the Cluster of Excellence - Machine Learning in Science (2021, 2022) and the workshop "Machine Learning meets Quantitative Social Sciences" in Tübingen (2023) as well as at the Reading Online Sport Economics Seminar (2021), the St. Gallen Economics Research Seminar (2023), a workshop hosted by the Erasmus School of Economics (2024), and the 53rd DGPs Congress/15th ÖGP Conference (2024) in Vienna.

Chapter 1

Introduction

Behavioral economics offers a framework for understanding human decision-making by incorporating psychological insights to explain systematic deviations from rationality. Unlike traditional economic models that assume fully rational agents, behavioral economics acknowledges that human behavior is often constrained by cognitive limitations, influenced by emotions, and shaped by social contexts (Chen 2024). This perspective enables the study of phenomena such as loss aversion, reference dependent preferences, and non-Bayesian judgment—the perception that the intuitive probability assessments of people may diverge from statistical principles like Bayes' rule. These topics have attracted significant empirical attention across fields like economics, psychology, and sociology, contributing to a more realistic understanding of decision-making processes (Camerer, Loewenstein, and Rabin 2006; Chen, Lakshminarayanan, and Santos 2006).

Particular focus has been placed on reference dependence and loss aversion, biases which have been explored in both experimental and real-world settings (e.g., Abdellaoui, Bleichrodt, and Paraschiv 2007; Marzilli Ericson and Fuster 2011; Pope and Schweitzer 2011). The prospect theory of Kahneman and Tversky (1979) has become a cornerstone for understanding how reference dependence and loss aversion influence decision-making under uncertainty. It suggests that people evaluate outcomes based on potential gains or losses relative to a reference point. A key insight from this theory is that individuals are generally loss averse, indicating that they experience losses more intensely than equivalent gains.

Building on this, Kahneman, Knetsch, and Thaler (1990) report that individuals tend to assign a higher value to objects they own. This phenomenon is known as the endowment effect and can be explained by emotional attachment, sentimental value, or symbolic importance. The implications are far-reaching. For instance, loss aversion in the context of the endowment effect can help explain systematic differences between selling and buying prices. In the real estate market, Genesove and Mayer (2001) show that sellers facing potential losses tend to set 25-35% higher asking prices than expected selling prices and achieve 3-18% higher final selling prices than if they were not subject to such losses.

Deviations from rationality are also evident in financial stock markets, where market participants often respond inconsistently to new information. Over- and underreactions serve as prominent behavioral mechanisms in the asset pricing literature. For instance, Barberis, Shleifer, and Vishny (1998) develop an investor sentiment model based on psychological evidence to explain stock price over- and underreactions to both positive and negative news. Similarly, Daniel, Hirshleifer, and Subrahmanyam (1998) argue that investors incline to overreact to private information and underreact to public news, explaining market anomalies such as long-term reversals,

momentum, excessive volatility, and return predictability following public events. This behavior is linked to two psychological biases: overconfidence in the reliability of private information and self-attribution bias, which inconsistently impacts investor confidence based on investment outcomes. Behavioral economics, therefore, enhances our understanding of why decision-makers often diverge from rational economic predictions, with significant implications for fields such as consumer behavior, public policy, and financial decision-making.

This dissertation builds on existing literature respecting the core principles outlined by Camerer, Loewenstein, and Rabin (2006), who emphasize the importance of incorporating empirically-driven models that account for real-world psychological influences on economic decision-making. The primary objective of this thesis is to quantify the short-term behavioral responses of the average market participant to new information and contribute robust empirical evidence through the application of advanced statistical methods. To achieve this, this present work uses sports betting markets as a controlled laboratory, which offer a unique setting characterized by exogenous terminal values, short and predetermined contract expiration dates based on event outcomes that are independent of participants' actions[1], and rich longitudinal data (Moskowitz 2021). This controlled environment provides a robust foundation for generalizing the findings to broader financial and economic markets.

Sporting events have sparked significant research interest due to several distinct advantages (Moskowitz 2021; Palacios-Huerta 2023). First, they enable the identification of effects, which is often challenging in other settings where direct physical or financial consequences confound the relationships being studied. Sports bets are isolated from systematic economic risks and have exogenous terminal values, allowing for a clearer analysis of investor behavior and pricing anomalies without the confounding effects of broader market risks. Second, sports are a regular part of everyday life for many people and evoke intense but "normal" emotions. Third, betting markets provide valuable data that facilitate the measurement of reference point-dependent emotions, which form a central focus of this dissertation. Additionally, this thesis highlights the importance of advanced statistical models—particularly new Bayesian regularization methods, for effectively managing large datasets and addressing complex, high-dimensional models—and emphasizes the critical role of sophisticated methods in extracting meaningful insights from data.

Chapter 2[2] examines the impact of immediate emotions on human behavior, by exploring whether and how experienced surprise and suspense are related to alcohol use as a mass phenomenon. Empirical research on how emotions affect behavior is scarce. Most existing studies

[1]This statement focuses on legitimate market behavior, explicitly excluding any illegal activities within betting markets, such as corruption, match-fixing, or bribery. For example, Wolfers (2006) highlights point-shaving in National Collegiate Athletic Association (NCAA) Division I basketball, where players intentionally limit their team's margin of victory to manipulate betting outcomes. The findings suggest that approximately 6% of strong favorites engage in such practices, affecting around 1% of all games.

[2]Chapter 2 is based on Fischer et al. (2024), available at SSRN.

rely on lab experiments that struggle to induce or control intense emotions (Tymula and Glimcher 2018). Unlike previous work, which often examines post-match behavior with significant time lags between the emotional trigger and the observed actions, the study focuses on the immediate behavioral responses to the reference point-dependent emotions surprise and suspense, as formalized by Ely, Frankel, and Kamenica (2015). Through an in-depth analysis of emotional drinking, the study provides the first field evidence, ensuring high external validity.

To explore the connection between emotions and alcohol use, the analysis employs a random intercept model, estimated using Bayesian methods with regularization. The dataset consists of high-frequency transaction data on beer sales from nearly 100 matches played by a top-division soccer team in Germany and detailed in-play event data and betting odds. The evaluation accounts for both negative and positive emotional states during the match—a critical distinction, as prior research suggests that individuals experiencing positive emotions tend to be more risk averse than those experiencing negative emotions (Isen 2008). The findings provide robust evidence of emotional drinking, with stable net effects, indicating that surprise leads to increased beer sales. This is in line with psychological theories (e.g., Cooper et al. 1995; Greeley and Oei 1999) that people turn to alcohol in emotionally charged situations and might help explain existing evidence on sporting event-related violence or traffic fatalities, for instance. Additionally, the results show that individuals tend to respond more quickly to negative emotional signals than to positive ones.

Chapter 3[3] builds on the previous chapter by addressing the challenges associated with estimating complex, high-dimensional models like those discussed earlier. In these contexts, traditional estimation methods like ordinary least squares often produce biased parameter estimates with low accuracy due to overfitting, overparameterization, or the failure of asymptotic assumptions when the number of predictors (D) exceeds the number of observations (N). Dimensionality reduction becomes especially important when the design matrix is sparse or the data exhibit high multicollinearity.

In Bayesian frameworks, regularization can be achieved through shrinkage priors. However, existing shrinkage priors come with limitations. For example, spike-and-slab priors, traditionally regarded as the gold standard in sparse Bayesian estimation, become computationally intensive in high-dimensional settings due to slow mixing and convergence (Mitchell and Beauchamp 1988; Bhadra et al. 2017). The Bayesian lasso, on the other hand, struggles to sufficiently shrink small coefficients (representing noise) while excessively shrinking large coefficients due to its relatively light-tailed distribution, making it less robust to outliers (Carvalho, Polson, and Scott 2008; Armagan, Dunson, and Lee 2013). Similarly, the horseshoe prior has practical limitations. As noted by Piironen and Vehtari (2017), it lacks a mechanism for incorporating prior knowledge about sparsity levels in the parameter vector, often leaving too many parameters unshrunk.

These limitations, along with the need for greater precision and adaptability, motivated the

[3]Chapter 3 is based on Nagel et al. (2024), published in Structural Equation Modeling: A Multidisciplinary Journal.

development of the Dirichlet-horseshoe prior, a novel approach that combines and extends the principles of the regularized horseshoe (Piironen and Vehtari 2017) and Dirichlet-Laplace priors (Bhattacharya et al. 2015) to overcome the shortcomings of existing methods. Compared to the regularized horseshoe, the Dirichlet-horseshoe applies stronger shrinkage to predictors that are less important than the average, while applying less shrinkage to those that are more important. As the explanatory power of the predictor variables becomes more uniform, the Dirichlet-horseshoe gradually approximates the regularized horseshoe, encompassing it as a special case when all predictors have equal importance. This adaptive approach allows the Dirichlet-horseshoe to act as a regularizer with enhanced selectivity, effectively distinguishing between significant and insignificant signals.

A series of simulations and empirical applications show that this new prior outperforms standard alternatives in terms of accuracy and precision, while also providing competitive coverage ratios. This approach offers a robust and highly accurate solution for high-dimensional estimation in finite samples across various sparsity levels, including both $N < D$ and $N > D$ scenarios, and under different levels of predictor correlation.

Chapter 4[4] addresses the dissertation's overarching research question by investigating how the average market participant and market makers behave and respond to new information—particularly when initial prices display biases and are imprecise—and quantifies how rapidly the market reacts to this information. Understanding the mechanisms behind price discovery and equilibrium dynamics is crucial in financial market research. The concept of tâtonnement, introduced by Walras (1889), describes a process in which agents continuously adjust their offers until supply and demand reach equilibrium—a key aspect of price discovery. Biais, Hillion, and Spatt (1999) apply this concept to the preopening period in the Paris Bourse, demonstrating that early preopening prices are noisy, but become more informative as the market approaches opening. This convergence towards equilibrium serves as a fundamental component for understanding price formation.

Research in this area is often constrained by the absence of a fundamental or terminal value for assets, necessitating the use of proxies—typically the last available price—which can introduce bias. To address this limitation, the study investigates the sports betting market as a setting where terminal values are known. Previous studies in sports settings often rely on in-play data or assess aggregated effects from opening to closing using cross-sectional data from a single bookmaker. In contrast, this analysis employs time series data from multiple bookmakers across the entire betting period, marking the first investigation of pre-match learning behavior in the sports betting market with such comprehensive data.

The findings offer new insights into price discovery and learning. While opening prices are often noisy and biased, they become more accurate as the market approaches closure, indicating significant learning primarily driven by the average bettor, who continuously updates their

[4]Chapter 4 is based on Nagel (2024), available at SSRN.

beliefs based on new information. Notably, evidence of learning emerges not immediately at the start of the betting period but increases toward the middle, continuing almost until closing. Learning rates, estimated using both the Generalized Method of Moments (GMM) and Bayesian methods, show substantial variation across bookmakers and market segments. Higher initial prices exhibit faster learning, and learning rates differ according to competition levels, with quicker adjustments observed in prominent, high-stakes matches. This heterogeneity suggests that the market's ability to incorporate information and correct mispricings varies with the level of competition, match prominence, and price ranges.

The final chapter synthesizes and integrates the key aspects from each study. It examines their implications for behavioral economics, economic decision-making, and broader financial contexts, while also discusses the limitations of the research. Additionally, it proposes potential avenues for future investigation, including applications in fields such as finance and labor economics, areas where cognitive biases and emotional reactions can also significantly influence decision-making. Investigations within these frameworks could contribute to further exploring key questions in behavioral economics and strengthen the understanding of economic behavior and decision-making processes.

Chapter 2

Emotional Drinking: Surprise, Suspense, and Alcohol Use During Soccer Matches[5]

2.1 Motivation

Surprise and *Suspense*, as formalized by Ely, Frankel, and Kamenica (2015), are reference point-dependent emotions[6] that can explain why people seek non-instrumental information by watching a sporting event live or following political news before presidential elections, for example. Consequently, *Surprise* and *Suspense* might explain demand in the entertainment industry but also inform the analysis of broader social issues like voting. In this paper, we explore whether and how experienced *Surprise* and *Suspense* are related to alcohol use as a mass phenomenon.[7]

For many years, economic research has extensively addressed the emotions that people might experience after making a decision, such as regret (Loomes and Sugden 1982, for example), but it has rarely considered the effects of immediate emotions on human behavior. More recent studies explicitly incorporate cues, such as the smell of alcohol, into models of human behavior though (e.g., Laibson 2001). Likewise, Loewenstein (2000) elaborates on the theoretical relevance of immediate emotions and the consequences of a wide range of visceral factors, such as anger or fear, for bargaining behavior, decision-making under risk and uncertainty, or intertemporal choice. He concludes that "visceral factors play an essential (probably the dominant) role in human behavior" (p. 427).

The anecdotic evidence for such a claim is manifold. This includes, for example, couples failing to work out prenuptial agreements in happy times, politicians entangling in sex scandals ruining their careers, as well as violent and risky compensatory behavior like emotional eating or drinking. Robust empirical evidence on how emotions influence behavior, however, is scarce. Moreover, existing studies have commonly relied on lab experiments, which by their very nature cannot induce (or control) intense emotions (Tymula and Glimcher 2018). Consequently, researchers have started to leverage settings that induce strong emotional responses,

[5] This chapter is a revised version of Fischer et al. (2024), available at SSRN.

[6] In detail, *Surprise* is a backward-looking emotion that relates events, such as goals scored during a soccer match, to anterior beliefs about the final outcome of a match, for instance. *Suspense* instead is an anticipatory forward-looking emotion attributed to the variance in the next period's beliefs.

[7] Previous studies explore the effects of *Surprise* and *Suspense* on sports demand, in the context of Wimbledon tennis (Bizzozero, Flepp, and Franck 2016), Premier League soccer (Buraimo et al. 2020), UEFA Champions League matches (Richardson, Nalbantis, and Pawlowski 2023), or eSports events (Simonov, Ursu, and Zheng 2023), as well as on social media use (Pawlowski et al. 2024). Kessler et al. (2022) examine how short-term fluctuations in incidental happiness affect economic behavior by conducting a series of experiments in a sports bar, but the authors consider self-reported measures of surprise and excitement.

such as weather extremes, terrorist attacks, political conflicts, or sporting events.

In recent years, especially sporting events have become an area of interest because they typically have some key advantages over the other mentioned settings (Palacios-Huerta 2023). First, the effect of emotions can be clearly identified, which is much more difficult (if not impossible) in the other settings, where direct physical or financial consequences confound the relation of interest. Second, sporting events in general, and regular league games in particular, are present in many peoples' everyday life. Therefore, they are associated with an intense but more "normal" set of emotions. Third, betting markets provide useful information to accurately measure reference point-dependent emotions which are in the focus of this research.

Overall, focused on post-match behavior, these studies exploit settings in which the temporal lag between the emotional trigger and observed behavior is comparably large.[8] This is different in our paper, in which we explore the immediate behavioral response to the reference point-related emotions *Surprise* and *Suspense*. We propose extending the framework by Ely, Frankel, and Kamenica (2015) to support our attempt to provide the first field evidence of the effects of immediate emotions on alcohol use.

Alcohol use in general is deeply rooted in society, widely recognized as a leading risk factor for death and disability (e.g., Griswold et al. 2018), and a major contributor to criminal behavior such as family violence (Klostermann and Fals-Stewart 2006). However, even though "virtually all major theories of drinking behavior and alcohol problems include an important role for emotional factors" (Lang, Patrick, and Stritzke 1999, p. 328ff.), field evidence of the immediate effects of emotions on alcohol use as a mass phenomenon is largely missing.

In this regard, sporting events are a particularly relevant setting to explore, given the propensity for increased alcohol consumption by spectators[9] and the established links between sporting events and traffic fatalities (Wood, McInnes, and Norton 2011), family violence (Card and Dahl 2011), and other forms of criminal behavior (Munyo and Rossi 2013). Due to data limitations, these studies are not able to comprehensively explore the role of alcohol directly. However, Klick and MacDonald (2021), for instance, provide some indirect evidence for the impact of game-related alcohol consumption on crime near the stadium by exploiting variation in Major League Baseball (MLB) match duration between the last alcohol call at the end of the seventh inning and the end of the match. Moreover, there is some survey evidence suggesting that a disturbing number of spectators with comparably high levels of blood alcohol levels claims to be driving

[8]Edmans, Garcia, and Norli (2007), identifying negative domestic stock market reactions following losses of national soccer teams, were among the first to investigate the effects of sports sentiment. Others have more explicitly focused on reference point-based behavior by exploiting upsets. For instance, upset losses in the National Football League (NFL) (Card and Dahl 2011) or the National Basketball Association (NBA) (Cardazzi et al. 2022) increase family violence, and those in college football leagues apparently increase the sentence lengths imposed by juvenile court judges (Eren and Mocan 2018), while upset wins in colleague football increase excessive partying and reports of rape (Lindo, Siminski, and Swensen 2018). In contrast, neither upset wins nor upset losses are associated with football-related violence in Germany (Andres, Fabel, and Rainer 2023).

[9]See https://adf.org.au/insights/alcohol-and-sport-potent-mix/, for example (retrieved on April 3, 2024).

home after a sporting event (Wolfe, Martinez, and Scott 1998).

To investigate the link between emotions and alcohol use, we combine high-fre-quency transaction data on beer sales during almost 100 matches played by a first division soccer team in Germany with the corresponding in-play event information and betting odds. These data allow us to measure *Surprise* and *Suspense* and to account for potential negative and positive emotional states during a match. Theoretically, this differentiation seems important because people experiencing positive (rather than negative) affect were found to be more risk averse (Isen 2008). As a result, a home team fan is likely to make different consumption decisions after a goal scored by the home team than after a goal scored by the away team, for example. Moreover, in contrast to North American sports where there are rules related to cutting of alcohol sales during the game, there are no such constraints on alcohol sales in our setting. If alcoholic beverages are forbidden, their sale in the stadium is prohibited before, during, and after the entire match. This is a rare occurrence in our sample, however, and we consider it in the analysis.

Our proposed random intercept model is estimated by means of Hamiltonian Monte Carlo and has two levels. Level 1 (within-level) consists of match minute-specific information. Level 2 (between-level) contains match-specific data to control for differences in alcohol sales across matches. This empirical design may help to justify the conditional independence assumption. However, it raises concerns about issues of reverse causality. In principle, if alcohol use influences fan behavior, it also might affect relevant match events. Home advantage is a well-established phenomenon in sports, and it can be attributed, at least partly, to crowd noise and social pressure on referees (Garicano, Palacios-Huerta, and Prendergast 2005). Fan support tends to be dynamic during a match, and some of that dynamism might reflect the alcohol level consumed by the crowd, such that alcohol use at time t might influence the style of play, for instance, which could ultimately affect match outcome. However, the likelihood of alcohol-induced enthusiastic cheering altering key events during the match seems negligible, given the average of 88.5 beers sold per minute during the match in our sample.

Overall, our findings provide some robust evidence for emotional drinking. First, we find stable net effects for *Surprise* increasing beer sales suggesting—in line with psychological theories (e.g., Cooper et al. 1995; Greeley and Oei 1999)—that people turn specifically to alcohol in emotionally charged situations. Second, the revealed temporal pattern of the effects of *Suspense* on beer sales—which are initially negative and then positive—can be attributed to the fear of missing important plays when *Suspense* increases and the desire to alleviate boredom by drinking alcohol when *Suspense* is low (Patrick and Schulenberg 2011). This finding lends some credibility to the claim by Wood, McInnes, and Norton (2011) that sporting event-related traffic fatalities after close matches are due to aggressive driving caused by high levels of testosterone rather than excessive drinking during the match. Finally, while we find that people seem to respond more quickly to negative (rather than positive) signals, our findings do not differ significantly between positive and negative emotional states.

2.2 Data and Measures

The underlying data consist of three pools. The first pool gathers transaction data (second-by-second sales of beer, water, hot wine punch, and so on) from VfB Stuttgart, a German first division soccer club. For the second pool, we obtain data regarding the sports betting market and basic match facts (e.g., minute-by-minute in-play and closing odds, date of the match, kickoff time) from NowGoal and OddsPortal. The third pool includes key match events, match day information, and further information on the teams (e.g., goals, red cards, ten minute-by-ten minute weather) from OptaSports, kicker.de, transfermarkt.de, Deutscher Wetterdienst (German weather service), Google Maps, kalender-online.com, schulferien.org, and cannsatter-volksfest.de. For more details and exact sources see appendix Tables A.3 and A.4, respectively.

2.2.1 Setting

We observe spectator behavior in the MHP (former Mercedes-Benz) Arena, where the first division (Bundesliga) soccer club VfB Stuttgart plays its home matches. Exploiting real in-play match events and using rarely accessed in-play transaction data (excluding sales during halftime, pre- and post-match) from the Mercedes-Benz Arena offers a promising approach, for several reasons. First, VfB Stuttgart's home matches draw some of the highest average attendances in Europe (e.g., 51.983 in season 2015/16[10]), so our analysis should be resilient, due to the high number of transactions. To track transactions, the club uses TCPOS[11] (point-of-sell technology), which documents every transaction taking place at any cash point in the stadium in a fully coherent, self-contained system.

Second, during the observed seasons, no stadium construction projects were taking place. Thus, no selection effects arise due to the unavailability of certain stands, for example. Moreover, no unexplained heterogeneity can arise as the result of changing beer vendors; for the entire sample, the vendor is the same, namely, Krombacher.

Third, this stadium imposes a strict segregation of spectators in sectors, surrounded by fences, during all matches, which is unlike the situation in many other stadiums in the Bundesliga. Across five areas, spectators cannot flow easily[12], though they can enter the catacombs in their section, which feature screens that allow them to watch the match and its key evens even if they are not in the stands, such that in-play information could affect their consumption decisions even

[10] See https://www.worldfootball.net/attendance/bundesliga-2015-2016/1/, for example (retrieved on July 19, 2024).

[11] TCPOS is one of the global leading providers in the point-of-sell technology sector for hospitality and retail industries and is the official service provider in the Mercedes-Benz Arena. We can rule out inaccuracy due to the absence of real-time information. We are able to perfectly match real time with the match minute, such that delayed kickoffs do not lead to mismatched data points.

[12] For some matches, the away fan area is extended by seats in the Untertürkheimer Kurve (opposite of the Cannstatter Kurve). We address this anomaly in the data preparation stage.

when they are waiting in line to make their purchases. This segregation enables us to estimate effects for different stands and isolate home team versus away team fans, who likely experience diametrically opposed emotions for the same match events.

Fourth, cashpoints in the stadium are located close to every stand, but it takes some time to reach them. Still, contemporaneous or shortly delayed effects, such as in the same minute of or shortly after a signal (e.g., a goal or a red card), could occur if people who are already at a cashpoint impulsively decide to purchase. The same holds true for spectators on their way, such that they are in closer proximity to a cashpoint. Beer sales also might decrease in the short time span surrounding an event if consumers and salespeople spontaneously turn their attention to a monitor to see a replay. To explore such temporal patterns in behavior, we estimate contemporaneous effects, as well as effects up to 9 minutes lagged.

Fifth, though we cannot draw conclusions at the individual level, the investigation of alcohol use as a mass phenomenon is beneficial, especially as a form of reference point behavior. According to wisdom-of-the-crowd effects, the real expectations of people in aggregate should be closer to rationally expected reference points than individual appraisals, which tend to be highly biased. Research analyzing the individual effect of emotional cues must consider this deviation, which can be challenging.

2.2.2 Beverage Sales

We operationalize alcohol consumption first in relation to beer sales, as the main dependent variable. Subsequently, we also consider shandy and soft drinks to test for the relevance of alcoholic strength. Water and hot wine punch sales provide controls for baseline consumption and substitution effects, respectively. Table 2.1 summarizes beverage sales on a per minute basis.

Table 2.1
Descriptive Statistics: Beverage Sales Per Minute

	count	mean	std	min	25%	50%	75%	max
Beer	8,820	88.5	57.4	0	44	76	123	320
Hot Wine	8,820	2.19	6.28	0	0	0	1	69
Shandy	8,820	8.64	8.35	0	3	6	12	68
Soft Drinks	8,820	16.7	17.1	0	5	11	23	130
Water	8,820	2.89	5.24	0	0	1	3	68

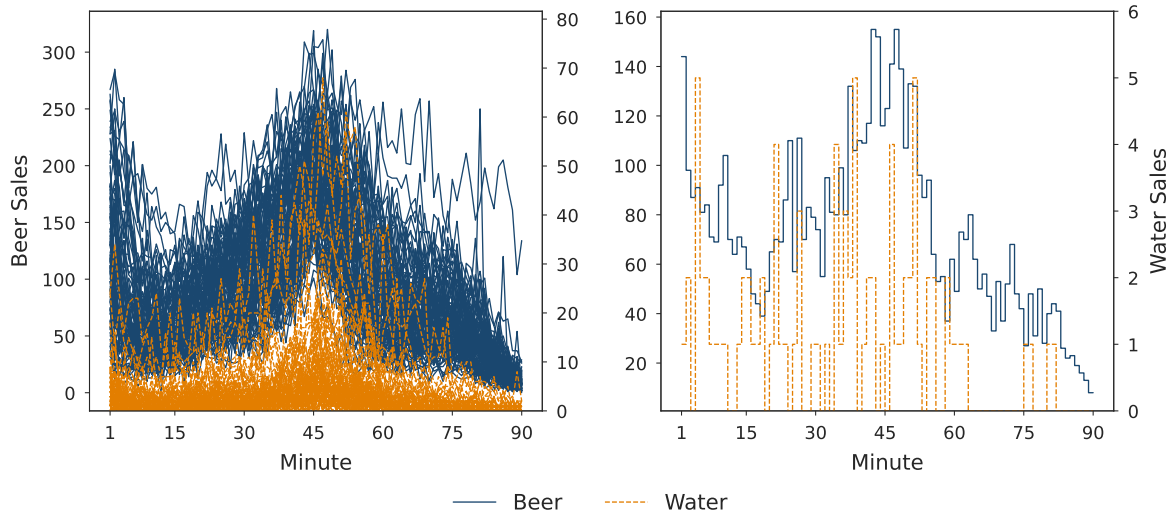
This table presents basic summary statistics of beverage sales on a per minute basis. The count variable is calculated by the number of matches (98 in the main specification) times 90 match minutes (halftime break excluded). On average there are 88.5 beer sales per minute in our sample. The number of beer sales serves as the dependent variable in the main specification and is substituted by shandy and soft drink sales in two alternative specifications. Water and hot wine (punch) control for baseline consumption and substitution effects, respectively. Percentages of 25%, 50%, and 75% denote the respective quantiles.

Clearly, beer accounts for the most transactions, with an average of 88.5 per minute. The corresponding, relatively high standard deviation of 57.4 give rise to the question of why beer sales vary so much during matches. For all beverages, we also observe match minutes with no sales. However, taking the lags, all observed minutes in the main specification actually exhibit

beer sales; the zero sales minutes occur at the very beginning of the match. A typical match in our sample exhibits 76 beer, 0 hot wine punch, 6 shandy, 11 soft drink, and 1 water sale(s) per minute (see 50% quantile in Table 2.1). The fewest transactions actually involve water, taking into account that hot wine punch is only sold in winter months.

Figure 2.1 illustrates the numbers in Table 2.1 by depicting the time series of beer and water sold in the whole stadium, except for the away fan areas, for all matches in the sample (left) and during a match between VfB Stuttgart and SC Freiburg on Sunday, February 3, 2019, at 6:00 PM (right), which ended in a draw (in stoppage time, which we do not consider in our analysis). For beer, we identify a third-degree polynomial sales pattern over time (sales during the break

Figure 2.1
Beer and Water Sales, Full Sample and Exemplary Match



This figure shows the development of beer and water sales over time for all matches in the sample (left) and for an exemplary German first division soccer match between VfB Stuttgart and SC Freiburg on February 3, 2019 at 6:00 PM (right). The underlying data reflect the whole stadium, except for away fan areas. Sales during the break are excluded. We do not consider stoppage time. The figure depicts a typical movement of beer sales, which first decrease, then peak around the half-time break, and decrease towards the end of the match.

excluded). Immediately after kickoff, beer sales begin to decrease signifying that fans who have paid to see live sports leave the cashpoint areas where they can buy alcohol. The observed delay likely reflects the queues at the cashpoints due to a high demand before kickoff. Also, fans arriving too late or just before kickoff and not wanting to forgo a beer can contribute to this pattern.

Then, just before halfway through the first half, beer sales reach their local minimum before they rise and find a maximum right after halftime break. Obviously, there is nothing to miss during the break in terms of live sports; halftime shows are either not presented or rather unattractive in German soccer (cf. many U.S. sports) and the vast majority of fans do not have access to convenience rooms. Directly after halftime (right next to the 45th minute), sales remain at their overall maximum, again likely due to delay effects. We attribute the increase in the first half to anticipation effects, such that consumers likely expect high demand and long lines during

the break. Economically speaking, people with lower opportunity costs of missing out go to buy early.

In the second half, beer sales drop considerably. Recall, there are no rules related to cutting off alcohol sales like in the NBA. Therefore, this overall decline in sales in the second half might occur because fans leave before the end of the match to avoid crowds or traffic jams (congestion). An unfinished beer extends time in the stadium, because there is a deposit for returned cups. Especially in cold winter months, people could substitute a drink in the second half with one after the match in a bar near the stadium or at home, which arguably is more comfortable than in the stadium.

All in all, water sales do exhibit a less clear pattern than beer sales. However, we can see a typical increase at the beginning of the match and decreasing volume towards the end. More relevant for our research, we do not find the distinct local maximum at the start across all matches, for example, and the dynamics of beer and water sales show systematically different dynamics on a per-match basis. By explaining beer sales, we are not summarizing transactions in general but alcohol sales in particular.

2.2.3 Reference Point-dependent Emotions

Emotional cues, as we construct and calculate them for this research, refer implicitly to reference point-based behavior. In this regard, we complement recent papers that use professional sports data in empirical tests of the theoretical prediction by Kőszegi and Rabin (2006) that utility from deviations between rationally expected references points and actual outcomes influence behavior under uncertainty (e.g., Card and Dahl 2011; Eren and Mocan 2018; Lindo, Siminski, and Swensen 2018). Whereas these prior studies focus on post-match sentiment, we explore the effects of emotional cues during a match.

As noted, we focus on the emotional cues *Surprise* and *Suspense*, theoretically introduced by Ely, Frankel, and Kamenica (2015). With respect to their formulation for a soccer context, we follow Buraimo et al. (2020). Our key explanatory variables are

$$\textit{Surprise}_t = u\left(\sum_{m\in H,D,A}\left[p_t^m - p_{t-1}^m\right]^2\right) \text{ and} \tag{2.1}$$

$$\textit{Suspense}_t = u\left(\sum_{m\in H,D,A} p_{t+1}^{HG}\left[\left(p_{t+1}^m \mid p_{t+1}^{HG}\right) - p_t^m\right]^2 + p_{t+1}^{AG}\left[\left(p_{t+1}^m \mid p_{t+1}^{AG}\right) - p_t^m\right]^2\right) \tag{2.2}$$

$$\text{with } u(\cdot) = \sqrt{\cdot} \tag{2.3}$$

in which variable *p* represents a probability. The match minute is given by *t*. The path $m \in \{H, D, A\}$ can refer to home (*H*), draw (*D*), or away (*A*). Superscript *HG* indicates a "Home Goal", such that $p_{t+1}^D \mid p_{t+1}^{HG}$ is the conditional probability in minute $t+1$ of the match ending in a draw given the home team scores in minute $t+1$, for example. The equivalent *AG* indicates an

"Away Goal."

As Equation (2.1) specifies, *Surprise* is a backward-looking measure. It is based on the path-specific difference between the current outcome probability and the outcome probability in the period just before the current one, which serves as reference point. To capture the entire movement of changes, we can square and sum the differences between probabilities over all possible outcomes. Squaring the components within the sum prevents the path-specific parts from canceling each other out and also weights large deviations more heavily. *Surprise* is high if the current match situation does not align with former expectations, such as when an underdog scores an opening goal just before the end of the match.

In contrast, *Suspense* is a forward-looking measure, suggesting a what-if scenario that reflects the difference between the current outcome probability and the outcome probability in the next minute, given that teams score in the next minute. Very unlikely goals might influence the outcome probabilities strongly, which is why deviations from the reference points are weighted by the probability of the event happening. Summing over the weighted probability space (home win, draw, and away win) shows the expected value character of *Suspense*. Finally, *Suspense* is high if the variance of potential outcomes for the next period is large. For example, toward the end of the match, the two teams might equally likely score with a high probability in the next period, and the current score indicates a draw.

The underlying utility function in Equation (2.3) controls for the curvature, reflecting the relative valuation of low versus high values. If $u(\cdot)$ is equal to the identity function, it implies that low cue portions are of similar importance to high portions. Root functions with low exponents near 0 instead denote strongly diminishing utility.

Both cues are based on match outcome probabilities, for which we use transformed in-play betting odds. However, both the unconditional probabilities and the conditional probabilities specify hypothetical scenarios we cannot directly derive from bookmaker odds. For the unconditional probabilities p_{t+1}^{HG} and p_{t+1}^{AG}, as well as the conditional probabilities $p_{t+1}^{m} \mid p_{t+1}^{HG}$ and $p_{t+1}^{m} \mid p_{t+1}^{AG}$ in *Suspense* (Equation (2.2)) we follow and refine the simulation approach proposed by Buraimo et al. (2020). For the exact procedure of calculating the cues, see Section A.1.

Endogeneity seemingly might be a concern, because emotions could trigger odds and therefore emotional cues. But we obtain the odds from the Asian bookmaker Crown, which should be irrelevant for the typical fan of VfB Stuttgart (in contrast with a regionally popular bookmaker like tipico). Therefore, even when liquidity is low, emotionally involved people are unlikely to account for significant betting volume. Instead, betting odds generally are likely to stem from the high stakes of professional bettors (or syndicates) that use statistical models to determine underpriced odds.

Outcome probabilities based on in-play betting odds inherently exhibit missing values, because the market closes around important match events like goals when the odds are updated. Another reason for missing odds is the low liquidity that occurs when a match is practically decided before the end. For the imputation of these values, we use a gated recurrent unit (GRU)

network via TensorFlow (Abadi et al. 2015), which includes goals, red cards, and previous outcome probabilities as features. In Appendix A.1.2.1, we describe this imputation procedure in detail.

2.2.3.1 Emotional States

We distinguish between negative and positive emotional states, according to the perspective of home fans. To be precise, we employ endogenous state switching (indicated by s) according to the definitions in Table 2.2, where $\Delta p_t = p_t - p_{t-1}$ and the current state is maintained when at least one of the probabilities did not change, such that for any path $\text{sgn}\,\Delta p_t = 0$. Irrespective of numerical fluctuations, a change in the probability of one path should imply changes in the probabilities of the other paths.

Table 2.2
Definition of State Switches

$\text{sgn}\,\Delta p_t^H$	$\text{sgn}\,\Delta p_t^D$	$\text{sgn}\,\Delta p_t^A$	s
-1	1	1	0
-1	-1	1	0
1	-1	-1	1
1	1	-1	1
-1	1	-1	$\begin{cases} 0, \text{if } \Delta p_t^H < \Delta p_t^A \\ 1, \text{if } \Delta p_t^H > \Delta p_t^A \end{cases}$
1	-1	1	

This table presents the criterions used to define emotional states. For example, when the sign of Δp_t is negative (i.e., the home probability decreased) and the sign of Δ for the other two paths is positive, we assume the home fans to be in negative state $S = 0$. For all other combinations that are not part of the table, the system stays in the current state (e.g., when at least one of the Δs is 0). Hence, we do not switch to the positive state $S = 1$ if the home probability has increased, the draw probability has stayed the same, and the away probability has decreased. This is because a change in the home and away probability should go along with a change in the probability for a draw. Contradicting cases of all signs being equal to 1 or -1 (probabilities need to sum up to 1) are not shown.

Related to that, if we are unable to define a state for $t = 1$, we choose the starting state at random, and both states have the same probability to be selected (p_0^H denotes the pre-match outcome probability for a home win, for example). Cases in which the sign is 1 or -1 for all three paths are contradicting, because the probabilities need to sum up to 1. Appendix Table A.5 provides descriptive statistics pertaining to how the states are represented in the sample; appendix Figure A.5 illustrates the distribution of the states during matches.

The state switching criteria in Table 2.2 based on changes in outcome probabilities are closely related to *Surprise*, as defined in Equation (2.1). Although the sign of *Surprise* is strictly positive, we can think of $S = 0$ and $S = 1$ as negative and positive *Surprise*, respectively, because the marginal effect is free to switch signs, depending on the best fit.

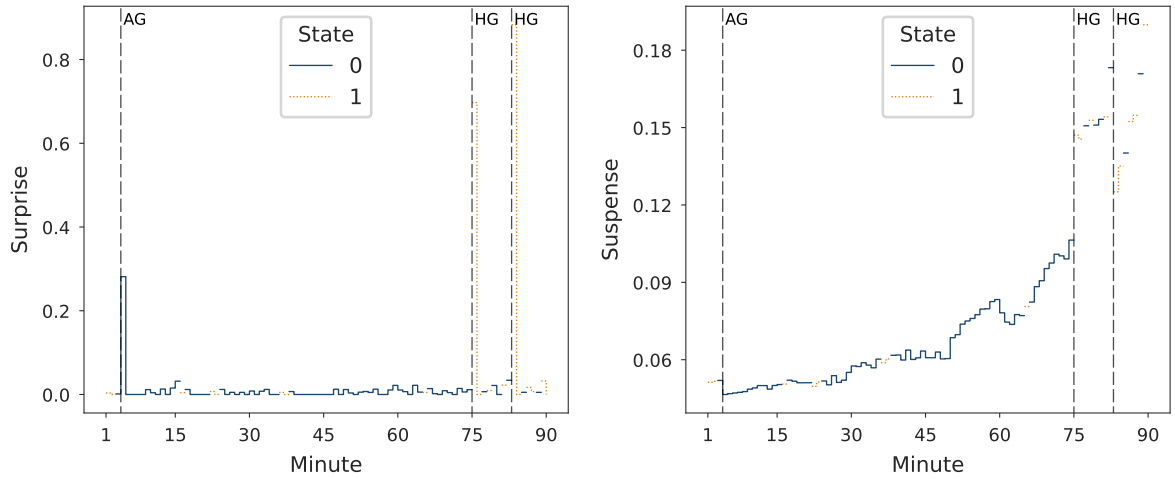
We do not explicitly distinguish between negative and positive *Suspense*. Instead, we assume that people are affected by *Suspense* regardless of the match event and the team they support, because it is a phenomenon referring to an outcome that is yet to happen, and thus partisanship should not create asymmetric responses. This implies that, independent of any preferences about

the timing of the resolution of uncertainty during the match (Palacios-Huerta 1999; Dillenberger 2010), realized relief after a tension phase is likely important for defining the emotional direction (i.e., negative or positive) rather than the expectation. Consequently, the emotional state in which people experience *Suspense* in t depends on the terminated and non-stochastic cue *Surprise*.

2.2.3.2 Visualization

In this section, we explicitly consider the entire stadium, excluding away fan areas. This sample is the largest possible block of data, accounting for emotional direction, and therefore smooths potential idiosyncrasies. Figure 2.2 reveals how cues reflect key match events like goals (*HG* and *AG*), which change the outcome probability of the match. The data depicted by Figure 2.2 refers to the same match between VfB Stuttgart and SC Freiburg in Figure 2.1.

Figure 2.2
Emotional Cues for Exemplary Match



This figure shows the development of emotional cues over time for an exemplary German first division soccer match between VfB Stuttgart and SC Freiburg on February 3, 2019 at 6:00 PM. Although VfB Stuttgart is the favorite, both cues exhibit a smaller jump when the away team scores at the beginning of the match (denoted *AG*) compared with the home team goals at the end (denoted *HG*). Key events like goals are more decisive when less time remains. Goals by the away team evoke the negative emotional state 0, from the perspective of home fans.

As we can see, *Surprise* as an adaptive variable returns to its pre-goal level very quickly after Freiburg scores in the first half, due to the varying reference point p_{t-1}. A goal for SC Freiburg (away) leads to the negative state 0 for home team fans. Stuttgart was favored in the match, so an away goal changes the match situation considerably. However, with a lot of time remaining, the jumps are higher when Stuttgart first equalizes and then breaks the tie deep in the second half.

For *Suspense*, the obvious pattern is its characteristic upward trend over time. The less time that remains, and the more information that is revealed, the greater the effect of a large variance of potential events in the next period; every goal becomes more crucial with respect to the final result.

2.3 Model

2.3.1 Formulae and Dimensions

To deal with the longitudinal multilevel data, we follow Asparouhov, Hamaker, and Muthén (2017) in determining

$$Y_{it} = Y_{1it} + Y_{2i} \ , \tag{2.4}$$

where *i* is the match index, and the match minute is captured by *t*. That is, Equation (2.4) represents the orthogonal decomposition of the dependent variable Y_{it} into two components, in which Y_{2i} (between-level) is the time average of Y_{it}, and Y_{1it} (within-level) is oscillating around Y_{2i}, such that $Y_{1it} = Y_{it} - Y_{2i}$. The explained variable $\boldsymbol{Y}$ is a matrix that comprises the number of beverage sales across matches in its *N* rows and the number of units sold over time, i.e., during the matches in it's *T* columns. The same applies to $\boldsymbol{Y_1}$. In contrast, $\boldsymbol{Y_2}$ is a $N \times 1$ vector, reflecting the lack of time specificity at level 2.

Having introduced the main components of the model, we can establish the within-level equation as

$$\begin{aligned}[Y_{1it} \mid S_{it} = s] = \nu_{1i} + \sum_{\mathrm{l}=1}^{L} Y_{1i(t-l)}\phi_{1sl} + \sum_{l=0}^{L} \left(\boldsymbol{X_{1i(t-l)}\beta_{1sl}} + \boldsymbol{Z_{1i(t-l)}\gamma_{1sl}}\right) \\ + W_{1it}\psi_{1ts} + \varepsilon_{1it} \ .\end{aligned} \tag{2.5}$$

Then at the between-level, we set up

$$\nu_{1i} = \boldsymbol{Z_{2i}\gamma_2} + \upsilon_{2i} \ ,$$

which is a random intercept model. In other words, mean sales vary between matches, which we explain with a fixed-effects vector ($\boldsymbol{Z_{2i}}$) and a random component (υ_{2i}).

We choose hierarchical modeling over a classical fixed-effects estimator because it offers a higher number of degrees of freedom and thus increased model precision. The slope parameters are more flexible, in the sense that they are not tied to a single intercept. The first-difference estimator, as common alternative, is not suitable, due to the trend-stationary dependent variable; to use it, we would have to take multiple differences to obtain a stationary time series, and data are costly in our rather small sample. Instead, we introduce match minute dummies, using $\boldsymbol{W_1}$, to capture general sales patterns, unrelated to events in the match (see Section 2.2.2).

Intercept $\boldsymbol{\nu_1}$ and random term $\boldsymbol{\upsilon_2}$ are $N \times 1$ vectors; the match minute dummy vector $\boldsymbol{W_1}$ is $N \times T$ and its coefficient vector $\boldsymbol{\psi_1}$ is $1 \times T$. Cues collected in $\boldsymbol{X_1}$ ($N \times T \times G_{1,1}$) share a marginal effect object $\boldsymbol{\beta_1}$ with dimensions $G_{1,1} \times 1$. The fixed-effects in $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ are $N \times T \times G_{1,2}$ and $G_2 \times N$ objects, such that $G_{1,2}$ and G_2 are the number of covariates in $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$. The marginal effect objects ϕ_1, $\boldsymbol{\gamma_1}$, and $\boldsymbol{\gamma_2}$ have corresponding dimensions: 1×1, $G_{1,2} \times 1$, and

$G_2 \times 1$, respectively. Finally, the residual $\boldsymbol{\varepsilon}$ is a $N \times T$ matrix[13].

As Equation (2.5) illustrates, the whole model is conditional on the state $S_{it} = s$. Lagged marginal effects thus apply to sales in t, given a certain state s, irrespective of whether the system was in a different state in the previous period.

2.3.2 Main Specification

For the main model specification, we use 98 matches from 6 seasons, spanning from 2013/14 to 2018/19.[14] We limit the set to league matches (i.e., Bundesliga and 2. Bundesliga, which are Germany's first and second divisions) to keep the sample homogeneous with respect to length and point in time at which the match outcome is determined. Cup matches, with potential overtime and penalty shootouts, thus are excluded. In addition, we do not consider stoppage time, so each time series lasts 90 minutes and has the exact same length for all matches. Next, we eliminate 4 high risk matches, for which the league decided that no alcohol, or only light beer, would be sold. The dependent variable includes beer sales and this main model specification refers to the entire stadium except for away fan areas.

2.3.2.1 Covariates

We control for weather conditions and substitutions at the within-level. For example, rain could simultaneously alter both the relative strength between teams and drinking behavior. Also, except for the potentially negligible effect of the entering player, there can be no updates to outcome probabilities during substitutions. That is, the systematic absence of match key events like goals and red cards during substitutions clearly affects outcome probabilities and simultaneously offers fans an opportunity to go for a drink. Therefore, we add substitutions as covariates. Other theoretically relevant variables at the between-level include the total number of spectators and the price of beverages, which can improve model precision. For a summary of all variables, see appendix Table A.3.

2.3.2.2 Priors

For the main specification, we use several prior distributions. Without any theoretical guidance related to most of the parameters, and considering that the level-1 sample size which is not extremely small, we apply weakly informative priors in most cases. First, we choose

$$\boldsymbol{\varepsilon} \sim \mathcal{N}\left(0, \sigma^2\right) \text{ with } \sigma \sim \mathrm{C}^{+}\left(0, 1\right) \ ,$$

[13] Throughout this paper, we represent vectors, matrices, and higher dimensional objects with bold letters to distinguish them from scalars.

[14] A regular season consists of 34 matches or 17 home matches per season. Subtracting the 4 high risk matches we exclude thus yields 98 matches.

where C^+ denotes the Half-Cauchy distribution. Second, we specify

$$\boldsymbol{\upsilon_2} \sim \mathcal{N}\left(\mathcal{N}\left(0,1\right), \mathcal{N}^+\left(0,1\right)\right) \ ,$$

where N^+ is the truncated normal distribution. Third, we use a continuous version of the spike-and-slab prior (Mitchell and Beauchamp 1988; George and McCulloch 1993; Ishwaran and Rao 2005), such that

$$\begin{aligned} \theta_j \mid \lambda_j, c &\sim \mathcal{N}\left(0, c^2\lambda_j\right) \text{ where } \theta_j \in \{\boldsymbol{\phi_{1s}}, \boldsymbol{\beta_{1s}}, \boldsymbol{\gamma_{1s}}, \boldsymbol{\psi_{1s}}, \boldsymbol{\gamma_2}\} \ , \\ \text{such that } \lambda_j &\sim \text{Beta}(0.5, 0.5) \text{ for } j = 1, \ldots, J, \text{ and} \\ c^2 &\sim \Gamma^{-1}(3,1) \ . \end{aligned} \tag{2.6}$$

Except for $\boldsymbol{\psi_1}$ (no lags) and $\boldsymbol{\gamma_2}$ (no states), all vectors affected by the spike-and-slab design contain lagged values and are state-dependent. The feature space is comparably large, and it is difficult to foresee which variables explain variation in the dependent variable, so we choose to regularize all parameters, which also helps prevent overfitting.

The spike-and-slab prior uses a binary indicator variable λ that determines whether the coefficient is zero (i.e., it comes from a "spike", $\lambda \to 0$) or non-zero ("slab", $\lambda \to 1$). The slab width is denoted by the parameter c. It is generally not flexible to fix c, so a common practice entails placing a hyperprior on c to allow for heavy-tailed slabs. The sparsity information of the coefficient vector can be controlled with the parameters of the Beta prior in Equation (2.6).

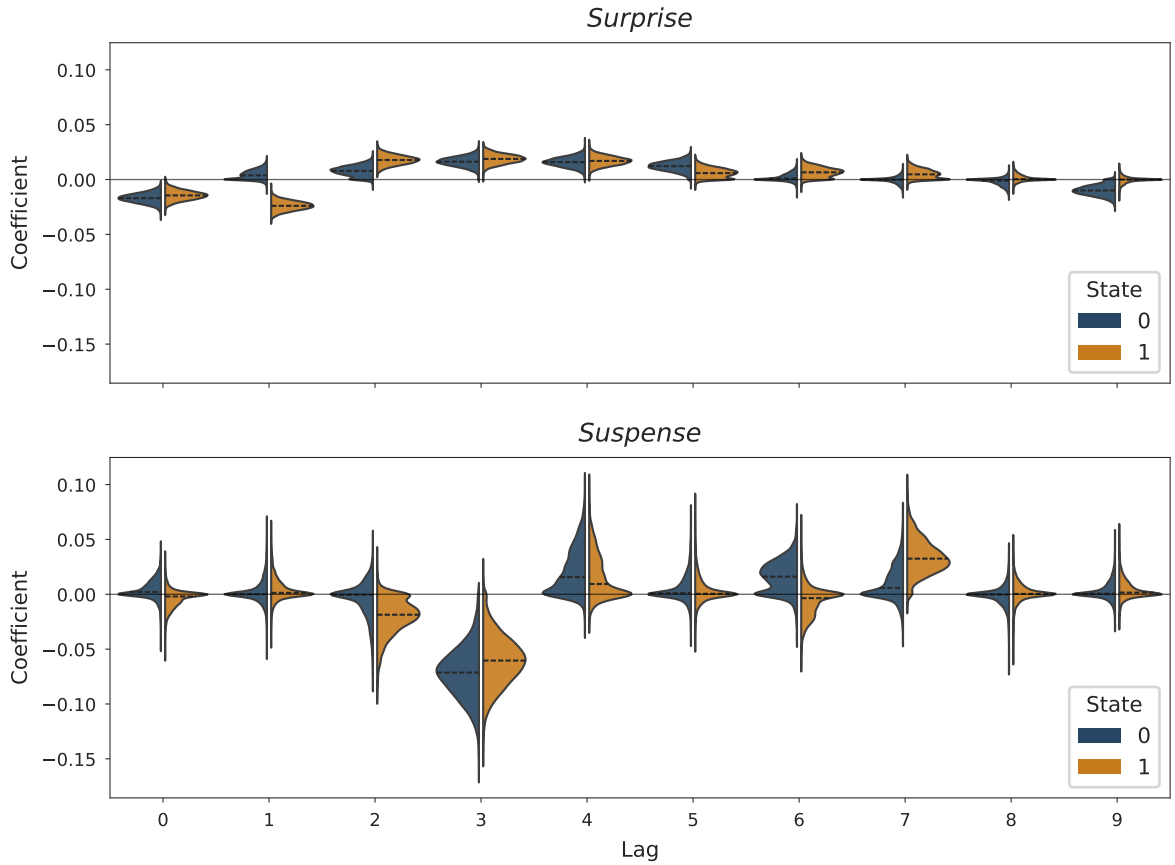
2.4 Empirical Results

2.4.1 Main Specification

Figure 2.3 depicts the main model posterior distributions of marginal effects $\boldsymbol{\beta_{11}}$ (*Surprise*) and $\boldsymbol{\beta_{12}}$ (*Suspense*) separately for negative (0) and positive (1) states. In addition to the contemporaneous effects ($l = 0$), we present all effects up to nine minutes lagged ($L = 9$).

The statistically and economically significant effects for *Surprise* and *Suspense* on the number of beer sales are most pronounced for $l = 3$. We use the median (dashed line) of the $l = 3$ marginal effect in positive state 1 for *Surprise* (upper panel) as an example; it is equal to 0.0190. A one standard deviation increase in positive *Surprise* then increases the conditional mean of the number of beers sold during minute 3 after by approximately 1.9%, ceteris paribus. If we take the average value of 85.5 from Table 2.1 as a basis, this increase corresponds to roughly two additional beers. At first glance, this effect might not seem notable, except that—as the exemplary match between VfB Stuttgart and SC Freiburg in Figure 2.2 reveals—the leading goal for Stuttgart increases *Surprise* more than 0.8. Considering the standard deviation of 0.07 for *Surprise* in state 1 (see appendix Table A.7), the marginal effect in this numerical example increases by a factor of approximately 10, such that 17 additional beers are sold (20% of 85.5).

Figure 2.3
Main Model Posterior Distributions



This figure shows the posterior distributions of the coefficients for *Surprise* and *Suspense*, based on the main specification. The posteriors are depicted for all lags in both states (negative state 0 and positive state 1). For example, a positive effect for *Surprise* appears in both states 3 minutes lagged (see distributions for the third lag in the upper panel). The median depicted by the dashed line, equal to 0.0190 for positive state 1, implies that a one standard deviation increase in *Surprise* in positive state 1 increases the conditional mean of the number of beers sold during minute 3 after a key match event by approximately 1.9%, ceteris paribus.

This effect for $l = 3$ corresponds to just a single minute; the total effect is longer lasting. For example, we find clearly positive posterior distributions for *Surprise* when $l = 2, l = 4$, and $l = 5$ as well. Besides, not all people in the stadium drink alcohol (e.g., children, teetotalers, pregnant women).

The coefficients for *Suspense* (lower panel) follow the same reasoning, but the main effect when $l = 3$ is clearly larger than that of *Surprise* for both states. However, due to the lower relative increase of *Suspense* for the second home goal though (Figure 2.2), we calculate that approximately 10 beers less are sold in $l = 3$ at state 1, if we follow the same procedure described for *Surprise*. Also, the posteriors are wider (less precise), so our conclusions with respect to *Suspense* reflect greater uncertainty. Overall though, we yield several notable findings.

First, both *Surprise* and *Suspense* initially decrease beer sales. While this pattern is immediately evident for *Surprise* and longer lasting in the positive state 1 (i.e., for $l = 0$ and $l = 1$) compared to the negative state 0 (only for $l = 0$), it occurs later for *Suspense* and is mainly evident

for $l = 3$ in both states.

Second, in the longer run, both *Surprise* and *Suspense* unfold positive effects at decreasing margins. Again, the switch from negative to positive posteriors occurs earlier for *Surprise* compared to *Suspense*. Likewise, we observe this switch for the negative state 0 earlier compared to the positive state 1. Overall, and even though this pattern is less clear for *Suspense* compared to *Surprise*, it might suggest that people respond more quickly to negative (rather than positive) signals in our setting.

Third, when aggregating all effects over $l = 0$ to $l = 9$, we do observe a clear positive net effect for *Surprise* in both states. For *Suspense*, however, the positive effects in the longer run can hardly compensate for the strong negative effects we observe for $l = 3$, thus suggesting rather a small negative net impact (if any).

Appendix Table A.8 contains the posteriors of all variables used in the main specification. The most important covariates for both states at the within-level are the first half dummy and the lagged dependent variable, which both control general time series pattern. The first half dummy is contemporaneously negatively associated with beer sales. The coefficients for the lagged dependent variable are positive and mostly decreasing in size with higher-order lags. Furthermore, we identify some isolated non-zero posterior distributions across lags in positive state 1 like the coefficients related to the number of hot wine punch units sold (positive) and the dummy for home substitutions (negative). Alcohol consumption in negative state 0 seems rather unaffected by covariates other than the away substitutions (one isolated negative effect across lags)[15].

At the between-level, the number of spectators (positive), the dummy for a first division match (negative), some dummies for kickoff times (positive), and the pre-match probability that the joint score exceeds 4.5 (positive) offer relatively great explanatory power with regard to the number of beer sales. In addition, the prices for beer and hot wine punch (both positive), the pre-match probability that the joint score exceeds 0.5 (negative), the dummies for the match taking place on Saturday (positive) or Sunday (negative), and the number of active cashpoints (positive) exhibit non-zero but more moderate coefficients. For a visualization of between-match variation in the number of beer sales, please see appendix Figure A.6[16].

[15] Apparently, when people are willing to buy beer, they also turn to hot wine punch, or else hot wine punch drinkers drive excess demand. Furthermore, substitutions of players during the match do not seem to serve as opportunities to go for a drink. In contrast, the negative association of home substitutions suggests fans are interested in who is being substituted or that substitutions are systematically conducted at crucial points of the match.

[16] We explain the positive coefficients of prices (beer and hot wine punch) with a combination of inflation and increasing sales figures over the years. Moreover, the negative association for first division matches might suggest that the prestige of matches (if relevant at all) is outweighed by more positive experiences in the second division, in which the average winning probability for VfB Stuttgart is significantly higher.

2.4.2 Alternative Specifications

We start assessing the robustness of our main findings by exploring alternative priors and extending the number of seasons under consideration.

2.4.2.1 Priors

Prior distributions can strongly determine the empirical results in Bayesian estimation, which may be a particular concern in relation to our comparably small sample setting with a lot of weakly informative priors. Therefore, we test different distributions to determine if the results change. In particular, we replace the spike-and-slab prior of the main specification with another regularization prior, namely, the regularized horseshoe prior (Piironen and Vehtari 2017). In turn, the definitions become

$$\begin{aligned}
\theta_j \mid \lambda_j, \tau, c &\sim \mathcal{N}\left(0, \tau^2 \tilde{\lambda}_j^2\right) \text{ where } \theta_j \in \{\boldsymbol{\phi_{1s}}, \boldsymbol{\beta_{1s}}, \boldsymbol{\gamma_{1s}}, \boldsymbol{\psi_{1s}}, \boldsymbol{\gamma_2}\} \text{ ,} \\
\text{with } \tilde{\lambda}_j^2 &= \frac{c^2 \lambda_j^2}{c^2 + \tau^2 \lambda_j^2} \text{ ,} \\
\lambda_j &\sim \mathrm{C}^+(0,1) \text{ for } j = 1, \ldots, J \text{ ,} \\
\tau &\sim \mathrm{C}^+(0, \tau_0^2) \text{ using } \tau_0 = \frac{\frac{J}{2}}{J - \frac{J}{2}} \frac{\sigma}{\sqrt{n}} \text{ , and} \\
c^2 &\sim \Gamma^{-1}(3,1)
\end{aligned} \tag{2.7}$$

relative to Equation (2.6). The regularized horseshoe shrinks large signals (coefficients far from zero) and therefore prevents flat posterior distributions. The largest parameters are regularized according to $\mathcal{N}(0, c^2)$. For small coefficients (in absolute value), the horseshoe estimator applies (Carvalho, Polson, and Scott 2009), because $\tilde{\lambda}_j^2 \to \lambda_j^2$. Generally, we follow Piironen and Vehtari (2017) with respect to distributional choices and their values. We assume a share of 50% of all features to be relevant (cf. $J/2$ in Equation (2.7), in which J denotes the size of the feature space).

Moreover, to investigate the effect of regularization, we also run a model with unregularized prior distributions, such that we consider

$$\phi_1, \boldsymbol{\beta_1}, \boldsymbol{\gamma_1}, \boldsymbol{\psi_{1s}}, \boldsymbol{\gamma_2} \sim \mathcal{N}\left(\mathcal{N}(0,1), \mathcal{N}^+(0,1)\right)$$

in contrast with the main specification. Overall, we confirm our three main findings (appendix Figures A.7 and A.8).

2.4.2.2 Seasons

Most of our data cover two additional seasons, that is, 2011/12 and 2012/13. Unfortunately, we lack bookmaker in-play odds for these two seasons (to the best of our knowledge). As a robustness check, we leverage these two seasons and predict the in-play odds for 2011/12 and 2012/13, using a feedforward neural network (appendix Table A.2) in TensorFlow (Abadi et al. 2015),

which entails an added set of 34 league matches. We describe the exact imputation procedure in Appendix A.1.2.2. With this larger sample, we reaffirm our three main findings (appendix Figure A.9).

2.4.3 Are Updating Beliefs Important?

Buraimo et al. (2020) propose another cue, *Shock*, which is similar to *Surprise* but applies to a different reference point. In concrete terms, the reference point in

$$Shock_t = \left(\sum_{m \in H,D,A} [p_t^m - p_0^m]^2 \right)^{0.5} \tag{2.8}$$

is p_0, which refers to the fixed pre-match outcome probability at the start of the match before any in-play information is revealed. Therefore, *Shock* is backward looking like *Surprise* but persistently affected by key match events. Often, *Shock* exhibits a long-term upward or downward trend, reflecting the low scores that characterize soccer. The more time that goes by, the lower the likelihood for changes, which reinforces the direction in which *Shock* already moves. When including *Shock* as regressor, our main findings hold. Moreover, *Shock* does not appear relevant for explaining the number of beer sales, across all combinations of states and lags (appendix Figure A.10).

2.4.4 Does Involvement Matter?

Including the whole stadium (cf. away fan areas) creates a trade-off between sample size and accuracy. Most transactions considered in the main specification involve home fans, but there is a margin for diffusion, because in some stands, away fans can mingle with home fans. To gain greater accuracy with respect to the direction of the effect of emotional cues on beer sales, we thus focus on the Cannstatter Kurve only, which represents a smaller but more homogeneous sample of home fans. We adjust variables related to the target stand accordingly. For example, the number of cashpoints is limited to those in the Cannstatter Kurve only, rather than the whole stadium. Apart from the slightly weaker effects for *Surprise* and the less clear temporal pattern for the negative state 0 and *Surprise*, our main findings remain (appendix Figure A.11).

2.4.5 Is it About Alcohol Use?

2.4.5.1 Shandy

In the main specification, we do not combine beer and shandy, due to their distinct alcoholic content and taste. If we take shandy as dependent variable, we can test if people consciously turn to alcohol to deal with sport-related emotions during the match, such that we expect smaller effects (in absolute terms) or generally less clear signals for the cues in this analysis. With the exception of *Surprise* when $l = 4$ (positive) in negative state 0, no effects differ unambiguously from zero

(appendix Figure A.12). Therefore, we do not find evidence for any of our main findings with this alternative dependent variable.

2.4.5.2 Soft Drinks

We also consider soft drinks as dependent variable, using aggregated sales of Coca-Cola, Coca-Cola Zero, Fanta, Mezzo Mix, Sprite, and Lift Apfelschorle (apple spritzer). Again, we seek to determine if people specifically turn to alcohol in emotional situations during the match or instead buy beverages in general when the match is not suspenseful, for example. In positive state 1, no effects unambiguously differ from zero. Furthermore, we only find a similar pattern to beer sales for *Surprise* in negative state 0. However, when aggregating all effects over $l = 0$ to $l = 9$, the net increase (if any) in soft drink sales is small (appendix Figure A.13).

2.5 Discussion

2.5.1 Observed Behavior

In conclusion, we find *Surprise* and *Suspense* to initially decrease beer sales while unfolding some positive effects at decreasing margins in the longer run. Overall, these patterns seem intuitively plausible. For *Surprise*, the immediate decrease in beer sales can result from tumult or distraction in the stands and at the cashpoints. After a key match event, people need time to stand up, go to the cashpoint, maybe wait in the line, and finish the transaction. For *Suspense*, the (compared to *Surprise*) slightly delayed decrease in beer sales might suggest that fans postpone consumption in suspenseful phases, then react to match key events when the suspense gets resolved. Likewise, saturation effects might explain why the positive effects we find in the longer run get smaller toward the end of the observation period: Once a fan has bought a beer, they are not likely to buy another one immediately.

When aggregating all effects over $l = 0$ to $l = 9$, we do observe a positive net effect for *Surprise* in both states. Because we find almost no effects when we replace beer by shandy and much weaker effects on soft drink purchases as the dependent variable, we conclude that people might turn specifically to alcohol in emotionally charged situations[17]. Overall, this interpretation is in line with psychological theories. Greeley and Oei (1999) provide an overview of the tension reduction theory (TRT), according to which negative (emotional) stress/tension increases alco-

[17] One might argue that the (mainly) zero-centered posteriors for *Surprise* and *Suspense* when analyzing shandy and soft drink sales may be related to differences in consumers. If, for instance a larger share of women prefer shandy or a larger share of children prefer soft drinks, and they would respond differently to emotional cues, selection effects might confound our results. Despite the heterogeneity in consumer groups between beer, shandy, and soft drink buyers, we should observe similar effects for all beverages if there is just a general pattern of opportunity (e.g., spectators believe that it would be a good time to get a drink immediately after a key match event because there is less likelihood of another event immediately following). However, we do not find such similarities across beverages.

hol use, because individuals consume alcohol for its stress-response dampening effects. Cooper et al. (1995) emphasize the importance of considering different psychological motives for drinking. While TRT focuses on alcohol as a potential moderator of negative affective states, they find alcohol to be used to regulate (and facilitate) positive emotions as well.

Moreover, the weaker effects for *Surprise* in the die-hard fans specification likely reflects their greater emotional involvement. A kind of emotional threshold may need to be surpassed to spark a consumption decision. This threshold could depend on the composition of the group of consumers. If the average die-hard fan is more involved than other fans and imposes a higher emotional threshold to be passed, because their opportunity costs of not watching are higher, a certain amount of *Surprise* would lead to comparably fewer beer sales. Furthermore, the die-hard fans in the standing room might systematically turn to mobile beer runners; due to the higher opportunity costs and in response to the poorer accessibility of cashpoints. This behavioral choice decreases the observed number of beer sales. As a consequence, the effects of the cues would be underestimated for die-hard fans.

In contrast to *Surprise*, we observe a small yet rather negative net effect for *Suspense* when aggregating all effects over $l = 0$ to $l = 9$. In principle, this can be attributed to the fear of missing important plays when *Suspense* increases and is in line with Liu, Shum, and Uetake (2020) who find that viewers of baseball games are much more attentive even to commercials in more suspenseful phases. Because, we do not find such effects when we replace beer with shandy or soft drinks, it could also well reflect the desire to alleviate boredom by drinking alcohol when *Suspense* is low. In other words, fans might systematically turn to alcohol (social drinking) when they are bored to make the experience more entertaining (Patrick and Schulenberg 2011). Independently of the concrete underlying mechanism for the (small) negative net effect of *Suspense* which we observe, it lends some credibility to the claim by Wood, McInnes, and Norton (2011), who relate traffic fatalities after close matches to aggressive driving due to an increased testosterone level rather than drinking during the match.

2.5.2 Identification

Identification of the aforementioned effects relies on the assumption that we are not missing variables violating strict exogeneity at the within-level, i.e., systematically changing both match outcome probabilities (i.e., the cues) and the number of beverage (alcohol) sales. At level 1, the missing control for different numbers of spectators during a match could be a source of endogeneity. If a very unbalanced outflow of either home or away team fans occurs, fan support for each team changes, relative to the other team, over the course of the match. The number of spectators at level 1 thus is an omitted variable that changes both the cues (outcome probabilities) and the number of beer sales. However, we assume that dynamic fan support from non-die-hard fans (die-hard fans will stay until the end of the match anyway), is responsive to the match, not the other way around.

Moreover, toward the end of a match, a systematic, not beverage-related difference might arise, such as more purchases of bottled water instead of beer. A bottle is easier to take home than a cup. In this case, water consumption would not be an appropriate control variable for baseline consumption. However, we consider this issue largely irrelevant, because all nonalcoholic beverages are poured into cups in our sample.

In addition to fixed cashpoints, mobile cashpoints and beer runners are available, whose locations and numbers vary across matches. Beer runners (exclusively) sell beer in the stands. After the match, they clear their transactions at a single master cashpoint. Before the 2017/18 season, mobile cashpoints deposited their cumulative transactions with the master cashpoint too, but after this season, each of them was equipped with their own cash register systems. Because sales through mobile cashpoints are labeled though, we can segregate them. Our main dependent variable does not refer to any sales through mobile cashpoints or beer runners. Although we lack minute-by-minute data about beer runner sales, the club indicated that they could be quite substantial. Because we can only control for the availability of beer runners, which reflects heterogeneity in supply, we assume that the null aggregated post-match numbers imply their absence.

Furthermore, adding beer runner sales at level 1 seems unlike to systematically adjust differently to emotional cues and thus alter our main results substantially, but the estimates could change. That is, we might identify stronger immediate effects, because fans do not have to walk to the cashpoint; the opportunity costs for buying from a runner also are lower. A longer wait time also could shift observed consumption back in time. Thus, our findings represent lower bound estimates, because they exclude transactions by people with higher opportunity costs.

Finally, goals change the match outcome and could also affect alcohol use, whether emotionally or indirectly due to the occasion, for example. Consequently, one might wonder whether additionally controlling for goal dummies is relevant in our setting. We argue, that *Surprise* and *Suspense* better reflect the match situation (probabilities rather than each goal treated equally) and are more fine-tuned. For example, if we assume that the effect of a goal reflects both the occasion and an emotion, and they both lead to alcohol use, the occasion part (long versus short lines) and the level of emotion depend on the match situation, such that they still can be better captured by *Surprise* and *Suspense* rather than a rough goal dummy, which exists irrespective of the score. Also, a specification additionally including goal dummies likely undermines the effect sizes of the coefficients for *Surprise* and *Suspense* due to multicollinearity. As a result, we refrain from controlling goal dummies in our models.

In summary and to emphasize missing concerns regarding both reverse causality issues or omitted variable bias at the within-level, we use the term "effect" instead of "association" for our key regressors. For further discussions and comprehensive details on the covariates, data engineering, and the algorithm used for estimation, please see Appendix A.1.3 to A.1.5 and appendix Table A.3.

2.6 Summary of Findings

We investigate the effects of emotional cues on alcohol sales during soccer matches, using real data pertaining to both emotional cues and alcohol use. Overall, our findings provide some robust evidence for emotional drinking. First, we find stable net effects for *Surprise* increasing beer sales. Because we find almost no effects when we replace beer by shandy and much weaker effects on soft drink purchases as the dependent variable, we conclude that people seem to turn specifically to alcohol in emotionally charged situations. This is in line with psychological theories suggesting that alcohol may facilitate positive emotional experiences and reduce tension or stress. Second, the revealed temporal pattern of the effects of *Suspense* on beer sales (which are initially negative and then positive) can be attributed to the fear of missing important plays when *Suspense* increases. Because we do not find this pattern when we replace beer with shandy or soft drinks, it could also reflect the desire to alleviate boredom by drinking alcohol when *Suspense* is low. Finally, while we find that people tend to respond more quickly to negative (rather than positive) signals, our findings do not differ significantly between positive and negative emotional states.

Broadly speaking, by providing empirical evidence for the influences of emotions experienced during the decision-making process for consumption decisions, we highlight the importance of short-term emotions in determining economic behavior. More specifically, we contribute to the sparse empirical literature on emotions and alcohol use as a mass phenomenon.

While one might be inclined to view the chosen setting of this study as not providing representative evidence, it is important to note that the key issue for any such choice should be the ability to observe and isolate the effects of interest (Falk and Heckman 2009). Furthermore, "Angrist and Pischke (2010) remind us that empirical evidence on any causal effect is always local." (Palacios-Huerta 2023, p. 5). We thus hope to encourage economists and psychologists to model utility as a function of both preferences and emotions and to further test reference point-dependent constructs of emotions in sports settings in the future.

A Appendix

A.1 Additional Information on Estimation

A.1.1 Calculation of Emotional Cues

To calculate the emotional cues *Shock*, *Surprise*, and *Suspense*, we need match outcome probabilities, because all three emotional cues reflect in-play probabilities for the three potential outcomes *H* (home win), *D* (draw), and *A* (away win) in each minute of the match. Furthermore, *Shock* and *Surprise* call for pre-match outcome probabilities, again for *H*, *D*, and *A*, in that the reference point for *Shock* is p_0 (Equation (2.8)) and that for *Surprise* in match minute 1 requires p_0 as reference point as well; *Suspense* also needs a starting point to simulate different potential match outcomes, as we clarify subsequently. The simulation for *Suspense* needs pre-match probabilities for the combined score to exceed a certain amount of goals.

We use betting odds to approximate these probabilities. In particular, we apply in-play odds from NowGoal (bookmaker Crown) and pre-match closing 1X2 (*H*, *D*, *A*) odds, as well as pre-match closing over/under odds[18] from OddsPortal (the sources are detailed in Table A.4). Our use of in-play odds diverges from the approach of Buraimo et al. (2020), who estimate outcome probabilities based on the time remaining, the current score, and the number of red cards. That is, they simulate different match outcomes by using historic goal distributions for the respective league and use the relative frequencies for *H*, *D*, and *A* from the scorelines at the end of the match as outcome probabilities. In contrast, we obtain the implied probabilities by taking the reciprocal of the odds, which usually add up to a value greater than 1, called the over-round. Then, we remove the difference between this over-round and 1 (considered the bookmaker commission or margin) that is inversely proportional to the size of the odds so that the resulting pseudo-probabilities, which we call normalized odds, sum up to 1.

The unconditional and conditional probabilities in *Suspense* (Equation (2.2)) outline hypothetical scenarios that cannot be directly obtained from bookmaker odds. We adopt the simulation method suggested by Buraimo et al. (2020) to calculate $p^m_{t+1} \mid p^{HG}_{t+1}$, and $p^m_{t+1} \mid p^{AG}_{t+1}$, but refine their approach.

Buraimo et al. (2020) estimate the scoreline of each encounter by exploiting a priori information about the teams' strengths, past performances, coaches, venue, and all other factors that have predictive power for the score. This information is embedded in pre-match odds. We use the last odds quoted before the match starts, referred to as closing odds. Closing odds most precisely reflect the market's assessment of the match outcome by accommodating more information than any other, previously quoted odds. Then we take the average across multiple bookmaker odds, to avoid possible idiosyncrasies linked to individual bookmakers. Thus, the data contain closing odds from up to 36 bookmakers and over/under odds from up to 34 bookmakers. To increase

[18] Over/under odds work as follows: If a bettor thinks there will be one or more goals for a given match, they will bet on over 0.5 and win the wager if at least one team scores. If they anticipate that there will be not more than two goals, they place their money on under 2.5.

the number of over/under odds, we also collect data about the combined score to exceed/subceed the statistics ± 0.5, ± 1.5, ± 2.5, ± 3.5, ± 4.5, and ± 5.5 goals.

Although Buraimo et al. (2020) proposes independent Poisson distributions to estimate the number of goals scored by each team, we use a bivariate Poisson distribution to estimate the scoreline and thereby relax the harsh independence assumption. Consequently, we assume that the number of goals scored by the home team, denoted by the random variable R_1, and the number of goals scored by the away team, denoted by the random variable R_2, are jointly Poisson distributed such that

$$P\left(R_1 = k_1, R_2 = k_2\right) = \exp\left(-\delta_1 - \delta_2 - \delta_3\right) \frac{\delta_1^{k_1}}{k_1!} \frac{\delta_2^{k_2}}{k_2!} \sum_{k=0}^{\min(k_1,k_2)} \binom{k_1}{k} \binom{k_2}{k} k! \left(\frac{\delta_3}{\delta_1 \delta_2}\right)^k$$

in which k_1 and k_2 are realizations of R_1 and R_2, respectively. Moreover, δ_1 and δ_2 refer to the scoring rates of the teams and

$$Cov(R_1, R_2) = \delta_3 \ .$$

If $\delta_3 = 0$, we end up at the independent Poisson distributions.

We use this in-play model to generate the probabilities for every scoreline of a given match. In addition, we calculate the probabilities for the match outcomes *H*, *D*, and *A* by summing the scoreline probabilities. Note that we restrict k_1 and k_2 to a maximum of 10, so we estimate the joint probabilities of all reasonable hypothetical scorelines for which the number of goals of each team does not exceed 10, such that $P(R_1 = 10, R_2 = 10)$ is the last probability estimated.

To estimate the scoring rates δ_1, δ_2, and δ_3, we minimize the squared difference between the transformed bookmaker odds and our estimated probabilities from the in-play model. The function we minimize is

$$F = \sum_{m \in H,D,A} \left(\boldsymbol{q}_m - \boldsymbol{p}_m\right)^2 ,$$

where $\boldsymbol{p}_m$ is a vector of probabilities obtained from our model, and $\boldsymbol{q}_m$ denotes the vector of normalized pre-match match win and over/under bookmaker odds. The match win odds provide information about which scoring rate should be larger; the over/under odds help reveal the exact size of the scoring rates when minimizing. For a match in which VfB Stuttgart is the clear favorite, for example, their scoring rate is naturally higher. However, the information on who is expected to win does not suffice to determine the magnitude of δ_1, δ_2, and δ_3.

In a next step, we distribute the estimated scoring rates $\delta_1 + \delta_3$ and $\delta_2 + \delta_3$ across minutes of the match. With the assumption that goals are uniformly distributed throughout a match, we could evenly split up the scoring rates, resulting in $\delta/90$ for each minute. However, scoring rates are not constant; more goals are scored toward the end and fewer goals at the beginning. Buraimo et al. (2020) propose spreading the scoring rates in proportion to the empirical distribution of goals per minute, which they generate by gathering the timing of goals scored by many teams in the past.

Although this empirical distribution better captures the average scoring patterns, it generally fails to represent the scoring patterns of an individual team adequately, which depend on the team's individual strength, its way of playing, and its coach's philosophy. The teams at the tails of the distribution can reveal scoring behaviors far away from the average. Therefore, we propose team-specific empirical goal distributions to distribute the scoring rates across the minutes, reflecting all goals scored by a given team during all games played before the current encounter.

To guarantee a smooth distribution, we use a 10-minute moving average and linearly extrapolate missing values at the beginning of the time series. The underlying data for the empirical goal distributions come from kicker (see Table A.4) and cover seasons 2013/14 to 2018/19 (inclusive). That is, the goal distribution for VfB Stuttgart consists of all league goals scored during that period (considering league affiliation due to relegations and promotions). The same is true for all opponents, such that each team takes its own weighting for the individual scoring rate. The squad and the way teams play can change significantly over the years, so it would be preferable to use goal data from the last few matches but due to the limited number of matches per season, we need to extend the observation period. Otherwise, we end up with a very sparse empirical goal distribution, with no observed goals for many match minutes.

With the per minute scoring rates, we can simulate the number of goals occurring in each minute of the match. We draw from a Bernoulli distribution with success probability equal to the per minute scoring rate in t and sum the final scoreline that thus results. To account for red cards, we rely on Vecer, Kopriva, and Ichiba (2009) who find that a red card decreases the affected team's scoring rate by $2/3$, while the opposite team's scoring rate increases by a factor of 1.2. For each match, we repeat this simulation 100,000 times so that there are $90 \times 100,000 = 9,000,000$ simulations per match. Then we can determine the probabilities required for the calculation of *Suspense* in t by evaluating the probabilities for the match outcome, given that the home or the away team, respectively, scores in the next minute. Without any next minute in regular match time in the 90th minute, we cannot calculate *Suspense* for it, such that we drop observations for this time point from the sample.

A.1.2 Imputation of In-Play Probabilities

A.1.2.1 Seasons 2013/14 to 2018/19

The betting market is closed when odds are updated or liquidity is low, so data on in-play odds exhibit missing values by nature, which in turn lead to missing in-play outcome probabilities. For the affected time points, we cannot transform the odds to obtain the outcome probabilities.

We handle the missing in-play outcome probabilities for seasons 2013/14 to 2018/19, as used in the main specification, by predicting their values using gated recurrent units (GRUs) and a separate training sample of approximately 4000 matches, played during the seasons 2013/14 to 2019/20 in the German first and second division. These GRUs represent an extension to recurrent neural networks (RNNs) that can account for dynamic behavior in the data by processing

sequences of inputs.

As introduced by Cho et al. (2014), a GRU consists of two gates, an update gate and a reset gate. The update gate determines which new information is added and which is discarded, similar to the forget and input gate of a long-short-term-memory (LSTM); the reset gate determines how much past information to forget. Because the GRU lacks the LSTM's output gate, there are fewer parameters to train, which promises less computational effort with similar or better performance than LSTMs. Our GRU networks consist of an input layer, 4 hidden GRU layers, a hidden dense layer, and an output layer. Moreover, we make use of bias nodes and alternate activation functions in the hidden layers between rectified linear unit (ReLU) and hyperbolic tangent (tanh). Table A.1 displays the model configuration.

Table A.1
Gated Recurrent Unit Network Model Configuration

Layer	Input	Units	Bias	Kernel Init.	Bias Init.	Activation	L2 Kernel
input_layer	(None, 10, 7)	NaN	NaN	NaN	NaN	NaN	NaN
gru_1	NaN	128	1	GlorotUniform	Zeros	relu	NaN
gru_2	NaN	96	1	GlorotUniform	Zeros	tanh	NaN
gru_3	NaN	64	1	HeUniform	Zeros	relu	NaN
gru_4	NaN	32	1	GlorotUniform	Zeros	tanh	NaN
dense	NaN	32	1	HeUniform	Zeros	relu	NaN
output	NaN	3	1	GlorotUniform	Zeros	softmax	0.0010

This table presents the configuration for the GRU network used to impute missing in-play outcome probabilities for seasons 2013/14 to 2018/19. The separate training sample consists of approximately 4000 matches from seasons 2013/14 to 2019/20 in the German first and second division. For each minute we, train 1 GRU network. The GRU networks include an input layer, 4 hidden GRU layers, a hidden dense layer, and an output layer. ``Init.'' in columns 5 and 6 is short for ``Initialization'', and NaN (``Not a Number'') labels unspecified parameters. As the first entry in the second column ``(None, 10, 7)'' shows, we use a full batch approach with a time series length of 10 and 7 features, namely, number of goals scored and red cards received by each team in t, as well as transformed in-play odds for the match outcome. Furthermore, the activation functions for the hidden layers are ReLU and tanh, alternating (column 7). The output layer squeezes the data between 0 and 1, according to the softmax function and is regularized by an L2 penalty term (last column).

For each minute, we train one GRU network, separately. The features (see "7" in "(None, 10, 7)" - second column of Table A.1) include the number of goals scored and red cards received by each team in t, as well as transformed in-play odds for the match outcome prior to t (for $t = 1$, transformed pre-match odds are used). The input sequences comprise this information in the last 10 minutes (see "10" in "(None, 10, 7)" - second column of Table A.1). Consequently, we use the information in $t - 10$ to t to predict the probabilities in t. For the periods prior to the 10th minute, this sequence reduces to t_0 to t. Note that there are no betting odds for stoppage time, and NaN ("Not a Number") denotes parameters that are not specified.

In-play odds are available to us on a minute-by-minute basis. However, these odds are obviously not synchronized with the match minutes. If multiple odds updates occur per match minute, we use the most recent update. Should there be a goal or a red card in a certain minute but the odds remain unchanged, because they were placed before the match event, we replace the odds by NaN and predict it using a GRU network. The underlying assumption is that goals or red cards must change outcome probabilities and odds. For the whole match, the last updated

odds are also the last usable odds. We replace all odds after them by NaN, because we do not know whether the odds do not change due to the course of the match or because the markets are closed because the match is almost decided, for example.

We standardize all features so that they are on the same scale, such that each feature has zero mean and unit variance. Specifically, we standardize the data by fitting parameters on the training set, then reuse them to transform the test data. In this way, we ensure there is no test set information in the training process, and, assuming training and test data come from the same distribution, we obtain more precise estimates for the mean and variance in the test set, because of the larger sample size of the training set. We carry out this transformation for each minute, separately. The targets are a vector of normalized odds, such that they give rise to a regression problem. Therefore, and because we want to penalize large deviations more strongly, we choose the mean squared error (MSE) as the objective loss function that needs to be minimized during the training process.

For the optimization, we use the AMSGrad variant (Reddi, Kale, and Kumar 2019) of Adam (Adaptive Moment) Estimation (Kingma and Ba 2014), which features stochastic gradient descent, based on the adaptive estimation of first-order and second-order moments. It thus is appropriate for settings with a relatively large number of parameters.

The number of epochs depends on the current match minute and is evenly spaced over the interval $[250, 500]$, rounded to the nearest integer, so that the network for minute 1 uses 250 epochs, and the network for minute 90 uses 500 epochs, because the higher the match minute, the more missing values there are, creating inaccuracies in prediction that stack over time. The batch size is equal to the size of the training data, such that the full data set is processed as one chunk (see "None" in "(None, 10, 7)" - second column of Table A.1). To prevent overfitting, we impose an L2 penalty of size 0.001 on the weights of the output layer.

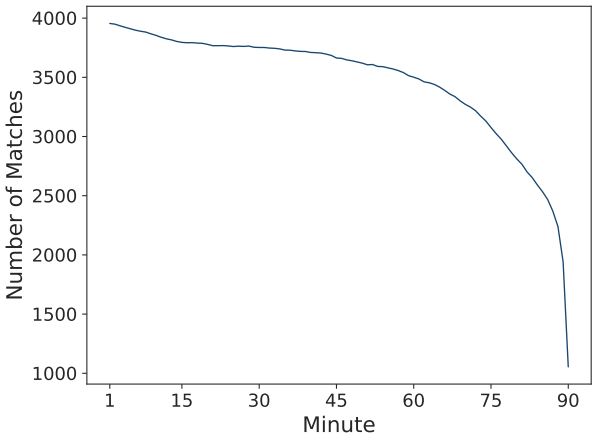
Figure A.1 displays the number of matches available to train the GRU networks for each minute.

The number of matches available decreases over time, because the betting market is often closed toward the end of a match, especially if the winner is clear before the very end but not if scorelines are close. As matches with close scorelines at the end tend to result in a draw, we confront a self-selection problem regarding missing in-play odds. Therefore, we randomly over-sample underrepresented scorelines (e.g., 4-0, 4-1, 0-4, or 1-4) by drawing samples with replacement from these minority classes until their relative frequencies are equal to the relative frequencies of the occurrence of the cases in the training sample in each minute. Analogously, we randomly over-sample the underrepresented red cards, to ensure the algorithm does not ignore red cards as feature.

To evaluate the accuracy of the predictions, we randomly split the data set into 5 folds and apply cross-validation. Figure A.2 displays various performance metrics, along with their 95% confidence intervals.

The solid line depicts the metric, averaged across the five folds; the shaded area illustrates the

Figure A.1
Gated Recurrent Unit Network Available Observations



This figure shows the available number of observations (matches) in the training sample for the GRU network. We clearly observe a decrease in the number of matches the higher the match minute.

associated confidence bounds. The upper left and right panel belong to loss and MSE, respectively. The lower left panel depicts the mean absolute error (MAE), and the lower right panel indicates the root mean squared error (RMSE). All metrics exhibit a general increasing trend over time. We offer three reasons for this pattern. First, the number of matches in the training set decreases over time. Second, prediction errors increase over time, because closing odds offer lower predictive power the more time goes by. Third, the number of key events (goals and red cards) increases toward the end of the match, which extends the directions a match can develop. All three aspects make the prediction more difficult.

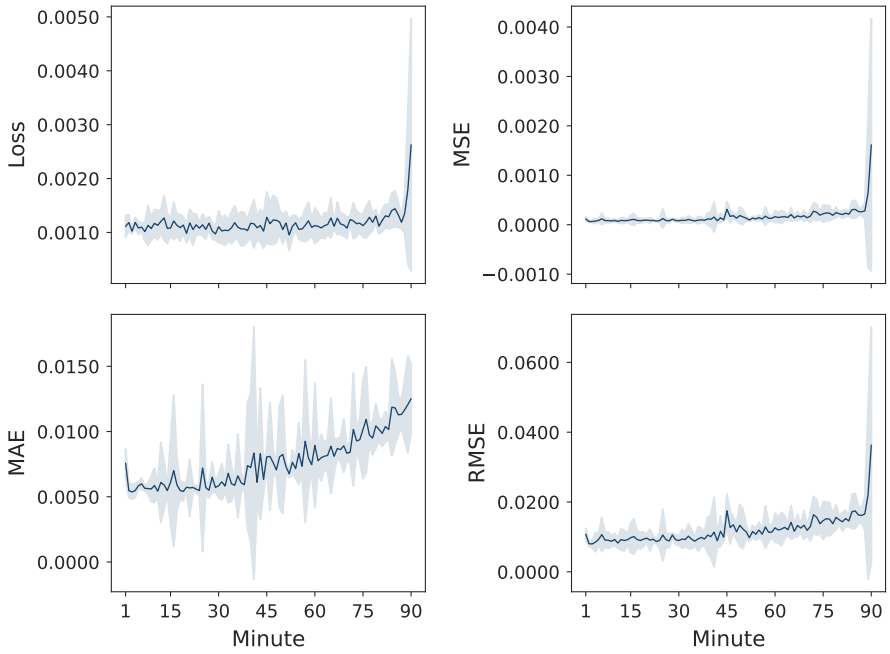
Because the targets are probabilities, we can focus on the MAE. However, similar deduction applies to the other performance metrics. The MAE can be interpreted as the mean absolute deviation between actual and predicted targets in percentage points. Consequently, predicted probabilities differ between 0.54 (minute 3) and 1.25 (minute 90) percentage points from the actual probabilities, with standard deviations of 0.03 (in minute 3) and 0.14 (in minute 90) percentage points. The MAE averaged across all minutes is 0.0077 (0.77 percentage points deviant from actual probabilities) with an average standard deviation of 0.0009 (0.09 percentage points).

A.1.2.2 Seasons 2011/12 and 2012/13

We find no in-play bookmaker odds available for seasons 2011/12 and 2012/13, from which probabilities could be derived. Therefore, we train a feedforward neural network (FFNN) for each minute to impute the missing probabilities. The training data are the same as for the GRU networks (Appendix A.1.2.1). Our predictors are normalized pre-match closing odds on the result (see Appendix A.1.1 on normalization) and in-play key events goals and red cards.

The FFNNs consist of an input layer, 4 hidden layers, and an output layer. The number of units in each layer declines from 128 to 32. The output layer contains 3 units, reflecting the number of

Figure A.2
Gated Recurrent Unit Network Performance Metrics



This figure shows GRU network performance metrics loss (top left), mean squared error (top right), mean absolute error (bottom left), and root mean squared error (bottom right) with their associated 95% confidence intervals (CI), based on 5-fold cross validation with the training set.

desired outcomes. We use a bias vector in each layer. The activations are alternately ReLU and tanh, as well as a softmax in the output layer to get results between 0 and 1, which then sum to 1. Generally, we initialize the weights using a Xavier uniform initialization (Glorot and Bengio 2010). For the hidden ReLU layers, however, we use He uniform to initialize, as proposed by He et al. (2015). To prevent overfitting, we impose an L2 penalty of size 0.001 on the weights of the output layer. As for the GRU networks, the number of epochs depends on the current match minute; the value is evenly spaced over the interval $[250, 500]$, rounded to the nearest integer, so that the network for minute 1 uses 250 epochs, and the network for minute 90 uses 500 epochs. Furthermore, we use full batch learning here.

We standardize the features to have zero mean and unit variance, analogous to Appendix A.1.2.1. Again, we apply this transformation for each minute, separately. The targets are a vector of normalized in-play probabilities for the potential outcomes *H*, *D*, and *A*, such that we face a regression problem. To minimize the MSE, we again use the AMSGrad variant of Adam Estimation. Moreover, we oversample underrepresented scores and red cards, as described for the GRU networks in Appendix A.1.2.1.

Table A.2 displays the complete architecture. Figure A.3 reveals the number of matches available to train the FFNN for each minute. Figure A.4 depicts several performance metrics and their associated 95% confidence intervals, which we obtain from 5-fold cross-validation.

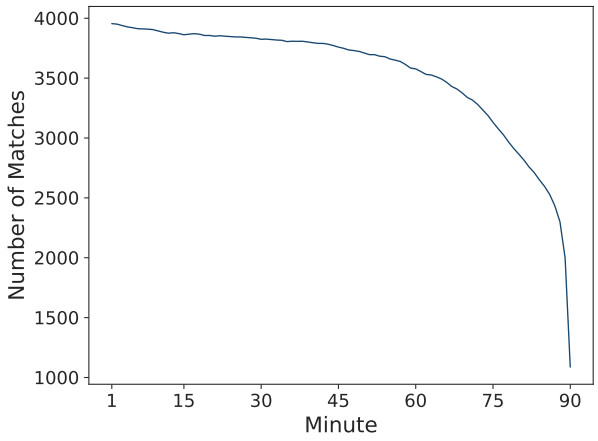
According to the MAE, the predicted probabilities on average differ between 0.73 (minute 1)

Table A.2
Feedforward Neural Network Model Configuration

Layer	Input	Units	Bias	Kernel Init.	Bias Init.	Activation	L2 Kernel
input_layer	(None, 7)	NaN	NaN	NaN	NaN	NaN	NaN
dense_1	NaN	128	1	GlorotUniform	Zeros	relu	NaN
dense_2	NaN	64	1	GlorotUniform	Zeros	tanh	NaN
dense_3	NaN	32	1	HeUniform	Zeros	relu	NaN
dense_4	NaN	16	1	GlorotUniform	Zeros	tanh	NaN
output	NaN	3	1	GlorotUniform	Zeros	softmax	0.0010

Equivalent description as for Table A.1, but using a feedforward neural network and transformed pre-match odds instead of in-play odds as features to predict transformed in-play odds for seasons 2011/12 and 2012/13.

Figure A.3
Feedforward Neural Network Available Observations



Equivalent description as for Figure A.3, but using a feedforward neural network.

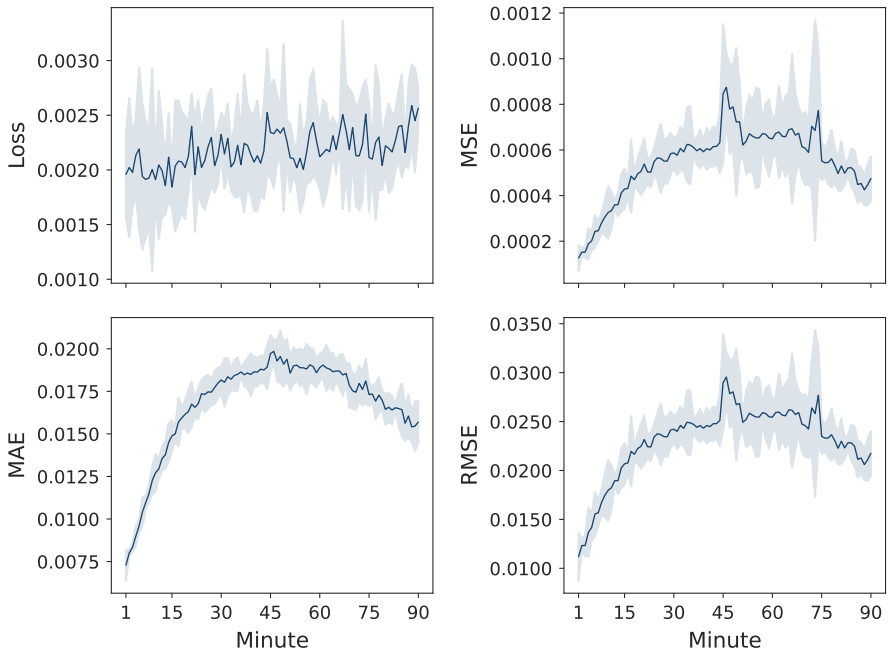
and 1.99 (minute 46) percentage points from the actual probabilities, with standard deviations of 0.05 (minute 1) and 0.03 (minute 46) percentage points. The MAE averaged across all minutes is 1.67 percentage points deviance from actual probabilities, with an average standard deviation of 0.05 percentage points.

A.1.3 Covariates

In addition to the cues, we add variables for air pressure, precipitation (dummy), relative humidity, temperature, and wind speed at level 1. These covariates aim to capture weather conditions around the stadium, which we consider relevant for drinking behavior and match outcome. For example, Ventura-Cots et al. (2019) find a positive association between alcohol consumption and colder weather, as well as fewer sunlight hours. Rao and Mohan (2021) provide evidence that injuries occur mostly when the weather conditions are hot and humid, cold, and/or wet and rainy. Adverse weather conditions also create opportunities to consume, because people tend to retreat further into the catacombs, where the cashpoints are.

Weather data are measured in 10 minute steps, instead of a minute-by-minute basis. There-

Figure A.4
Feedforward Neural Network Performance Metrics



Equivalent description as for Figure A.2, but using a feedforward neural network.

fore, we linearly interpolate air pressure, temperature, relative humidity, and wind speed. The precipitation dummy uses information on precipitation duration. Whenever the 10-minute precipitation duration is 0 or 10, the precipitation dummy is 0 or 1, respectively. If precipitation duration lies between 0 and 10 minutes (e.g., 8 minutes), we split the dummy into two blocks, comprised of one block with zero values and another block with ones, such that the order of the blocks is assigned at random. For example, for 8 minutes of rain, for the first minutes of a 10-minute interval, the precipitation dummy equals 1, and then for the last 2 minutes the dummy takes a value of 0.

In winter months, spectators can buy hot wine punch in addition to the other drinks, so we include hot wine punch sales to control for substitution effects. Furthermore, a dummy variable separating the first and second half of each match is necessary, because the two halves are structurally different, as Figure 2.1 depicts.[19]

In addition to the substitutions discussed in Section 2.3.2.1, water sales at level 1 control for baseline consumption, such that any effect on beer sales is more clearly attributable to the conscious decision to buy alcohol. That is, the effects on alcohol sales become cleaner when we also measure the sales of fans who are "just thirsty". Model precision increases too, such that baseline consumption reflects any global trend of variation on beverage sales.

[19] Analyzing both halves completely separate from each other would be interesting. Due to the restricted length of the time series though, this approach is not feasible.

At level 2, we include several probability estimates based on average closing over/under odds (normalized over/under odds). By adding the variables adapted from over/under odds, we seek to cover different spectator sentiments across matches. Pre-match sales and sales at the beginning of the match (due to delay effects) likely exhibit heterogeneity, reflecting the different levels of excitement in expectation of high versus low scoring matches.

We calculate pre-match outcome probability estimates using average closing odds for away team wins and draws (we drop home team wins due to multicollinearity). A match with a clear favorite is fundamentally different from a match with an uncertain outcome. These variables aim to account for the fan's basic state of mind. For example, fans who usually buy one beer might skip it when they anticipate a very close match, if their fear of missing important events exceeds the loss of a beer not consumed, which results in a systematic difference in mean sales between matches.

Beer prices also could be an important driver for demand and are therefore included at match level, to refine the (sports) seasonal dummies. Beer prices are stable within each season, but there is some variation between seasons over time.

The geodesic distance, defined as the shortest path between two points on a surface (the earth), between the away team's stadium and the home stadium of VfB Stuttgart (latitude-longitude data) can affect match relevance, due to its strong correlation with local rivalries (derbies). Neal and Fromme (2007) and Barry et al. (2014) find that the importance of a match (e.g., due to a rivalry) increases alcohol consumption by college football fans. We control for seasonal patterns beyond weather by using monthly dummies. For example, match relevance likely varies between the start of the season in autumn and its end in spring, which is not perfectly reflected by weather (e.g., snow in April). Pre-match team rankings and the round further contribute to match relevance. In general, matches between similarly ranked teams toward the end of the season are more important, because they can be decisive for how the teams will be ranked.

A dummy indicates whether the match is a first or second division match. During the observation period, VfB Stuttgart played in both divisions. The second division holds less prestigious matches, which might lead to systematic heterogeneity. For example, important first division matches might prompt additional consumption, due to perceptions that the event is special.

Two further dummies indicate whether the day after the match is a public holiday and if the match takes place during school holidays. Public holidays likely lead to disparities in drinking behavior. People might alter their drinking habits when the following day is a public holiday, during which excursions and festivities are popular pastimes. Zonda et al. (2009) identify a significant increase in alcohol consumption during holidays. Furthermore, people do not have to work on holidays, or on Sundays, so the public holiday dummy only refers to weekdays in this regard. School holidays likely change the composition of spectators, because family vacations typically take place during school holidays, which might alter per capita alcohol use.

Another potential factor influencing in-play consumption is the pre-match alcohol level, so

we add a dummy denoting when the Cannstatter Volksfest and the Stuttgart Spring Festival take place. These festivals, hosted by the Cannstatter Wasen, a 35 hectare area near the stadium, are of great regional importance. With more than four million visitors (pre COVID-19), the Cannstatter Volksfest is considered the second largest beer festival in the world (after the Oktoberfest in Munich). We also control for systematic substitution effects when people forgo their beer in the stadium to drink later at the festival.

Drinking behavior also depends on the time of day (Room et al. 2013), such that across cultures, it is more common after 5:00 p.m. Therefore, we include the kickoff time. Another potential source of significant differences might be the interval between meals.

With the recognition that supply potentially drives demand, we control for the number of cashpoints where people can buy drinks and food. Wait times likely are longer if fewer cashpoints are open, ceteris paribus, which might decrease transactions due to higher opportunity costs. Variance in the number of cashpoints mainly occurs when single cashpoints open or close only for certain matches, rather than due to newly constructed cashpoints, for which the seasonal dummies would be sufficient, assuming the construction occurs between seasons.

In any case, the number of spectators relates to alcohol sales (assuming enough variation). We add the information at level 2, because VfB Stuttgart does not collect this data at level 1.

Next, we add (sports) seasonal dummies to deal with several sources of heterogeneity over years. Rule changes across seasons, such as shifting definitions of accidental handballs, can lead to more or less controversial referee decisions. According to Meij et al. (2015), fans exhibit more aggressive behavior when they perceive the match as unfair. Another source of heterogeneity over seasons involves payment options, such that during the 2011/12 season, fans could only pay with a so-called fan card. In 2012/13, cash was added as an option. From 2013/14 onward, people could pay cash or with a credit card.

We use weekday dummies to account for the structural difference between a match on Saturday versus Friday, for example. Playing simultaneously with a lot of other teams on Saturday probably evokes a different excitement level, because the consequences (e.g., rankings) of simultaneous scores can be incorporated immediately into consumption decisions. Moreover, people tend to drink more on weekends (Room et al. 2013).

Tables A.3 and A.4 present the entire feature space of the main specification (including further variables) with descriptions, sources, and additional remarks.

A.1.4 Data Engineering

We standardize all explanatory variables (including the lagged dependent variable and all dummy variables). Standardization is highly recommended when using Markov chain Monte Carlo methods. Especially in a regularization context, it is important that large-scale features do not overwhelm features of smaller scales. For our data, varying scales are very prominent (e.g., air pressure in hPa versus emotional cues). The dependent variable is demeaned (following the basic idea of the decomposition described in Section 2.3.1) and transformed by the inverse hyperbolic

sine (IHS) to allow a ceteris paribus comparison between the different specifications[20]. We do not log-transform, because for the alternative specifications Die-Hard Fans, Shandy, Soft Drinks, and Seasons in Section 2.4.2 we observe match minutes with no sales even after taking the lags. Yet we obtain marginal effects in %. Finally, we drop the first match minute dummy, due to perfect multicollinearity, resulting in a total of $T-1$ minute dummies for which coefficient ψ_{11} serves as reference group.

A.1.5 Algorithm

We estimate the model using the probabilistic programming package PyMC (Salvatier, Wiecki, and Fonnesbeck 2016), which relies on PyTensor in its computational backend. In particular, we apply the No-U-Turn (NUTS) sampler (Hoffman and Gelman 2014), a recursive algorithm for continuous variables based on Hamiltonian mechanics that extends Hamiltonian Monte Carlo (HMC) by eliminating the need to set a number of steps through the inbuilt automatic stoppage once the sampler starts to make a U-turn. We set a comparably large acceptance probability of 0.9, which is associated with a small step size, so that we can achieve non-diverged trajectories for the samples. We run 3 chains with 2500 iterations each and an additional 1000 burn-in samples per chain that we discard. The Gelman-Rubin statistic $\hat{R}$ (Gelman and Rubin 1992; Brooks and Gelman 1998) monitors convergence.

[20] The slope of $\sinh^{-1}(y) = \ln\left(y^2 + \sqrt{y^2+1}\right)$ is approximately equal to that of $\ln(y)$ for $y \geq 2$, so we interpret the coefficients as for a log-transformed dependent variable, because the scale of our dependent variable is large enough.

A.2 Data Overview

Table A.3
Variable Summary

Name	Description	Comment	Source
Shock	Shock	Generated from in-play odds and closing odds	7
Surprise	Surprise	Generated from in-play odds and closing odds	7
Suspense	Suspense	Generated from in-play odds, closing odds, historical scoring rates, timing of goals, and red cards	5, 6, 7, and 8
BeerSls	Number of beverage units sold	Dependent variable; considered are the sales of a certain drink (beer, shandy or soft drinks depending on the specification) within the target stand(s) (e.g., whole stadium except for the away fan area)	11
AirPress	Air pressure (in hPa)	Linearly interpolated, because the measuring station (5km away from the stadium) records in 10-minute steps	2
HotWineSls	Number of hot wine punch units sold	Considered are only the sales within the target stand	11
Is1stHalf	Dummy for first half	Generated from match data	6
IsPrecip	Dummy for precipitation	Derived from precipitation duration on a 10-minute basis with the measuring station (5km away from the stadium) recording in 10 minute steps	2
IsSubAway	Dummy for an away substitution		8
IsSubHome	Dummy for a home substitution		8
RelHumid	Relative humidity (in %)	Linearly interpolated, because the measuring station (5km away from the stadium) records in 10-minute steps	2
Temp	Temperature (in Celsius)	Linearly interpolated, because the measuring station (5km away from the stadium) records in 10-minute steps	2
WaterSls	Number of water units sold		11
Wind	Wind speed (in m/s)	Linearly interpolated, because the measuring station (5km away from the stadium) records in 10-minute steps	2
AvgClsOver05Prob	Pre-match probability that joint score exceeds 0.5	Uses average closing odds (multiple bookmakers)	7
AvgClsOver15Prob	Pre-match probability that joint score exceeds 1.5	Uses average closing odds (multiple bookmakers)	7
AvgClsOver25Prob	Pre-match probability that joint score exceeds 2.5	Uses average closing odds (multiple bookmakers)	7

AvgClsOver35Prob	Pre-match probability that joint score exceeds 3.5	Uses average closing odds (multiple bookmakers)	7
AvgClsOver45Prob	Pre-match probability that joint score exceeds 4.5	Uses average closing odds (multiple bookmakers)	7
AvgClsOver55Prob	Pre-match probability that joint score exceeds 5.5	Uses average closing odds (multiple bookmakers)	7
AvgClsProbAway	Pre-match probability that away team wins	Uses average closing odds (multiple bookmakers)	7
AvgClsProbHome	Pre-match probability that home team wins	Uses average closing odds (multiple bookmakers)	7
BeerPrice	Price of beverage (in Euro)	The corresponding beverage is beer, shandy or soft drink depending on which one is set as dependent variable in the underlying specification	11
GeoDist	Geodesic distance between home and away team's stadium (in km)	Generated using latitude and longitude of each team's stadium	3
HotWinePrice	Price of hot wine punch (in Euro)		11
Is1stDiv	Dummy for a first division match		6
IsMobCpoint	Dummy for the availability of mobile cashpoints	Generated from the number of units sold by mobile cashpoints; considered are the mobile cashpoints within the target stand(s) (e.g., whole stadium except for the away fan area) using the assumption that no sales imply no availability	11
IsPromoAway	Dummy for the away team promoted last season		10
IsPromoHome	Dummy for the home team promoted last season		10
IsPubHoliday	Dummy for a legal or church holiday at the day after the match		4
IsRelegAway	Dummy for the away team relegated last season		10
IsRelegHome	Dummy for the home team relegated last season		10
IsRunner	Dummy for the availability of mobile salesmen	Generated from the number of units sold by mobile salesmen (beer runners) who sell beer and softdrinks only (no shandy, for example); considered are the sales in the whole stadium (due to the data at hand) using the assumptions that no sales imply no availability and non-zero sales indicate availability in every stand	11
IsSchlHoliday	Dummy for school holidays at the match day		9
IsSoldOut	Dummy for a sold-out match		5
IsVar	Dummy for the presence of a video assistant referee		6
IsWasen	Dummy for Wasen - a local festival		1
Kickoff15:30	Dummy for kickoff time 15:30		6
Kickoff15:45	Dummy for kickoff time 15:45		6
Kickoff17:30	Dummy for kickoff time 17:30		6

Kickoff18:00	Dummy for kickoff time 18:00		6
Kickoff18:30	Dummy for kickoff time 18:30		6
Kickoff20:00	Dummy for kickoff time 20:00		6
Kickoff20:15	Dummy for kickoff time 20:15		6
Kickoff20:30	Dummy for kickoff time 20:30		6
MonthAug	Dummy for the match taking place in August		6
MonthDec	Dummy for the match taking place in December		6
MonthFeb	Dummy for the match taking place in February		6
MonthJan	Dummy for the match taking place in January		6
MonthMar	Dummy for the match taking place in March		6
MonthMay	Dummy for the match taking place in May		6
MonthNov	Dummy for the match taking place in November		6
MonthOct	Dummy for the match taking place in October		6
MonthSep	Dummy for the match taking place in September		6
NumCpoints	Number of active cashpoints	Considered are the cashpoints within the target stand(s) (e.g., whole stadium except for the away fan area)	11
NumSpects	Number of spectators	Refers always to all spectators in the whole stadium	5
RankAwayLast	Final rank of the away team last season		10
RankAwayPre	Rank of the away team before the match		5
RankHomeLast	Final rank of the home team last season		5
RankHomePre	Rank of the home team before the match		5
Round	Round of the season		5
Season2014/15	Dummy for season 2014/2015		6
Season2015/16	Dummy for season 2015/2016		6
Season2016/17	Dummy for season 2016/2017		6
Season2017/18	Dummy for season 2017/2018		6
Season2018/19	Dummy for season 2018/2019		6
WkdayMo	Dummy for the match taking place on a Monday		8
WkdaySa	Dummy for the match taking place on a Saturday		8
WkdaySu	Dummy for the match taking place on a Sunday		8
WkdayTu	Dummy for the match taking place on a Tuesday		8

WkdayWe	Dummy for the match taking place on a Wednesday	8

This table presents details on dependent and independent variables, used in the main specification, excluding the match minute dummies. Note that, depending on the specification, isolated variables may be discarded due to multicollinearity. Also, the set of variables might differ slightly in specifications where the sample changes, such as in the Seasons specification (e.g., additional kickoff time or season dummies). We code all binary indicators 0: false and 1: true. The reference groups for kickoff time, month, season, and weekday are 13:30, Apr, 2013/14, and Fr, respectively. Table A.4 particularizes the numbered data sources in the last column.

Table A.4

Data Sources

Source	Name	URL	Retrieved On
1	cannstatter-volksfest.de	https://www.cannstatter-volksfest.de/de/landing-page/	October 02, 2019
2	Deutscher Wetterdienst	https://www.dwd.de/	November 9, 2021 at 11:50:34 AM
3	Google Maps	https://www.google.de/maps/	September 26, 2019
4	kalender-online.com	https://kalender-online.com/	November 12, 2021 at 8:43:36 PM
5	kicker	https://www.kicker.de/	October 21, 2020 at 8:08:46 PM
6	NowGoal	https://www.nowgoal.com/	September 8, 2020 at 10:55:08 to October 21, 2020 12:03:04 PM
7	OddsPortal	https://www.oddsportal.com/	October 22, 2020 4:34:22 PM
8	OptaSports		August 16, 2021 12:45:28 PM
9	schulferien.org	https://www.schulferien.org/	November 13, 2021 5:14:34 PM
10	transfermarkt.de	https://www.transfermarkt.de/	September 26, 2019
11	VfB Stuttgart		November 7, 2018 and July 12, 2019

This table presents information on the data sources specified in Table A.3.

Table A.5

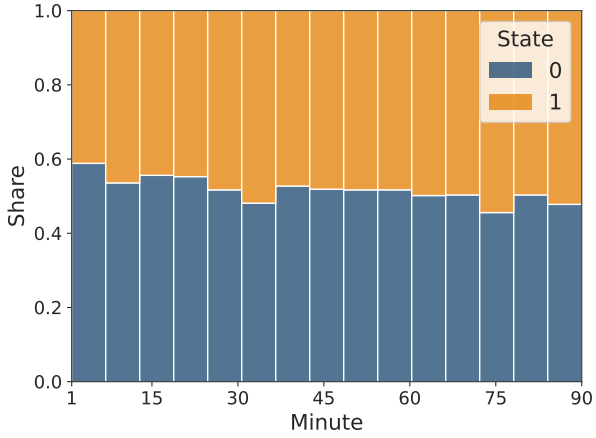
Descriptive Statistics: States

	count	sum	mean	std	min	25%	50%	75%	max
State 0	98	4,559	46.5	17.5	16	33.0	44.5	60.8	82
State 1	98	4,261	43.5	17.5	8	29.2	45.5	57.0	74

This table presents basic summary statistics for the negative and the positive state 0 and 1, respectively. Overall, we observe the positive state 1 4,261 times (sum) in 98 matches (count). There is at least 1 match in which we record the negative state 0 only 16 times; 25%, 50%, and 75% denote the respective quantiles.

Figure A.5

In-Play Distribution of States



This figure shows shares of both negative state 0 and positive state 1 during matches. Overall, the states are very consistently observed over time. If at all, the positive state 1 exhibits a slight positive trend, whereas the occurences of the negative state 0 slightly decrease.

Table A.6

Descriptive Statistics: Explanatory Variables Per Minute in State 0

		count	mean	std	min	25%	50%	75%	max
Shock	$X_{1,0}$	4,559	0.3034	0.2338	0.0011	0.0755	0.2949	0.4704	0.9319
Surprise	$X_{1,1}$	4,559	0.0157	0.0633	0	0	0.0042	0.0115	0.9422
Suspense	$X_{1,2}$	4,511	0.0701	0.0410	0.0100	0.0496	0.0603	0.0818	0.2544
AirPress	$Z_{1,0}$	4,559	979	9.25	948	976	981	986	998
HotWineSls	$Z_{1,1}$	4,559	2.48	6.90	0	0	0	1	69
Is1stHalf	$Z_{1,2}$	4,559	0.5183	0.4997	0	0	1	1	1
IsPrecip	$Z_{1,3}$	4,559	0.1000	0.3001	0	0	0	0	1
IsSubAway	$Z_{1,4}$	4,559	0.0235	0.1599	0	0	0	0	2
IsSubHome	$Z_{1,5}$	4,559	0.0327	0.1874	0	0	0	0	2
RelHumid	$Z_{1,6}$	4,559	68.3	17.3	24.7	57.2	67.1	82.8	99.2
Temp	$Z_{1,7}$	4,559	10.8	7.17	-3.04	5.01	10.3	15.3	31
WaterSls	$Z_{1,8}$	4,559	2.78	4.80	0	0	1	3	61
Wind	$Z_{1,9}$	4,559	3.17	1.47	0.2000	2.20	2.96	4.02	8.47
AvgClsOver05Prob	$Z_{2,0}$	4,559	0.9181	0.0145	0.8787	0.9085	0.9193	0.9278	0.9513
AvgClsOver15Prob	$Z_{2,1}$	4,559	0.7691	0.0364	0.6706	0.7465	0.7730	0.7958	0.8450
AvgClsOver25Prob	$Z_{2,2}$	4,559	0.5481	0.0531	0.4165	0.5144	0.5471	0.5835	0.6682

AvgClsOver35Prob	$Z_{2,3}$	4,559	0.3402	0.0512	0.2285	0.3068	0.3397	0.3706	0.4672
AvgClsOver45Prob	$Z_{2,4}$	4,559	0.1879	0.0373	0.1145	0.1605	0.1888	0.2099	0.2851
AvgClsOver55Prob	$Z_{2,5}$	4,559	0.0993	0.0230	0.0541	0.0843	0.0983	0.1112	0.1605
AvgClsProbAway	$Z_{2,6}$	4,559	0.3142	0.1345	0.1376	0.2237	0.2768	0.3668	0.7721
AvgClsProbHome	$Z_{2,7}$	4,559	0.4264	0.1281	0.0772	0.3466	0.4495	0.5156	0.6481
BeerPrice	$Z_{2,8}$	4,559	4.18	0.0907	4	4.20	4.20	4.20	4.30
GeoDist	$Z_{2,9}$	4,559	294	140	55.6	152	326	398	536
HotWinePrice	$Z_{2,10}$	4,559	3.93	0.2841	3.50	3.50	4	4.20	4.30
Is1stDiv	$Z_{2,11}$	4,559	0.8596	0.3474	0	1	1	1	1
IsMobCpoint	$Z_{2,12}$	4,559	0.4431	0.4968	0	0	0	1	1
IsPubHoliday	$Z_{2,13}$	4,559	0.0180	0.1329	0	0	0	0	1
IsRelegAway	$Z_{2,14}$	4,559	0.0121	0.1092	0	0	0	0	1
IsRunner	$Z_{2,15}$	4,559	0.6080	0.4882	0	0	1	1	1
IsSchlHoliday	$Z_{2,16}$	4,559	0.1450	0.3521	0	0	0	0	1
IsSoldOut	$Z_{2,17}$	4,559	0.2288	0.4201	0	0	0	0	1
IsWasen	$Z_{2,18}$	4,559	0.1535	0.3605	0	0	0	0	1
Kickoff15:30	$Z_{2,19}$	4,559	0.6319	0.4823	0	0	1	1	1
Kickoff15:45	$Z_{2,20}$	4,559	0.0118	0.1082	0	0	0	0	1
Kickoff17:30	$Z_{2,21}$	4,559	0.0660	0.2483	0	0	0	0	1
Kickoff18:00	$Z_{2,22}$	4,559	0.0254	0.1575	0	0	0	0	1
Kickoff18:30	$Z_{2,23}$	4,559	0.0770	0.2666	0	0	0	0	1
Kickoff20:30	$Z_{2,24}$	4,559	0.0980	0.2974	0	0	0	0	1
MonthAug	$Z_{2,25}$	4,559	0.0726	0.2595	0	0	0	0	1
MonthDec	$Z_{2,26}$	4,559	0.1132	0.3169	0	0	0	0	1
MonthFeb	$Z_{2,27}$	4,559	0.1222	0.3275	0	0	0	0	1
MonthJan	$Z_{2,28}$	4,559	0.0818	0.2741	0	0	0	0	1
MonthMar	$Z_{2,29}$	4,559	0.1141	0.3179	0	0	0	0	1
MonthMay	$Z_{2,30}$	4,559	0.0660	0.2483	0	0	0	0	1
MonthNov	$Z_{2,31}$	4,559	0.0805	0.2721	0	0	0	0	1
MonthOct	$Z_{2,32}$	4,559	0.0919	0.2889	0	0	0	0	1
MonthSep	$Z_{2,33}$	4,559	0.1507	0.3578	0	0	0	0	1
NumCpoints	$Z_{2,34}$	4,559	157	20.5	113	135	157	178	188
NumSpects	$Z_{2,35}$	4,559	52,214	6,180	36,800	47,125	54,068	58,000	60,000
RankAwayLast	$Z_{2,36}$	4,559	7.63	4.85	1	3	7	12	18
RankAwayPre	$Z_{2,37}$	4,559	9.51	5.04	1	5	9	14	18
RankHomePre	$Z_{2,38}$	4,559	13	4.90	1	11	15	16.5	18
Round	$Z_{2,39}$	4,559	16.8	9.66	1	8	17	24	34
Season2014/15	$Z_{2,40}$	4,559	0.1807	0.3848	0	0	0	0	1
Season2015/16	$Z_{2,41}$	4,559	0.1895	0.3920	0	0	0	0	1
Season2017/18	$Z_{2,42}$	4,559	0.1542	0.3612	0	0	0	0	1
Season2018/19	$Z_{2,43}$	4,559	0.1660	0.3722	0	0	0	0	1
WkdayMo	$Z_{2,44}$	4,559	0.0511	0.2202	0	0	0	0	1
WkdaySa	$Z_{2,45}$	4,559	0.5863	0.4925	0	0	1	1	1
WkdaySu	$Z_{2,46}$	4,559	0.2165	0.4119	0	0	0	0	1
WkdayTu	$Z_{2,47}$	4,559	0.0066	0.0809	0	0	0	0	1
WkdayWe	$Z_{2,48}$	4,559	0.0035	0.0591	0	0	0	0	1

This table presents basic summary statistics of explanatory variables (Table A.3), excluding the lagged dependent variable and match minute dummies, on a per minute basis in negative state 0. For example, the minimum value of *Surprise* ($X_{1,1}$) per minute in state 0 is 0. Note that 25%, 50%, and 75% denote the respective quantiles.

Table A.7

Descriptive Statistics: Explanatory Variables Per Minute in State 1

		count	mean	std	min	25%	50%	75%	max
Shock	$X_{1,0}$	4,261	0.3035	0.2056	0.0023	0.1207	0.3054	0.4361	0.9617
Surprise	$X_{1,1}$	4,261	0.0162	0.0660	0	0	0.0034	0.0109	1.04
Suspense	$X_{1,2}$	4,211	0.0669	0.0377	0.0083	0.0470	0.0599	0.0807	0.2439
AirPress	$Z_{1,0}$	4,261	980	9.16	948	975	981	986	998
HotWineSls	$Z_{1,1}$	4,261	1.87	5.53	0	0	0	1	62
Is1stHalf	$Z_{1,2}$	4,261	0.4804	0.4997	0	0	0	1	1
IsPrecip	$Z_{1,3}$	4,261	0.1251	0.3309	0	0	0	0	1
IsSubAway	$Z_{1,4}$	4,261	0.0390	0.2098	0	0	0	0	2
IsSubHome	$Z_{1,5}$	4,261	0.0275	0.1677	0	0	0	0	2
RelHumid	$Z_{1,6}$	4,261	68	19.1	24.9	54.4	69.3	84.8	100
Temp	$Z_{1,7}$	4,261	10.6	7.18	-3.16	4.68	10.1	15.2	30.9
WaterSls	$Z_{1,8}$	4,261	3.00	5.68	0	0	1	3	68
Wind	$Z_{1,9}$	4,261	3.09	1.43	0	2.10	3	3.99	8.50
AvgClsOver05Prob	$Z_{2,0}$	4,261	0.9201	0.0144	0.8787	0.9100	0.9201	0.9299	0.9513
AvgClsOver15Prob	$Z_{2,1}$	4,261	0.7739	0.0361	0.6706	0.7510	0.7790	0.7983	0.8450
AvgClsOver25Prob	$Z_{2,2}$	4,261	0.5554	0.0531	0.4165	0.5209	0.5547	0.5882	0.6682
AvgClsOver35Prob	$Z_{2,3}$	4,261	0.3469	0.0511	0.2285	0.3121	0.3430	0.3716	0.4672
AvgClsOver45Prob	$Z_{2,4}$	4,261	0.1928	0.0378	0.1145	0.1657	0.1905	0.2106	0.2851
AvgClsOver55Prob	$Z_{2,5}$	4,261	0.1023	0.0232	0.0541	0.0872	0.0985	0.1143	0.1605
AvgClsProbAway	$Z_{2,6}$	4,261	0.3361	0.1543	0.1376	0.2237	0.2788	0.4031	0.7721
AvgClsProbHome	$Z_{2,7}$	4,261	0.4087	0.1409	0.0772	0.3238	0.4459	0.5221	0.6481
BeerPrice	$Z_{2,8}$	4,261	4.18	0.0915	4	4.20	4.20	4.20	4.30
GeoDist	$Z_{2,9}$	4,261	289	143	55.6	152	326	398	536
HotWinePrice	$Z_{2,10}$	4,261	3.96	0.2732	3.50	3.99	4	4.20	4.30
Is1stDiv	$Z_{2,11}$	4,261	0.8334	0.3727	0	1	1	1	1
IsMobCpoint	$Z_{2,12}$	4,261	0.4130	0.4924	0	0	0	1	1
IsPubHoliday	$Z_{2,13}$	4,261	0.0019	0.0433	0	0	0	0	1
IsRelegAway	$Z_{2,14}$	4,261	0.0082	0.0903	0	0	0	0	1
IsRunner	$Z_{2,15}$	4,261	0.6801	0.4665	0	0	1	1	1
IsSchlHoliday	$Z_{2,16}$	4,261	0.1195	0.3244	0	0	0	0	1
IsSoldOut	$Z_{2,17}$	4,261	0.2621	0.4399	0	0	0	1	1
IsWasen	$Z_{2,18}$	4,261	0.1948	0.3961	0	0	0	0	1
Kickoff15:30	$Z_{2,19}$	4,261	0.6123	0.4873	0	0	1	1	1
Kickoff15:45	$Z_{2,20}$	4,261	0.0084	0.0915	0	0	0	0	1
Kickoff17:30	$Z_{2,21}$	4,261	0.0772	0.2670	0	0	0	0	1
Kickoff18:00	$Z_{2,22}$	4,261	0.0150	0.1216	0	0	0	0	1
Kickoff18:30	$Z_{2,23}$	4,261	0.0866	0.2813	0	0	0	0	1
Kickoff20:30	$Z_{2,24}$	4,261	0.0641	0.2449	0	0	0	0	1
MonthAug	$Z_{2,25}$	4,261	0.0490	0.2160	0	0	0	0	1
MonthDec	$Z_{2,26}$	4,261	0.1112	0.3145	0	0	0	0	1
MonthFeb	$Z_{2,27}$	4,261	0.1439	0.3510	0	0	0	0	1
MonthJan	$Z_{2,28}$	4,261	0.0603	0.2381	0	0	0	0	1
MonthMar	$Z_{2,29}$	4,261	0.1103	0.3133	0	0	0	0	1
MonthMay	$Z_{2,30}$	4,261	0.0983	0.2978	0	0	0	0	1
MonthNov	$Z_{2,31}$	4,261	0.0828	0.2757	0	0	0	0	1

MonthOct	$Z_{2,32}$	4,261	0.0918	0.2887	0	0	0	0	1
MonthSep	$Z_{2,33}$	4,261	0.1134	0.3171	0	0	0	0	1
NumCpoints	$Z_{2,34}$	4,261	158	20.2	113	140	155	178	188
NumSpects	$Z_{2,35}$	4,261	52,332	6,725	36,800	46,600	54,300	58,569	60,000
RankAwayLast	$Z_{2,36}$	4,261	7.85	4.88	1	3	8	12	18
RankAwayPre	$Z_{2,37}$	4,261	8.98	5.07	1	4	9	13	18
RankHomePre	$Z_{2,38}$	4,261	12.7	5.17	1	11	14	17	18
Round	$Z_{2,39}$	4,261	18.5	9.67	1	10	19	28	34
Season2014/15	$Z_{2,40}$	4,261	0.1446	0.3517	0	0	0	0	1
Season2015/16	$Z_{2,41}$	4,261	0.1563	0.3632	0	0	0	0	1
Season2017/18	$Z_{2,42}$	4,261	0.1941	0.3955	0	0	0	0	1
Season2018/19	$Z_{2,43}$	4,261	0.1603	0.3669	0	0	0	0	1
WkdayMo	$Z_{2,44}$	4,261	0.0720	0.2586	0	0	0	0	1
WkdaySa	$Z_{2,45}$	4,261	0.5766	0.4942	0	0	1	1	1
WkdaySu	$Z_{2,46}$	4,261	0.2330	0.4228	0	0	0	0	1
WkdayTu	$Z_{2,47}$	4,261	0.0141	0.1178	0	0	0	0	1
WkdayWe	$Z_{2,48}$	4,261	0.0174	0.1306	0	0	0	0	1

Equivalent description as for Table A.6, but given positive state 1.

A.3 Further Results

A.3.1 Main Specification

Table A.8
Main Model Results

			mean	std	median	$\text{hdi}^{5\%}$	$\text{hdi}^{95\%}$	$\hat{R}$
Surprise	State 0	$\hat{\beta}_{1,1,0,0}$	-0.0170	0.0040	-0.0170	-0.0240	-0.0100	1
		$\hat{\beta}_{1,1,0,1}$	0.0040	0.0040	0.0040	-0.0020	0.0110	1
		$\hat{\beta}_{1,1,0,2}$	0.0080	0.0050	0.0080	-0.0010	0.0140	1
		$\hat{\beta}_{1,1,0,3}$	0.0160	0.0050	0.0160	0.0090	0.0230	1
		$\hat{\beta}_{1,1,0,4}$	0.0160	0.0040	0.0160	0.0090	0.0230	1
		$\hat{\beta}_{1,1,0,5}$	0.0120	0.0050	0.0120	0.0040	0.0190	1
		$\hat{\beta}_{1,1,0,6}$	0.0020	0.0030	0.0010	-0.0030	0.0080	1
		$\hat{\beta}_{1,1,0,7}$	-0.0010	0.0030	0.0000	-0.0060	0.0050	1
		$\hat{\beta}_{1,1,0,8}$	-0.0020	0.0040	-0.0010	-0.0080	0.0030	1
		$\hat{\beta}_{1,1,0,9}$	-0.0100	0.0050	-0.0100	-0.0160	0.0000	1
	State 1	$\hat{\beta}_{1,1,1,0}$	-0.0140	0.0040	-0.0140	-0.0210	-0.0070	1
		$\hat{\beta}_{1,1,1,1}$	-0.0240	0.0040	-0.0240	-0.0310	-0.0170	1
		$\hat{\beta}_{1,1,1,2}$	0.0180	0.0040	0.0180	0.0100	0.0250	1
		$\hat{\beta}_{1,1,1,3}$	0.0190	0.0040	0.0190	0.0110	0.0250	1
		$\hat{\beta}_{1,1,1,4}$	0.0170	0.0040	0.0170	0.0100	0.0240	1
		$\hat{\beta}_{1,1,1,5}$	0.0060	0.0040	0.0060	-0.0010	0.0130	1
		$\hat{\beta}_{1,1,1,6}$	0.0070	0.0050	0.0070	0.0000	0.0140	1
		$\hat{\beta}_{1,1,1,7}$	0.0050	0.0040	0.0050	-0.0010	0.0120	1
		$\hat{\beta}_{1,1,1,8}$	0.0010	0.0030	0.0000	-0.0040	0.0070	1
		$\hat{\beta}_{1,1,1,9}$	0.0000	0.0030	0.0000	-0.0060	0.0040	1

Suspense	State 0	$\hat{\beta}_{1,2,0,0}$	0.0040	0.0080	0.0020	-0.0080	0.0180	1
		$\hat{\beta}_{1,2,0,1}$	0.0020	0.0100	0.0000	-0.0130	0.0180	1
		$\hat{\beta}_{1,2,0,2}$	-0.0040	0.0150	0.0000	-0.0300	0.0200	1
		$\hat{\beta}_{1,2,0,3}$	-0.0720	0.0230	-0.0710	-0.1120	-0.0370	1
		$\hat{\beta}_{1,2,0,4}$	0.0210	0.0200	0.0160	-0.0050	0.0550	1
		$\hat{\beta}_{1,2,0,5}$	0.0040	0.0110	0.0010	-0.0110	0.0240	1
		$\hat{\beta}_{1,2,0,6}$	0.0160	0.0140	0.0160	-0.0040	0.0390	1
		$\hat{\beta}_{1,2,0,7}$	0.0100	0.0150	0.0060	-0.0110	0.0370	1
		$\hat{\beta}_{1,2,0,8}$	-0.0010	0.0090	0.0000	-0.0150	0.0130	1
		$\hat{\beta}_{1,2,0,9}$	0.0030	0.0080	0.0010	-0.0090	0.0170	1
	State 1	$\hat{\beta}_{1,2,1,0}$	-0.0040	0.0080	-0.0020	-0.0190	0.0080	1
		$\hat{\beta}_{1,2,1,1}$	0.0040	0.0100	0.0010	-0.0090	0.0220	1
		$\hat{\beta}_{1,2,1,2}$	-0.0200	0.0160	-0.0190	-0.0450	0.0040	1
		$\hat{\beta}_{1,2,1,3}$	-0.0600	0.0230	-0.0600	-0.1010	-0.0260	1
		$\hat{\beta}_{1,2,1,4}$	0.0160	0.0200	0.0090	-0.0090	0.0500	1
		$\hat{\beta}_{1,2,1,5}$	0.0030	0.0110	0.0000	-0.0120	0.0230	1
		$\hat{\beta}_{1,2,1,6}$	-0.0070	0.0140	-0.0040	-0.0320	0.0130	1
		$\hat{\beta}_{1,2,1,7}$	0.0340	0.0160	0.0320	0.0090	0.0620	1
		$\hat{\beta}_{1,2,1,8}$	0.0020	0.0090	0.0000	-0.0120	0.0160	1
		$\hat{\beta}_{1,2,1,9}$	0.0040	0.0080	0.0020	-0.0080	0.0190	1
AirPress	State 0	$\hat{\gamma}_{1,0,0,0}$	0.0000	0.0170	0.0000	-0.0250	0.0250	1
		$\hat{\gamma}_{1,0,0,1}$	0.0000	0.0170	0.0000	-0.0250	0.0260	1
		$\hat{\gamma}_{1,0,0,2}$	0.0030	0.0180	0.0010	-0.0230	0.0290	1
		$\hat{\gamma}_{1,0,0,3}$	-0.0030	0.0180	-0.0010	-0.0350	0.0200	1
		$\hat{\gamma}_{1,0,0,4}$	0.0030	0.0190	0.0010	-0.0250	0.0310	1
		$\hat{\gamma}_{1,0,0,5}$	-0.0050	0.0190	-0.0030	-0.0360	0.0230	1
		$\hat{\gamma}_{1,0,0,6}$	-0.0010	0.0170	0.0000	-0.0270	0.0230	1
		$\hat{\gamma}_{1,0,0,7}$	0.0010	0.0180	0.0000	-0.0230	0.0310	1
		$\hat{\gamma}_{1,0,0,8}$	-0.0010	0.0180	0.0000	-0.0280	0.0230	1
		$\hat{\gamma}_{1,0,0,9}$	0.0020	0.0170	0.0000	-0.0230	0.0280	1
	State 1	$\hat{\gamma}_{1,0,1,0}$	0.0020	0.0170	0.0000	-0.0220	0.0280	1
		$\hat{\gamma}_{1,0,1,1}$	0.0010	0.0170	0.0000	-0.0230	0.0270	1
		$\hat{\gamma}_{1,0,1,2}$	-0.0020	0.0180	-0.0010	-0.0310	0.0230	1
		$\hat{\gamma}_{1,0,1,3}$	0.0020	0.0180	0.0010	-0.0230	0.0300	1
		$\hat{\gamma}_{1,0,1,4}$	-0.0030	0.0190	-0.0010	-0.0320	0.0240	1
		$\hat{\gamma}_{1,0,1,5}$	0.0050	0.0190	0.0030	-0.0190	0.0400	1
		$\hat{\gamma}_{1,0,1,6}$	0.0010	0.0170	0.0000	-0.0220	0.0280	1
		$\hat{\gamma}_{1,0,1,7}$	-0.0010	0.0180	0.0000	-0.0260	0.0290	1
		$\hat{\gamma}_{1,0,1,8}$	0.0020	0.0180	0.0010	-0.0240	0.0280	1
		$\hat{\gamma}_{1,0,1,9}$	-0.0020	0.0170	0.0000	-0.0270	0.0240	1
HotWineSls	State 0	$\hat{\gamma}_{1,1,0,0}$	-0.0010	0.0060	0.0000	-0.0110	0.0090	1
		$\hat{\gamma}_{1,1,0,1}$	-0.0030	0.0060	-0.0010	-0.0150	0.0060	1
		$\hat{\gamma}_{1,1,0,2}$	0.0000	0.0060	0.0000	-0.0100	0.0110	1
		$\hat{\gamma}_{1,1,0,3}$	0.0010	0.0060	0.0000	-0.0090	0.0110	1
		$\hat{\gamma}_{1,1,0,4}$	-0.0030	0.0070	-0.0010	-0.0160	0.0070	1
		$\hat{\gamma}_{1,1,0,5}$	0.0050	0.0070	0.0040	-0.0040	0.0190	1
		$\hat{\gamma}_{1,1,0,6}$	0.0010	0.0060	0.0000	-0.0080	0.0130	1
		$\hat{\gamma}_{1,1,0,7}$	0.0000	0.0060	0.0000	-0.0120	0.0090	1
		$\hat{\gamma}_{1,1,0,8}$	-0.0010	0.0060	0.0000	-0.0120	0.0080	1

		$\hat{\gamma}_{1,1,0,9}$	0.0010	0.0050	0.0000	-0.0080	0.0100	1
	State 1	$\hat{\gamma}_{1,1,1,0}$	0.0140	0.0090	0.0140	-0.0010	0.0270	1
		$\hat{\gamma}_{1,1,1,1}$	0.0040	0.0070	0.0020	-0.0060	0.0180	1
		$\hat{\gamma}_{1,1,1,2}$	0.0020	0.0060	0.0000	-0.0080	0.0130	1
		$\hat{\gamma}_{1,1,1,3}$	0.0030	0.0070	0.0010	-0.0070	0.0160	1
		$\hat{\gamma}_{1,1,1,4}$	-0.0080	0.0090	-0.0070	-0.0230	0.0040	1
		$\hat{\gamma}_{1,1,1,5}$	0.0010	0.0070	0.0000	-0.0100	0.0130	1
		$\hat{\gamma}_{1,1,1,6}$	-0.0040	0.0070	-0.0010	-0.0170	0.0070	1
		$\hat{\gamma}_{1,1,1,7}$	-0.0050	0.0080	-0.0040	-0.0190	0.0050	1
		$\hat{\gamma}_{1,1,1,8}$	-0.0030	0.0070	-0.0010	-0.0140	0.0080	1
		$\hat{\gamma}_{1,1,1,9}$	0.0000	0.0060	0.0000	-0.0110	0.0090	1
Is1stHalf	State 0	$\hat{\gamma}_{1,2,0,0}$	-0.1270	0.0370	-0.1320	-0.1880	-0.0670	1
		$\hat{\gamma}_{1,2,0,1}$	-0.0040	0.0260	0.0000	-0.0370	0.0400	1
		$\hat{\gamma}_{1,2,0,2}$	0.0020	0.0150	0.0000	-0.0170	0.0280	1
		$\hat{\gamma}_{1,2,0,3}$	0.0010	0.0150	0.0000	-0.0230	0.0260	1
		$\hat{\gamma}_{1,2,0,4}$	0.0180	0.0230	0.0110	-0.0100	0.0580	1
		$\hat{\gamma}_{1,2,0,5}$	0.0160	0.0230	0.0070	-0.0110	0.0560	1
		$\hat{\gamma}_{1,2,0,6}$	0.0110	0.0200	0.0030	-0.0130	0.0480	1
		$\hat{\gamma}_{1,2,0,7}$	0.0280	0.0280	0.0210	-0.0050	0.0740	1
		$\hat{\gamma}_{1,2,0,8}$	0.0200	0.0260	0.0100	-0.0100	0.0640	1
		$\hat{\gamma}_{1,2,0,9}$	0.0230	0.0230	0.0200	-0.0070	0.0610	1
	State 1	$\hat{\gamma}_{1,2,1,0}$	-0.0760	0.0360	-0.0800	-0.1230	0.0010	1
		$\hat{\gamma}_{1,2,1,1}$	-0.0100	0.0260	-0.0020	-0.0450	0.0320	1
		$\hat{\gamma}_{1,2,1,2}$	0.0020	0.0150	0.0000	-0.0190	0.0260	1
		$\hat{\gamma}_{1,2,1,3}$	0.0090	0.0160	0.0040	-0.0140	0.0340	1
		$\hat{\gamma}_{1,2,1,4}$	0.0160	0.0230	0.0070	-0.0130	0.0560	1
		$\hat{\gamma}_{1,2,1,5}$	0.0170	0.0230	0.0080	-0.0110	0.0580	1
		$\hat{\gamma}_{1,2,1,6}$	0.0110	0.0200	0.0030	-0.0150	0.0460	1
		$\hat{\gamma}_{1,2,1,7}$	0.0190	0.0270	0.0100	-0.0160	0.0640	1
		$\hat{\gamma}_{1,2,1,8}$	0.0190	0.0260	0.0090	-0.0110	0.0650	1
		$\hat{\gamma}_{1,2,1,9}$	0.0270	0.0240	0.0250	-0.0040	0.0630	1
IsPrecip	State 0	$\hat{\gamma}_{1,3,0,0}$	0.0030	0.0060	0.0010	-0.0060	0.0120	1
		$\hat{\gamma}_{1,3,0,1}$	0.0060	0.0080	0.0040	-0.0040	0.0190	1
		$\hat{\gamma}_{1,3,0,2}$	-0.0090	0.0090	-0.0080	-0.0230	0.0030	1
		$\hat{\gamma}_{1,3,0,3}$	0.0060	0.0080	0.0030	-0.0050	0.0200	1
		$\hat{\gamma}_{1,3,0,4}$	-0.0050	0.0070	-0.0030	-0.0190	0.0040	1
		$\hat{\gamma}_{1,3,0,5}$	0.0000	0.0060	0.0000	-0.0110	0.0100	1
		$\hat{\gamma}_{1,3,0,6}$	0.0020	0.0060	0.0000	-0.0070	0.0130	1
		$\hat{\gamma}_{1,3,0,7}$	0.0050	0.0070	0.0030	-0.0050	0.0170	1
		$\hat{\gamma}_{1,3,0,8}$	0.0020	0.0060	0.0000	-0.0080	0.0130	1
		$\hat{\gamma}_{1,3,0,9}$	-0.0010	0.0050	0.0000	-0.0110	0.0070	1
	State 1	$\hat{\gamma}_{1,3,1,0}$	-0.0010	0.0050	0.0000	-0.0100	0.0070	1
		$\hat{\gamma}_{1,3,1,1}$	-0.0020	0.0060	0.0000	-0.0140	0.0080	1
		$\hat{\gamma}_{1,3,1,2}$	0.0010	0.0070	0.0000	-0.0110	0.0120	1
		$\hat{\gamma}_{1,3,1,3}$	-0.0040	0.0070	-0.0020	-0.0170	0.0050	1
		$\hat{\gamma}_{1,3,1,4}$	0.0010	0.0060	0.0000	-0.0100	0.0120	1
		$\hat{\gamma}_{1,3,1,5}$	-0.0010	0.0060	0.0000	-0.0110	0.0100	1
		$\hat{\gamma}_{1,3,1,6}$	0.0020	0.0070	0.0010	-0.0070	0.0150	1
		$\hat{\gamma}_{1,3,1,7}$	-0.0030	0.0070	-0.0010	-0.0160	0.0070	1

		$\hat{\gamma}_{1,3,1,8}$	0.0030	0.0070	0.0010	-0.0060	0.0150	1
		$\hat{\gamma}_{1,3,1,9}$	0.0010	0.0050	0.0000	-0.0070	0.0110	1
IsSubAway	State 0	$\hat{\gamma}_{1,4,0,0}$	0.0010	0.0040	0.0000	-0.0040	0.0080	1
		$\hat{\gamma}_{1,4,0,1}$	0.0060	0.0050	0.0050	-0.0010	0.0140	1
		$\hat{\gamma}_{1,4,0,2}$	-0.0120	0.0050	-0.0130	-0.0210	-0.0030	1
		$\hat{\gamma}_{1,4,0,3}$	0.0000	0.0030	0.0000	-0.0060	0.0060	1
		$\hat{\gamma}_{1,4,0,4}$	0.0020	0.0040	0.0010	-0.0030	0.0090	1
		$\hat{\gamma}_{1,4,0,5}$	-0.0010	0.0030	0.0000	-0.0070	0.0050	1
		$\hat{\gamma}_{1,4,0,6}$	-0.0060	0.0050	-0.0060	-0.0140	0.0010	1
		$\hat{\gamma}_{1,4,0,7}$	-0.0040	0.0050	-0.0040	-0.0120	0.0020	1
		$\hat{\gamma}_{1,4,0,8}$	-0.0040	0.0050	-0.0030	-0.0120	0.0020	1
		$\hat{\gamma}_{1,4,0,9}$	0.0030	0.0040	0.0020	-0.0030	0.0100	1
	State 1	$\hat{\gamma}_{1,4,1,0}$	0.0000	0.0030	0.0000	-0.0050	0.0050	1
		$\hat{\gamma}_{1,4,1,1}$	-0.0010	0.0030	0.0000	-0.0070	0.0030	1
		$\hat{\gamma}_{1,4,1,2}$	0.0010	0.0030	0.0000	-0.0040	0.0050	1
		$\hat{\gamma}_{1,4,1,3}$	0.0000	0.0030	0.0000	-0.0040	0.0050	1
		$\hat{\gamma}_{1,4,1,4}$	0.0010	0.0030	0.0000	-0.0040	0.0060	1
		$\hat{\gamma}_{1,4,1,5}$	0.0010	0.0030	0.0000	-0.0040	0.0060	1
		$\hat{\gamma}_{1,4,1,6}$	-0.0050	0.0040	-0.0050	-0.0110	0.0010	1
		$\hat{\gamma}_{1,4,1,7}$	0.0000	0.0030	0.0000	-0.0050	0.0040	1
		$\hat{\gamma}_{1,4,1,8}$	0.0000	0.0030	0.0000	-0.0040	0.0050	1
		$\hat{\gamma}_{1,4,1,9}$	0.0000	0.0030	0.0000	-0.0050	0.0040	1
IsSubHome	State 0	$\hat{\gamma}_{1,5,0,0}$	-0.0030	0.0040	-0.0030	-0.0100	0.0020	1
		$\hat{\gamma}_{1,5,0,1}$	0.0020	0.0030	0.0010	-0.0030	0.0070	1
		$\hat{\gamma}_{1,5,0,2}$	-0.0020	0.0030	-0.0010	-0.0070	0.0030	1
		$\hat{\gamma}_{1,5,0,3}$	0.0010	0.0030	0.0000	-0.0040	0.0060	1
		$\hat{\gamma}_{1,5,0,4}$	0.0000	0.0030	0.0000	-0.0040	0.0060	1
		$\hat{\gamma}_{1,5,0,5}$	0.0000	0.0030	0.0000	-0.0040	0.0050	1
		$\hat{\gamma}_{1,5,0,6}$	-0.0040	0.0040	-0.0030	-0.0100	0.0020	1
		$\hat{\gamma}_{1,5,0,7}$	-0.0090	0.0040	-0.0090	-0.0160	-0.0020	1
		$\hat{\gamma}_{1,5,0,8}$	-0.0020	0.0030	-0.0010	-0.0080	0.0030	1
		$\hat{\gamma}_{1,5,0,9}$	0.0000	0.0030	0.0000	-0.0050	0.0050	1
	State 1	$\hat{\gamma}_{1,5,1,0}$	0.0000	0.0030	0.0000	-0.0050	0.0070	1
		$\hat{\gamma}_{1,5,1,1}$	0.0040	0.0040	0.0030	-0.0020	0.0110	1
		$\hat{\gamma}_{1,5,1,2}$	-0.0030	0.0040	-0.0020	-0.0100	0.0020	1
		$\hat{\gamma}_{1,5,1,3}$	0.0020	0.0040	0.0010	-0.0040	0.0080	1
		$\hat{\gamma}_{1,5,1,4}$	-0.0090	0.0050	-0.0100	-0.0170	0.0000	1
		$\hat{\gamma}_{1,5,1,5}$	-0.0140	0.0050	-0.0140	-0.0220	-0.0060	1
		$\hat{\gamma}_{1,5,1,6}$	-0.0010	0.0040	0.0000	-0.0080	0.0040	1
		$\hat{\gamma}_{1,5,1,7}$	0.0100	0.0050	0.0100	0.0000	0.0170	1
		$\hat{\gamma}_{1,5,1,8}$	-0.0040	0.0040	-0.0030	-0.0120	0.0020	1
		$\hat{\gamma}_{1,5,1,9}$	-0.0010	0.0030	0.0000	-0.0070	0.0050	1
RelHumid	State 0	$\hat{\gamma}_{1,6,0,0}$	0.0020	0.0170	0.0000	-0.0240	0.0280	1
		$\hat{\gamma}_{1,6,0,1}$	0.0000	0.0180	0.0000	-0.0270	0.0280	1
		$\hat{\gamma}_{1,6,0,2}$	0.0090	0.0210	0.0050	-0.0220	0.0400	1
		$\hat{\gamma}_{1,6,0,3}$	-0.0040	0.0180	-0.0010	-0.0320	0.0230	1
		$\hat{\gamma}_{1,6,0,4}$	-0.0010	0.0180	0.0000	-0.0280	0.0270	1
		$\hat{\gamma}_{1,6,0,5}$	-0.0050	0.0200	-0.0020	-0.0350	0.0230	1
		$\hat{\gamma}_{1,6,0,6}$	-0.0020	0.0180	0.0000	-0.0290	0.0240	1

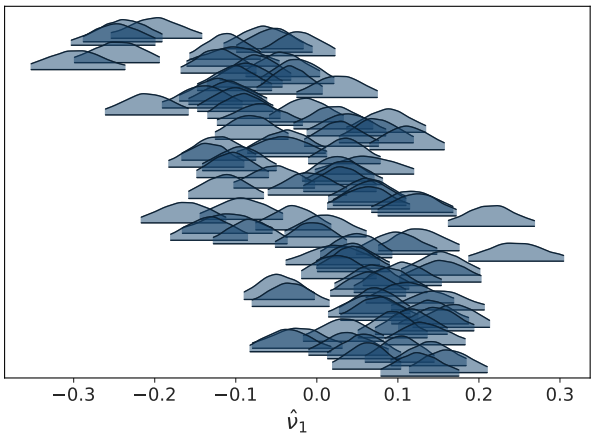
		$\hat{\gamma}_{1,6,0,7}$	-0.0080	0.0200	-0.0050	-0.0410	0.0210	1
		$\hat{\gamma}_{1,6,0,8}$	0.0030	0.0170	0.0010	-0.0240	0.0290	1
		$\hat{\gamma}_{1,6,0,9}$	0.0020	0.0160	0.0000	-0.0210	0.0280	1
	State 1	$\hat{\gamma}_{1,6,1,0}$	0.0040	0.0170	0.0000	-0.0200	0.0330	1
		$\hat{\gamma}_{1,6,1,1}$	0.0060	0.0180	0.0020	-0.0190	0.0350	1
		$\hat{\gamma}_{1,6,1,2}$	-0.0040	0.0200	-0.0020	-0.0340	0.0290	1
		$\hat{\gamma}_{1,6,1,3}$	0.0030	0.0180	0.0010	-0.0200	0.0340	1
		$\hat{\gamma}_{1,6,1,4}$	-0.0010	0.0180	0.0000	-0.0270	0.0280	1
		$\hat{\gamma}_{1,6,1,5}$	0.0040	0.0200	0.0010	-0.0230	0.0360	1
		$\hat{\gamma}_{1,6,1,6}$	0.0010	0.0180	0.0000	-0.0260	0.0280	1
		$\hat{\gamma}_{1,6,1,7}$	0.0060	0.0200	0.0030	-0.0200	0.0420	1
		$\hat{\gamma}_{1,6,1,8}$	-0.0040	0.0170	-0.0010	-0.0320	0.0210	1
		$\hat{\gamma}_{1,6,1,9}$	0.0000	0.0160	0.0000	-0.0240	0.0250	1
Temp	State 0	$\hat{\gamma}_{1,7,0,0}$	-0.0070	0.0210	-0.0040	-0.0420	0.0230	1
		$\hat{\gamma}_{1,7,0,1}$	-0.0020	0.0190	0.0000	-0.0290	0.0290	1
		$\hat{\gamma}_{1,7,0,2}$	0.0010	0.0200	0.0000	-0.0250	0.0340	1
		$\hat{\gamma}_{1,7,0,3}$	-0.0060	0.0220	-0.0030	-0.0410	0.0250	1
		$\hat{\gamma}_{1,7,0,4}$	0.0040	0.0200	0.0000	-0.0240	0.0350	1
		$\hat{\gamma}_{1,7,0,5}$	0.0070	0.0200	0.0030	-0.0200	0.0420	1
		$\hat{\gamma}_{1,7,0,6}$	0.0010	0.0190	0.0000	-0.0280	0.0320	1
		$\hat{\gamma}_{1,7,0,7}$	-0.0010	0.0200	0.0000	-0.0310	0.0310	1
		$\hat{\gamma}_{1,7,0,8}$	0.0020	0.0190	0.0000	-0.0250	0.0330	1
		$\hat{\gamma}_{1,7,0,9}$	0.0010	0.0180	0.0000	-0.0300	0.0270	1
	State 1	$\hat{\gamma}_{1,7,1,0}$	0.0050	0.0210	0.0030	-0.0250	0.0400	1
		$\hat{\gamma}_{1,7,1,1}$	0.0010	0.0190	0.0000	-0.0250	0.0320	1
		$\hat{\gamma}_{1,7,1,2}$	0.0000	0.0190	0.0000	-0.0270	0.0310	1
		$\hat{\gamma}_{1,7,1,3}$	0.0080	0.0220	0.0040	-0.0230	0.0440	1
		$\hat{\gamma}_{1,7,1,4}$	0.0000	0.0200	0.0000	-0.0290	0.0300	1
		$\hat{\gamma}_{1,7,1,5}$	-0.0040	0.0200	-0.0010	-0.0370	0.0260	1
		$\hat{\gamma}_{1,7,1,6}$	0.0030	0.0190	0.0000	-0.0270	0.0330	1
		$\hat{\gamma}_{1,7,1,7}$	0.0050	0.0200	0.0020	-0.0220	0.0390	1
		$\hat{\gamma}_{1,7,1,8}$	0.0010	0.0190	0.0000	-0.0300	0.0280	1
		$\hat{\gamma}_{1,7,1,9}$	0.0010	0.0180	0.0000	-0.0250	0.0310	1
WaterSls	State 0	$\hat{\gamma}_{1,8,0,0}$	0.0080	0.0080	0.0070	-0.0030	0.0210	1
		$\hat{\gamma}_{1,8,0,1}$	0.0000	0.0060	0.0000	-0.0090	0.0100	1
		$\hat{\gamma}_{1,8,0,2}$	0.0030	0.0060	0.0010	-0.0060	0.0140	1
		$\hat{\gamma}_{1,8,0,3}$	-0.0010	0.0060	0.0000	-0.0110	0.0080	1
		$\hat{\gamma}_{1,8,0,4}$	-0.0030	0.0060	-0.0010	-0.0140	0.0070	1
		$\hat{\gamma}_{1,8,0,5}$	-0.0010	0.0060	0.0000	-0.0110	0.0090	1
		$\hat{\gamma}_{1,8,0,6}$	0.0010	0.0060	0.0000	-0.0090	0.0110	1
		$\hat{\gamma}_{1,8,0,7}$	0.0100	0.0080	0.0100	-0.0020	0.0230	1
		$\hat{\gamma}_{1,8,0,8}$	0.0020	0.0060	0.0010	-0.0060	0.0140	1
		$\hat{\gamma}_{1,8,0,9}$	0.0000	0.0050	0.0000	-0.0090	0.0090	1
	State 1	$\hat{\gamma}_{1,8,1,0}$	0.0060	0.0070	0.0050	-0.0030	0.0190	1
		$\hat{\gamma}_{1,8,1,1}$	0.0030	0.0060	0.0010	-0.0060	0.0140	1
		$\hat{\gamma}_{1,8,1,2}$	0.0000	0.0060	0.0000	-0.0100	0.0100	1
		$\hat{\gamma}_{1,8,1,3}$	-0.0020	0.0060	-0.0010	-0.0140	0.0060	1
		$\hat{\gamma}_{1,8,1,4}$	0.0010	0.0060	0.0000	-0.0090	0.0110	1
		$\hat{\gamma}_{1,8,1,5}$	0.0030	0.0060	0.0010	-0.0060	0.0150	1

		$\hat{\gamma}_{1,8,1,6}$	-0.0010	0.0060	0.0000	-0.0110	0.0090	1
		$\hat{\gamma}_{1,8,1,7}$	0.0050	0.0070	0.0030	-0.0050	0.0180	1
		$\hat{\gamma}_{1,8,1,8}$	0.0010	0.0060	0.0000	-0.0090	0.0110	1
		$\hat{\gamma}_{1,8,1,9}$	-0.0040	0.0070	-0.0030	-0.0160	0.0040	1
Wind	State 0	$\hat{\gamma}_{1,9,0,0}$	-0.0070	0.0110	-0.0040	-0.0260	0.0090	1
		$\hat{\gamma}_{1,9,0,1}$	0.0010	0.0120	0.0000	-0.0180	0.0210	1
		$\hat{\gamma}_{1,9,0,2}$	0.0010	0.0120	0.0000	-0.0170	0.0210	1
		$\hat{\gamma}_{1,9,0,3}$	-0.0030	0.0140	-0.0010	-0.0260	0.0170	1
		$\hat{\gamma}_{1,9,0,4}$	0.0000	0.0130	0.0000	-0.0220	0.0200	1
		$\hat{\gamma}_{1,9,0,5}$	0.0120	0.0180	0.0080	-0.0110	0.0430	1
		$\hat{\gamma}_{1,9,0,6}$	-0.0010	0.0200	-0.0010	-0.0340	0.0300	1
		$\hat{\gamma}_{1,9,0,7}$	0.0060	0.0160	0.0040	-0.0170	0.0340	1
		$\hat{\gamma}_{1,9,0,8}$	-0.0100	0.0160	-0.0060	-0.0370	0.0110	1
		$\hat{\gamma}_{1,9,0,9}$	-0.0050	0.0140	0.0000	-0.0290	0.0160	1
	State 1	$\hat{\gamma}_{1,9,1,0}$	0.0000	0.0110	0.0000	-0.0160	0.0180	1
		$\hat{\gamma}_{1,9,1,1}$	-0.0070	0.0130	-0.0030	-0.0280	0.0090	1
		$\hat{\gamma}_{1,9,1,2}$	-0.0020	0.0120	0.0000	-0.0210	0.0160	1
		$\hat{\gamma}_{1,9,1,3}$	0.0060	0.0140	0.0020	-0.0150	0.0270	1
		$\hat{\gamma}_{1,9,1,4}$	0.0030	0.0140	0.0000	-0.0170	0.0250	1
		$\hat{\gamma}_{1,9,1,5}$	0.0020	0.0170	0.0000	-0.0250	0.0310	1
		$\hat{\gamma}_{1,9,1,6}$	0.0200	0.0210	0.0180	-0.0070	0.0550	1
		$\hat{\gamma}_{1,9,1,7}$	-0.0090	0.0170	-0.0070	-0.0360	0.0150	1
		$\hat{\gamma}_{1,9,1,8}$	0.0000	0.0150	0.0000	-0.0250	0.0240	1
		$\hat{\gamma}_{1,9,1,9}$	-0.0110	0.0150	-0.0070	-0.0350	0.0080	1
AvgClsOver05Prob	-	$\hat{\gamma}_{2,0}$	-0.0200	0.0300	-0.0120	-0.0760	0.0190	1
AvgClsOver15Prob	-	$\hat{\gamma}_{2,1}$	-0.0110	0.0340	-0.0030	-0.0700	0.0440	1
AvgClsOver25Prob	-	$\hat{\gamma}_{2,2}$	-0.0160	0.0420	-0.0050	-0.0900	0.0480	1
AvgClsOver35Prob	-	$\hat{\gamma}_{2,3}$	0.0060	0.0440	0.0010	-0.0640	0.0860	1
AvgClsOver45Prob	-	$\hat{\gamma}_{2,4}$	0.0340	0.0460	0.0240	-0.0270	0.1130	1
AvgClsOver55Prob	-	$\hat{\gamma}_{2,5}$	0.0050	0.0330	0.0010	-0.0440	0.0660	1
AvgClsProbAway	-	$\hat{\gamma}_{2,6}$	-0.0030	0.0190	-0.0020	-0.0350	0.0260	1
AvgClsProbHome	-	$\hat{\gamma}_{2,7}$	0.0070	0.0190	0.0030	-0.0200	0.0370	1
BeerPrice	-	$\hat{\gamma}_{2,8}$	0.0180	0.0170	0.0170	-0.0060	0.0470	1
GeoDist	-	$\hat{\gamma}_{2,9}$	0.0080	0.0070	0.0080	-0.0010	0.0180	1
HotWinePrice	-	$\hat{\gamma}_{2,10}$	0.0170	0.0170	0.0160	-0.0060	0.0450	1
Is1stDiv	-	$\hat{\gamma}_{2,11}$	-0.0450	0.0240	-0.0430	-0.0780	0.0010	1
IsMobCpoint	-	$\hat{\gamma}_{2,12}$	-0.0010	0.0100	0.0000	-0.0180	0.0170	1
IsPubHoliday	-	$\hat{\gamma}_{2,13}$	0.0020	0.0050	0.0000	-0.0060	0.0100	1
IsRelegAway	-	$\hat{\gamma}_{2,14}$	0.0020	0.0050	0.0010	-0.0050	0.0110	1
IsRunner	-	$\hat{\gamma}_{2,15}$	0.0010	0.0080	0.0000	-0.0120	0.0160	1
IsSchlHoliday	-	$\hat{\gamma}_{2,16}$	0.0010	0.0060	0.0000	-0.0090	0.0110	1
IsSoldOut	-	$\hat{\gamma}_{2,17}$	0.0020	0.0060	0.0010	-0.0060	0.0150	1
IsWasen	-	$\hat{\gamma}_{2,18}$	-0.0020	0.0050	-0.0010	-0.0120	0.0060	1
Kickoff15:30	-	$\hat{\gamma}_{2,19}$	0.0290	0.0250	0.0240	-0.0070	0.0700	1
Kickoff15:45	-	$\hat{\gamma}_{2,20}$	0.0100	0.0090	0.0090	-0.0020	0.0240	1
Kickoff17:30	-	$\hat{\gamma}_{2,21}$	0.0110	0.0150	0.0070	-0.0070	0.0370	1
Kickoff18:00	-	$\hat{\gamma}_{2,22}$	0.0070	0.0090	0.0040	-0.0050	0.0240	1
Kickoff18:30	-	$\hat{\gamma}_{2,23}$	0.0300	0.0150	0.0300	0.0070	0.0550	1
Kickoff20:30	-	$\hat{\gamma}_{2,24}$	0.0430	0.0250	0.0410	-0.0010	0.0800	1

MonthAug	-	$\hat{\gamma}_{2,25}$	0.0040	0.0080	0.0010	-0.0080	0.0170	1
MonthDec	-	$\hat{\gamma}_{2,26}$	0.0020	0.0060	0.0010	-0.0060	0.0120	1
MonthFeb	-	$\hat{\gamma}_{2,27}$	0.0000	0.0060	0.0000	-0.0090	0.0100	1
MonthJan	-	$\hat{\gamma}_{2,28}$	-0.0050	0.0060	-0.0040	-0.0160	0.0040	1
MonthMar	-	$\hat{\gamma}_{2,29}$	0.0000	0.0050	0.0000	-0.0090	0.0070	1
MonthMay	-	$\hat{\gamma}_{2,30}$	0.0030	0.0060	0.0010	-0.0050	0.0130	1
MonthNov	-	$\hat{\gamma}_{2,31}$	-0.0040	0.0060	-0.0030	-0.0150	0.0050	1
MonthOct	-	$\hat{\gamma}_{2,32}$	0.0030	0.0060	0.0010	-0.0060	0.0130	1
MonthSep	-	$\hat{\gamma}_{2,33}$	0.0030	0.0080	0.0010	-0.0080	0.0160	1
NumCpoints	-	$\hat{\gamma}_{2,34}$	0.0130	0.0170	0.0100	-0.0090	0.0430	1
NumSpects	-	$\hat{\gamma}_{2,35}$	0.0510	0.0110	0.0510	0.0320	0.0690	1
RankAwayLast	-	$\hat{\gamma}_{2,36}$	0.0000	0.0040	0.0000	-0.0080	0.0070	1
RankAwayPre	-	$\hat{\gamma}_{2,37}$	-0.0010	0.0060	0.0000	-0.0130	0.0080	1
RankHomePre	-	$\hat{\gamma}_{2,38}$	-0.0010	0.0080	0.0000	-0.0150	0.0130	1
Round	-	$\hat{\gamma}_{2,39}$	0.0020	0.0090	0.0000	-0.0130	0.0160	1
Season2014/15	-	$\hat{\gamma}_{2,40}$	0.0000	0.0090	0.0000	-0.0140	0.0170	1
Season2015/16	-	$\hat{\gamma}_{2,41}$	0.0000	0.0100	0.0000	-0.0180	0.0140	1
Season2017/18	-	$\hat{\gamma}_{2,42}$	0.0040	0.0120	0.0010	-0.0130	0.0260	1
Season2018/19	-	$\hat{\gamma}_{2,43}$	0.0090	0.0150	0.0040	-0.0100	0.0360	1
WkdayMo	-	$\hat{\gamma}_{2,44}$	-0.0100	0.0130	-0.0070	-0.0310	0.0070	1
WkdaySa	-	$\hat{\gamma}_{2,45}$	0.0170	0.0230	0.0150	-0.0130	0.0570	1
WkdaySu	-	$\hat{\gamma}_{2,46}$	-0.0180	0.0190	-0.0170	-0.0480	0.0090	1
WkdayTu	-	$\hat{\gamma}_{2,47}$	-0.0040	0.0060	-0.0020	-0.0160	0.0050	1
WkdayWe	-	$\hat{\gamma}_{2,48}$	0.0040	0.0090	0.0010	-0.0090	0.0220	1
BeerSls	State 0	$\hat{\phi}_{1,0,1}$	0.1930	0.0100	0.1930	0.1770	0.2080	1
		$\hat{\phi}_{1,0,2}$	0.1730	0.0100	0.1730	0.1570	0.1890	1
		$\hat{\phi}_{1,0,3}$	0.0580	0.0100	0.0580	0.0420	0.0740	1
		$\hat{\phi}_{1,0,4}$	0.0270	0.0110	0.0280	0.0100	0.0450	1
		$\hat{\phi}_{1,0,5}$	0.0340	0.0110	0.0340	0.0170	0.0520	1
		$\hat{\phi}_{1,0,6}$	0.0100	0.0100	0.0090	-0.0030	0.0270	1
		$\hat{\phi}_{1,0,7}$	0.0210	0.0100	0.0210	0.0040	0.0380	1
		$\hat{\phi}_{1,0,8}$	-0.0140	0.0090	-0.0140	-0.0280	0.0010	1
		$\hat{\phi}_{1,0,9}$	-0.0040	0.0070	-0.0020	-0.0170	0.0070	1
	State 1	$\hat{\phi}_{1,1,1}$	0.1820	0.0100	0.1820	0.1660	0.1980	1
		$\hat{\phi}_{1,1,2}$	0.1760	0.0100	0.1760	0.1600	0.1930	1
		$\hat{\phi}_{1,1,3}$	0.0660	0.0100	0.0660	0.0500	0.0830	1
		$\hat{\phi}_{1,1,4}$	0.0240	0.0110	0.0250	0.0060	0.0430	1
		$\hat{\phi}_{1,1,5}$	0.0170	0.0110	0.0170	-0.0010	0.0330	1
		$\hat{\phi}_{1,1,6}$	0.0250	0.0100	0.0250	0.0080	0.0420	1
		$\hat{\phi}_{1,1,7}$	0.0090	0.0090	0.0070	-0.0040	0.0250	1
		$\hat{\phi}_{1,1,8}$	0.0060	0.0080	0.0040	-0.0040	0.0220	1
		$\hat{\phi}_{1,1,9}$	-0.0100	0.0090	-0.0100	-0.0240	0.0020	1
Noise	-	$\hat{\sigma}$	0.2560	0.0020	0.2560	0.2530	0.2600	1
SlabWitdh	-	c^2	0.0110	0.0010	0.0110	0.0090	0.0130	1

This table presents summary statistics on the posterior distributions of the coefficients, estimated in the main specification. We exclude the match minute dummies for presentation reasons. Columns 7 and 8 include the 10% and 9% quantiles of the highest density interval, respectively. The last column provides the Gelman-Rubin statistic. Indices of coefficients denote the level, the variable, the state, and the lag. For ϕ, we drop the variable index, because it is redundant. The median for $\beta_{1,1,1,3}$ equal to 0.0190 implies: a one standard deviation increase of *Surprise* in positive state 1 increases the conditional mean of beer sales during minute 3 after a key match event by approximately 1.9%, ceteris paribus.

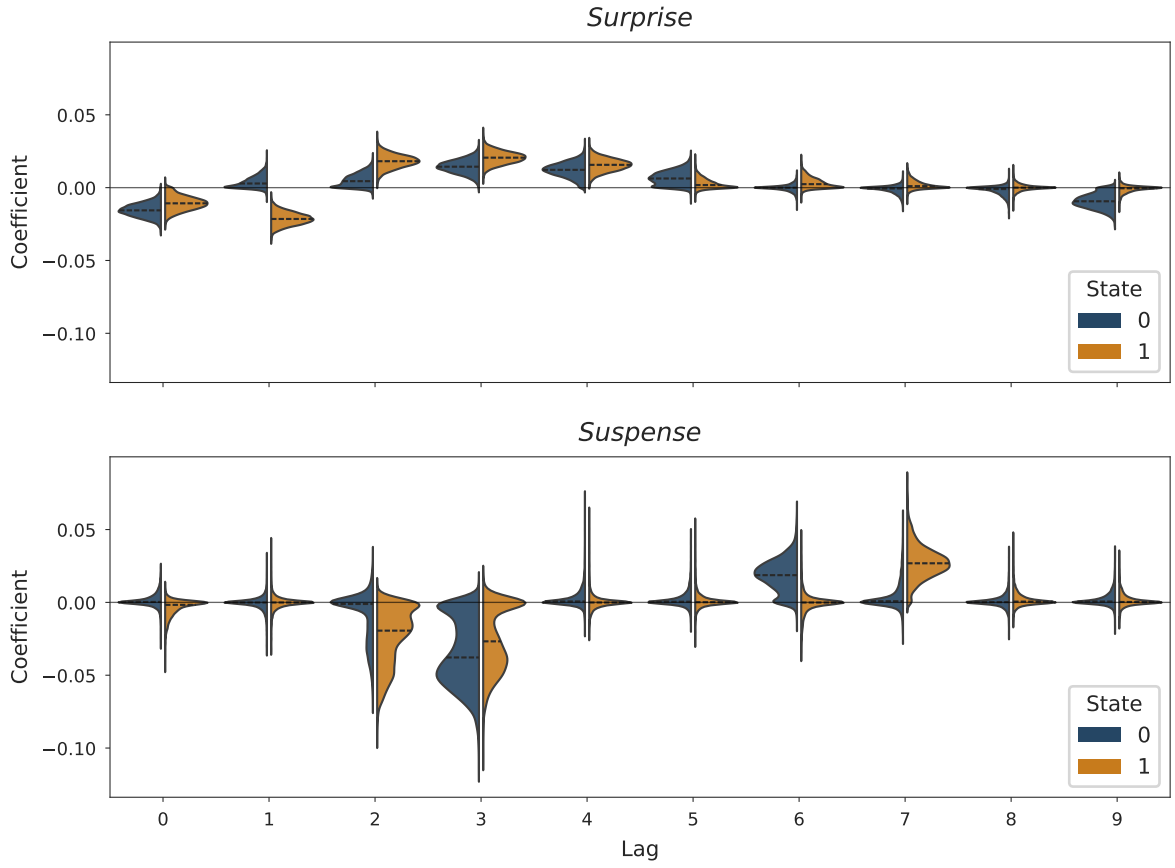
Figure A.6
Main Model Forest Plot



This figure shows posterior distributions of the intercept $\hat{\nu}_1$ for each match. The matches are sorted chronologically from top to bottom (seasons 2013/14 to 2018/19).

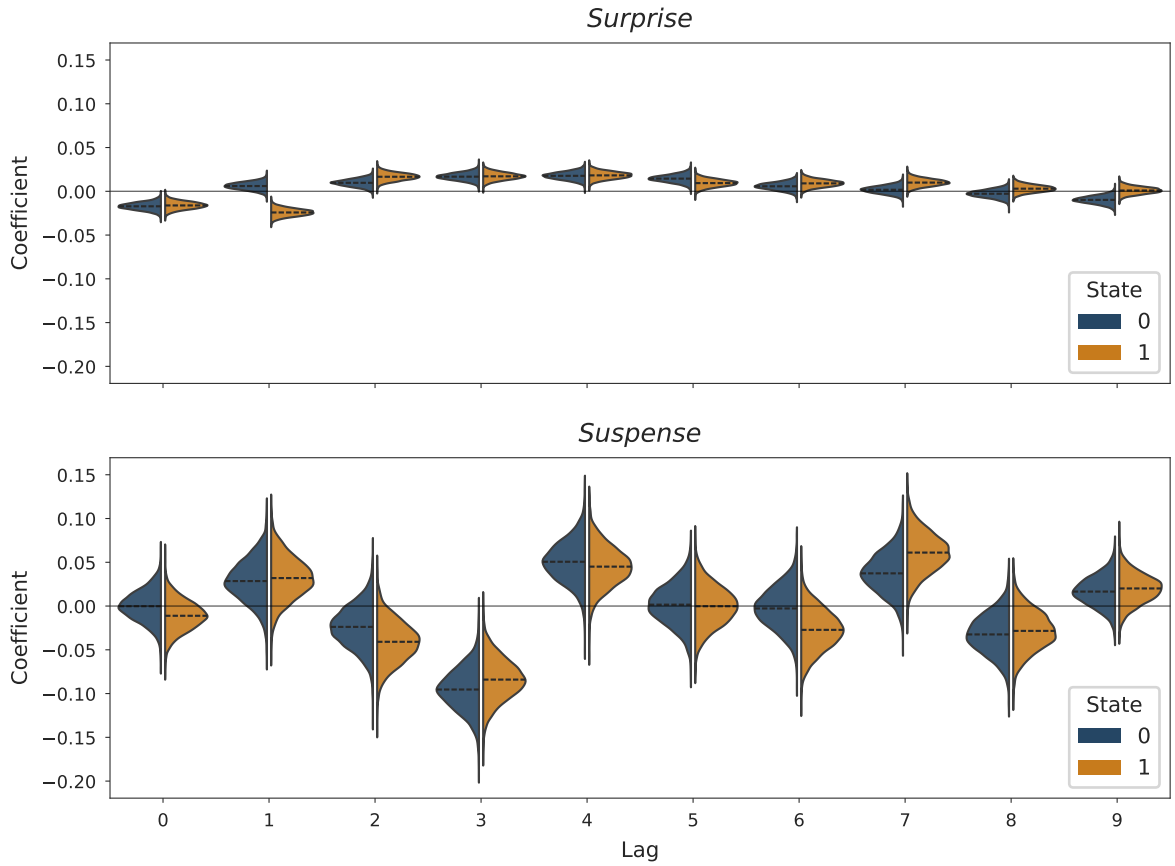
A.3.2 Extended Specifications and Robustness Checks

Figure A.7
Regularized Horseshoe Posterior Distributions



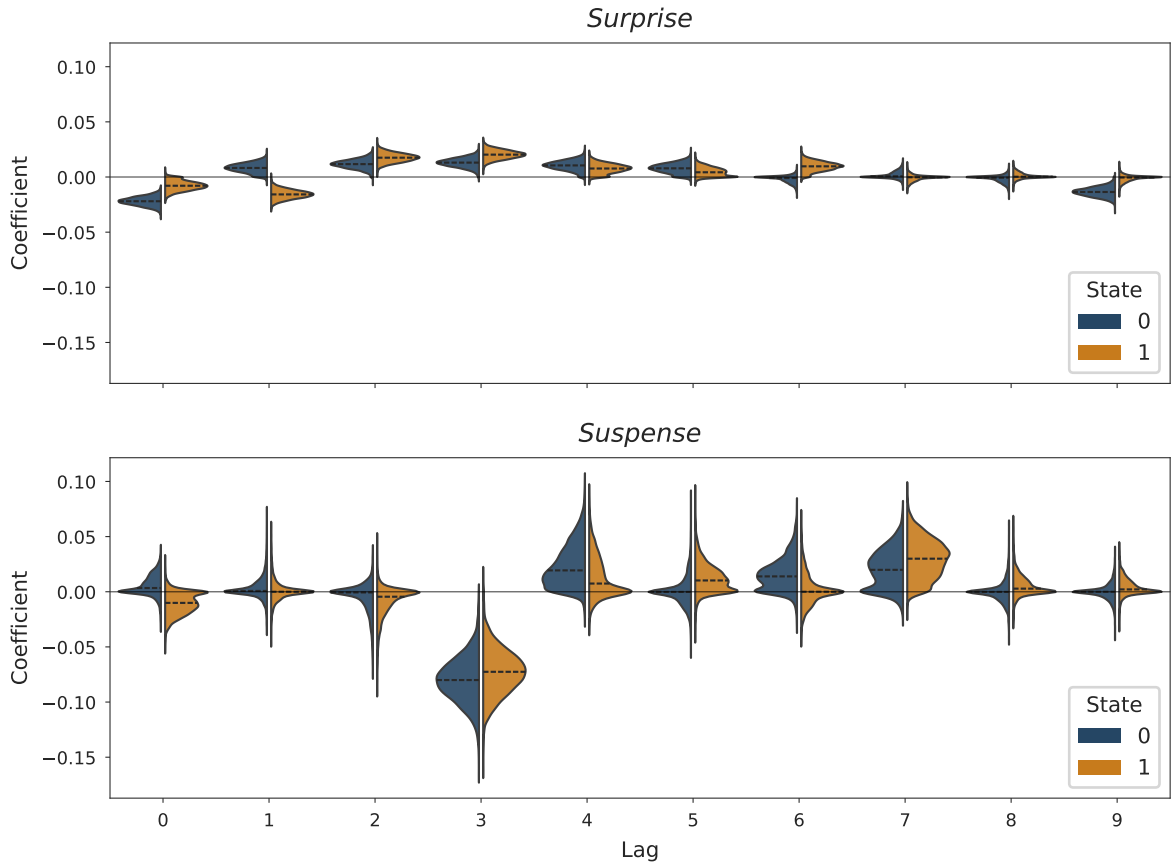
Equivalent description as for Figure 2.3, but for the regularized horseshoe specification.

Figure A.8
Normal Prior Posterior Distributions



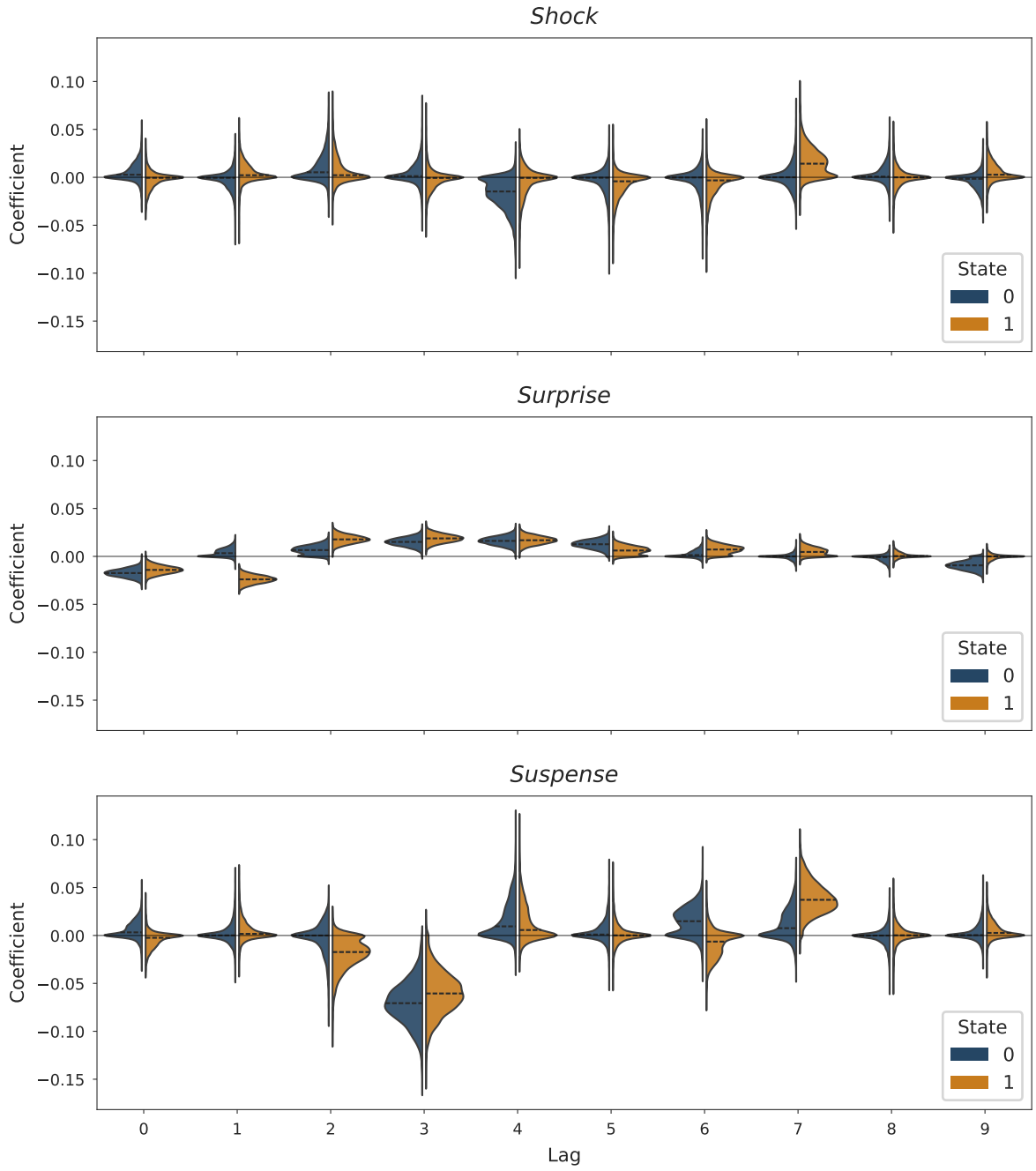
Equivalent description as for Figure 2.3, but for the normal prior specification.

Figure A.9
Seasons Posterior Distributions



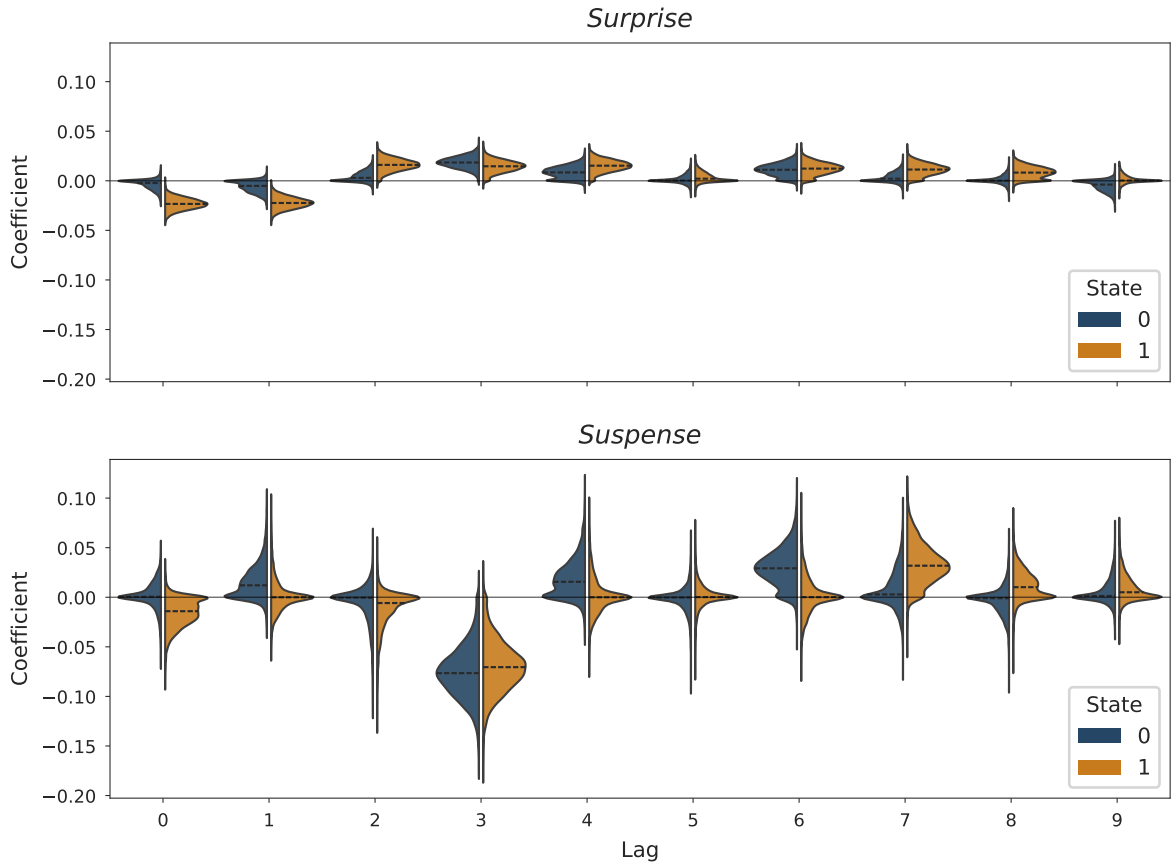
Equivalent description as for Figure 2.3, but for the seasons specification.

Figure A.10
Shock Posterior Distributions



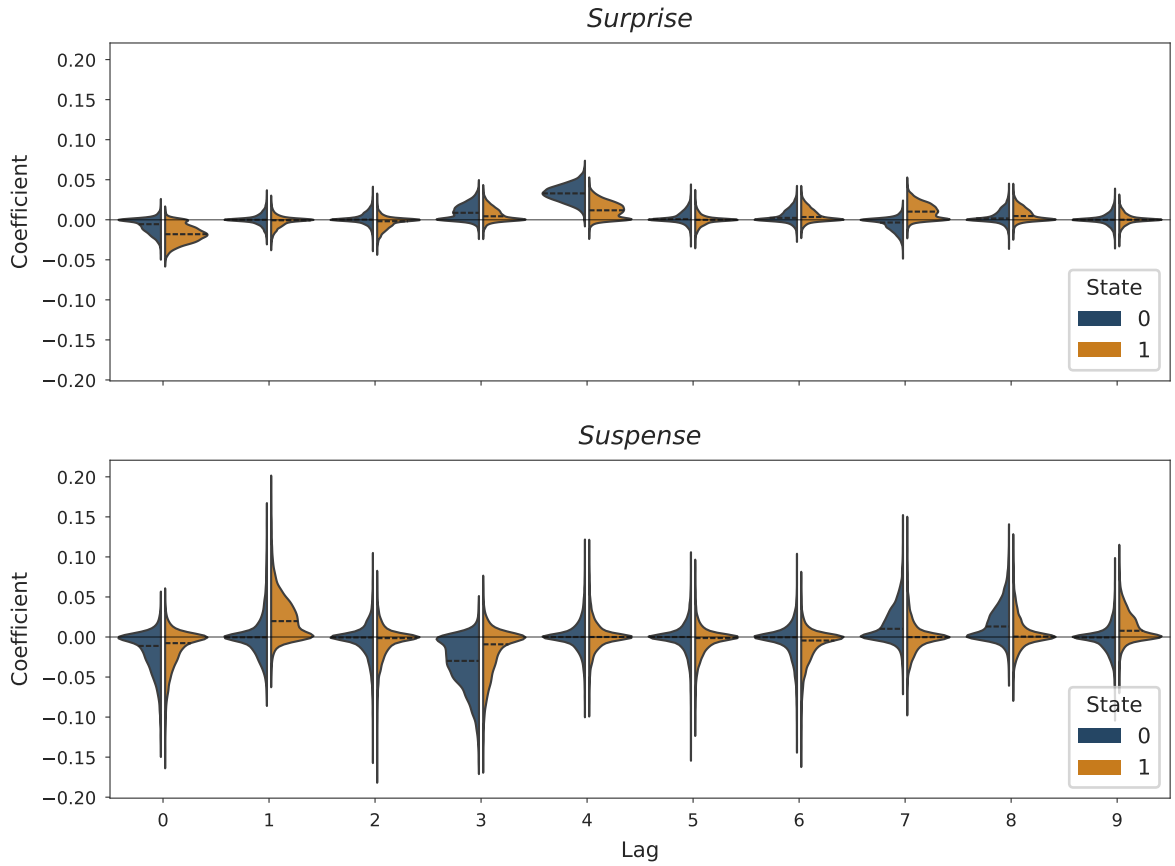
Equivalent description as for Figure 2.3, but for the shock specification.

Figure A.11
Die-Hard Fans Posterior Distributions



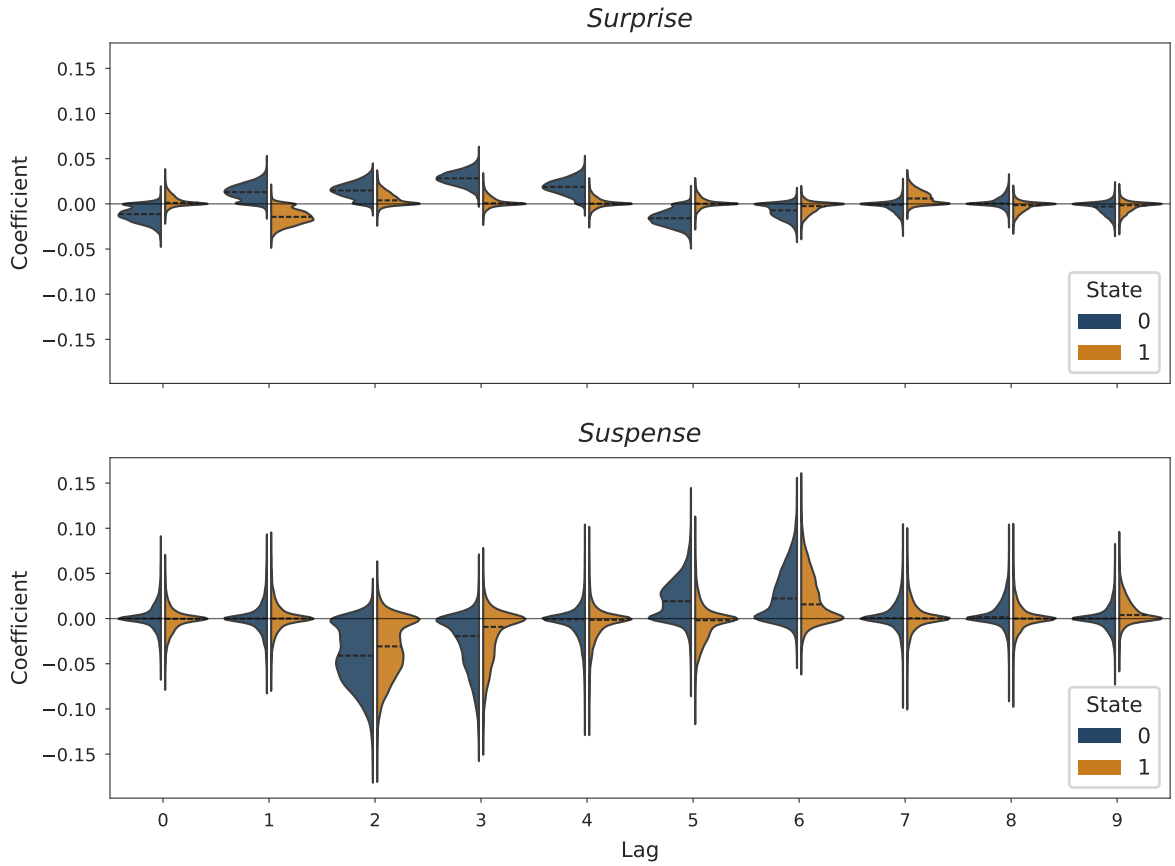
Equivalent description as for Figure 2.3, but for the die-hard fans specification.

Figure A.12
Shandy Posterior Distributions



Equivalent description as for Figure 2.3, but for the shandy specification.

Figure A.13
Soft Drinks Posterior Distributions



Equivalent description as for Figure 2.3, but for the soft drinks specification.

Chapter 3

An Alternative Prior for Estimation in High-Dimensional Settings[21]

3.1 Motivation

Applied research in both natural and social sciences, such as genetics, psychology, and economics, often deals with intricate models using high-dimensional parameter spaces that incorporate several interactions, splines, and quadratic effects (e.g., Watt 2004; Engle and Rangel 2008; Brandt, Cambria, and Kelava 2018). In such settings, estimation techniques such as ordinary least squares tend to provide biased parameter estimates with low accuracy, because asymptotic theories do not hold when the number of predictors D is greater than the number of observations N. When the design matrix is sparse or the data feature high multicollinearity, dimensionality reduction is of particular importance.

To address problems associated with overfitting and overparametrization (Williams and Rodriguez 2022), frequentist approaches apply penalized regression methods such as L1 or L2 regularization (Marquardt and Snee 1975; Tibshirani 1996). In Bayesian frameworks, regularization can be induced through the specification of prior distributions, which allocate substantial probability mass around 0 to shrink small coefficients toward 0. The heavy tails of these distributions create a means to accommodate large coefficients. Renowned examples of such approaches include the spike-and-slab prior (Mitchell and Beauchamp 1988; George and McCulloch 1993; Ishwaran and Rao 2005) and the Bayesian lasso (Park and Casella 2008). Although these approaches serve as good starting points, various issues emerge, related to their practical use.

Specifically, spike-and-slab priors were considered the gold standard in sparse Bayesian estimation for a long time because of their well-understood theoretical properties and good performance in a variety of settings related to sparsity. However, they become computationally intensive, particularly when used in high-dimensional spaces, because posterior sampling requires a stochastic search over a tremendous space, leading to slow mixing and convergence (Mitchell and Beauchamp 1988; Bhadra et al. 2017). As noted by Piironen and Vehtari (2017), their results also can be sensitive to prior choices (i.e., slab width and prior inclusion probability), which can be difficult to determine (Ishwaran and Rao 2005). Moreover, Mitchell and Beauchamp (1988) point out challenges in modeling correlated predictors and class variables with more than two levels.

The popularity of the Bayesian lasso stems largely from its simplicity and computational efficiency (Piironen and Vehtari 2017). As Park and Casella (2008) note, the Bayesian lasso ensures a

[21] This chapter is a revised version of Nagel et al. (2024), published in Structural Equation Modeling: A Multidisciplinary Journal. A replication package is available at https://github.com/michael-nagel/dhs

unimodal posterior, which facilitates fast convergence and produces meaningful estimates. However, Carvalho, Polson, and Scott (2008) and Armagan, Dunson, and Lee (2013) caution that it also insufficiently shrinks small coefficients (representing noise) while overshrinking large coefficients due to its relatively light tails, such that it is less robust to outliers—a notable drawback.

Carvalho, Polson, and Scott (2009) also identify two other factors that hamper the performance of the Bayesian lasso method: overestimation of the signal share in the parameter vector and underestimation of the presence of large coefficients. Finally, due to its reliance on a single hyperparameter in the double exponential (Laplace) prior, the Bayesian lasso suffers restricted flexibility and adaptivity to different problems (Brown and Griffin 2010).

Some more recently developed priors include, for example, continuous shrinkage priors, which sometimes are expressed as one-group global-local scale mixtures of Gaussians, like the horseshoe (Carvalho, Polson, and Scott 2010). These priors achieve good performance in high-dimensional settings with regard to sparsity (Van Der Pas, Kleijn, and Van Der Vaart 2014), exhibit adaptivity to unknown sparsity and signal-to-noise ratios (Carvalho, Polson, and Scott 2008), and offer computational efficiency, in the sense that their complexity depends on the sparsity ratio rather than the number of observations (Van Der Pas, Kleijn, and Van Der Vaart 2014).

Despite its popularity, some shortcomings of the horseshoe limit its practical use. In particular, Piironen and Vehtari (2017) highlight the lack of a possibility to incorporate prior knowledge about the sparsity level in the parameter vector for the global shrinkage hyperparameter. They assert that the default choices encourage solutions in which too many parameters remain unshrunk. Van Der Pas, Kleijn, and Van Der Vaart (2014) also report both over- and underestimation of the sparsity ratio when applying full Bayes implementation with a Cauchy prior on the global hyperparameter. Notably, the horseshoe's shrinkage profile keeps large coefficients unshrunk, which can be a favorable property, but it also can deteriorate performance when the parameters are only weakly identified by the data, such as when dealing with flat likelihoods emerging in logistic regressions with data separation (Piironen and Vehtari 2017).

In addition, it can cause the means of the posteriors to vanish, as observed for the Cauchy prior (Ghosh, Li, and Mitra 2018). Another concern stems from the multimodality of the posterior, caused by correlated predictors, which can raise sampling and convergence issues for the inference algorithms (Piironen and Vehtari 2017). Noticeable convergence issues have also been reported by Piironen and Vehtari (2015), even for simple regression problems.

More advanced priors include the regularized horseshoe (Piironen and Vehtari 2017) and Dirichlet-Laplace priors (Bhattacharya et al. 2015). The former is a generalization of the horseshoe and nests the horseshoe as a special case. Accordingly, the regularized horseshoe is commonly applied to high-dimensional regression problems in which the number of predictors is much greater than the number of observations, due to its strong performance in such settings. Furthermore, it can incorporate prior knowledge about the degree of sparsity in the global hyperparameter.

By adding a small degree of regularization to the large coefficients, the regularized horseshoe yields more reasonable parameter estimates, ensures that the posterior mean exists, and avoids sampling issues related to divergent transitions (Piironen and Vehtari 2017). Although these improvements lead to improved performance and applicability of the horseshoe, as Piironen and Vehtari (2017) emphasize, the issues related to multimodality due to correlated predictors still persist. Furthermore, as they demonstrate, a misspecification of the hyperparameters can result in overly strong regularization.

Dirichlet-Laplace priors unify Dirichlet and Laplace priors to induce sparsity in models. Whereas the Laplacian part controls the overall amount of shrinkage, concentrating the Dirichlet elements constrains the number of non-zero coefficients. Thus, Dirichlet-Laplace priors are theoretically sound; they provide a closed-form representation that is appealing theoretically and also allows for an exact representation of the posterior.

According to Bhattacharya et al. (2015), Dirichlet-Laplace priors yield efficient posterior computations, with an appropriate chosen concentration rate, such that they offer an optimal posterior concentration (minimax rate of convergence). Nevertheless, their practical use is limited, because the results depend powerfully on the commonly unknown, hard-to-determine concentration parameter, which can exert substantial effects on the results, as demonstrated by Bhattacharya et al. (2015). In Table 3.1, we provide an overview of the properties of these priors.

Table 3.1
Comparison of Properties of Prior Distributions

Spike-and-Slab	Bayesian (Adaptive) Lasso
• discrete (point-mass mixture) prior; continuous versions exist • strong sparsity-inducing performance in many settings • well-understood theoretical properties • computationally intensive in high-dimensional settings • highly sensitive to hyperprior choices • struggles with modeling correlated predictors • challenges handling class variables with more than two levels	• continuous shrinkage prior • simple and computationally efficient • guarantees unimodal posterior distributions • limited flexibility, less adaptive to varying signal strengths • tends to undershrink noise and overshrink large coefficients • may overestimate signal density
Horseshoe	**Dirichlet-Laplace**
• continuous shrinkage prior • strong performance in high-dimensional, sparse problems • adapts to unknown sparsity and signal-to-noise ratios • performance not dependent on the number of observations • does not allow inclusion of prior knowledge about sparsity • over- and underestimation of sparsity ratios possible • large coefficients left unshrunk • correlated predictors may result in multimodal posteriors • potential convergence and sampling issues	• continuous shrinkage prior with closed-form representation • efficient posterior computation • optimal posterior concentration • strong sensitivity to unknown concentration parameter
Regularized Horseshoe	
• continuous shrinkage prior • strong performance in high-dimensional settings • improved sampling stability • some sensitivity to hyperparameter choices	• generalization of the horseshoe that encompasses it • allows incorporation of prior knowledge about sparsity levels • correlated predictors may still lead to multimodal posteriors

This table summarizes the key properties of various prior distributions used in high-dimensional models.

3.2 Article Aims

To overcome some of these limitations, we suggest a novel prior approach, which we call Dirichlet-horseshoe, that combines and expands on the principles of the regularized horseshoe and the Dirichlet-Laplace priors. We test its performance by conducting two replicated simulation studies, along with a real-data example. In the first simulation study, which reflects the normal means problem (see Bhattacharya et al. 2015), Gaussian white noise gets added to a high-dimensional vector of signals with varying signal strengths and different sparsity ratios. Then the second simulation study features a multiple regression setting, such that we draw predictor variables from a multivariate normal distribution with varying correlation strengths and generate a sparse vector of regression coefficients from a log-normal distribution.

In the real-data example, we replicate the estimation procedure implemented by Fischer et al. (2023) but apply our proposed Dirichlet-horseshoe prior. Fischer et al. (2023) estimate the effects of emotional cues on minute-by-minute, aggregated alcohol consumption in a soccer stadium, on the basis of intensive longitudinal data and by employing a random intercept regime-switching model. The data set consists of 8,820 observations, with 10 predictors at the within-level and 49 predictors at the between-level.

These illustrative studies confirm that the Dirichlet-horseshoe can deal with different sparsity ratios, handle high multicollinearity, and cope well with small N, large D settings. In all setups, its performance, measured in terms of loss, is either superior or similar to that of competing regularization priors. The results suggest that our proposed prior is best suited for sparse, small sample settings (e.g., 100 observations with 5% relevant predictors) that require precise estimates and a high probability of isolating important predictors from less important ones.

3.3 Common Priors

To distinguish the Dirichlet-horseshoe prior from existing ones, we briefly introduce five well-established approaches, in chronological publication order. More precisely, we introduce three central concepts ("baseline priors"), which are prone to the issues we outlined previously. Then, we present two more recent ("advanced") priors, on which we build to construct the Dirichlet-horseshoe.

Throughout this article, we use N to denote the number of observations, with $i = 1, \ldots, N$, and D to indicate dimensionality, equal to the number of predictor variables in a regression, where $j = 1, \ldots, D$. The vector of regression coefficients to be estimated is signified by $\boldsymbol{\beta}$, and σ^2 represents the noise variance. Vectors and matrices are highlighted in bold.

3.3.1 Baseline Priors

3.3.1.1 Spike-and-Slab

The spike-and-slab prior (Mitchell and Beauchamp 1988; George and McCulloch 1993; Ishwaran and Rao 2005) belongs to the class of discrete mixture priors and is based on a binary decision. Either the point mass for β_j is at 0 (i.e., it comes from a spike), or the coefficient is drawn from a diffuse uniform distribution (i.e., it comes from a slab).

This prior is often written as a two-component mixture of Gaussian functions (Piironen and Vehtari 2017) such that

$$\begin{aligned} \beta_j \mid \lambda_j, c, \varepsilon \sim & \lambda_j \mathcal{N}\left(0, c^2\right) + (1 - \lambda_j)\, \mathcal{N}\left(0, \varepsilon^2\right) \ , \\ \lambda_j \sim & \text{Ber}\left(\pi\right) \ , \end{aligned}$$

in which the hyperparameter c controls the shape of the distribution (slab width), and π is the prior inclusion probability. This prior establishes an ideal theoretical benchmark, due to its notable properties. First, it fully reflects the concept of regularization, because coefficients can become exactly 0. Second, there exists a closed-form solution for the posterior distribution. By choosing $\varepsilon = 0$ and replacing the discrete Bernoulli distribution with a continuous Beta, the prior can be expressed as

$$\begin{aligned} \beta_j \mid \lambda_j, c \sim & \lambda_j \mathcal{N}\left(0, c^2\right) \ , \\ \lambda_j \sim & \text{Beta}\left(\alpha, \beta\right) \ . \end{aligned}$$

For example, Beta $(.5, .5)$ would be an uninformative Jeffrey's prior.

3.3.1.2 Bayesian (Adaptive) Lasso

Park and Casella (2008) introduce the Bayesian lasso,

$$\pi\left(\boldsymbol{\beta} \mid \sigma^2\right) = \prod_{j=1}^{D} \frac{\lambda}{2\sqrt{\sigma^2}} e^{-\lambda|\beta_j|/\sqrt{\sigma^2}} \ ,$$

for which the shape of the coefficient's distribution is governed by σ^2 and the fixed shrinkage parameter λ. According to Park and Casella (2008), estimates have both lasso and ridge regression properties, and the posterior mode is equivalent to the frequentist lasso estimator introduced by Tibshirani (1996). In an attempt to allow for different shrinkage for different coefficients, while maintaining parsimony, Leng, Tran, and Nott (2014) propose the Bayesian adaptive lasso,

$$\pi\left(\boldsymbol{\beta} \mid \sigma^2\right) = \prod_{j=1}^{D} \frac{\lambda_j}{2\sqrt{\sigma^2}} e^{-\lambda_j|\beta_j|/\sqrt{\sigma^2}} \ ,$$

which also ensures that the posterior is unimodal, given any choice of λ_j (Park and Casella 2008; Leng, Tran, and Nott 2014).

3.3.1.3 Horseshoe

Carvalho, Polson, and Scott (2010) propose the horseshoe,

$$\begin{aligned} \beta_j \mid \lambda_j, \tau &\sim \mathcal{N}\left(0, \lambda_j^2 \tau^2\right) \ , \\ \lambda_j &\sim \mathcal{C}^+\left(0, 1\right) \ , \end{aligned}$$

a prior distribution with a global (τ) and a local (λ_j) shrinkage parameter. The general degree of regularization τ (often referred to as the global regularization component) shrinks all coefficients toward 0; the local regularization component λ_j allows large coefficients to escape shrinkage. The horseshoe's heavy tails, resulting from the half-Chauchy distribution, allows the prior to accommodate even the largest coefficients.

3.3.2 Advanced Priors

3.3.2.1 Dirichlet-Laplace

Dirichlet-Laplace priors (Bhattacharya et al. 2015) are another class of shrinkage priors that extend the global-local approach by introducing relative predictor importance according to a concentration parameter a, as in

$$\begin{aligned} \beta_j \mid \phi_j, \tau &\sim \text{DE}\left(0, \phi_j \tau\right) \ , \\ \boldsymbol{\phi} &\sim \text{Dir}\left(a, \ldots, a\right) \ , \\ \tau &\sim \text{Gamma}\left(Da, 0.5\right) \ . \end{aligned}$$

The components of the Dirichlet distribution ϕ_j add up to 1 (i.e., the Dirichlet is a multivariate generalization of the Beta distribution). In these equations, DE refers to the double exponential (Laplace) distribution. The concentration parameter a guides the sparsity allocation of the probability mass, such that the larger a is, the more mass gets concentrated on one component, and the smaller a is, the sparser the resulting distribution.

3.3.2.2 *Regularized Horseshoe*

The regularized horseshoe, as proposed by Piironen and Vehtari (2017), is a generalization of the horseshoe, expressed as:

$$\begin{aligned} \beta_j \mid \lambda_j, \tau, c &\sim \mathcal{N}\left(0, \tau^2 \tilde{\lambda}_j^2\right) \ , \\ \tilde{\lambda}_j^2 &= \frac{c^2 \lambda_j^2}{c^2 + \tau^2 \lambda_j^2} \ , \\ \lambda_j &\sim \mathcal{C}^+ (0, 1) \ . \end{aligned} \tag{3.1}$$

If $\tau^2 \lambda_j^2 \ll c^2$, such that β_j is close to 0, then $\tilde{\lambda}_j^2 \underset{p}{\to} \lambda_j^2$, and Equation (3.1) approaches the original horseshoe. But if $\tau^2 \lambda_j^2 \gg c^2$, such that β_j is far from 0, then $\tilde{\lambda}_j^2 \underset{p}{\to} \frac{c^2}{\tau^2}$, and Equation (3.1) approaches $\mathcal{N}(0, c^2)$, which is equivalent to the spike-and-slab estimator for the case in which the coefficient comes from a slab.

The regularized version enhances the horseshoe in several ways. First, the global shrinkage parameter τ depends on prior information about the sparsity ratio in the design matrix, because

$$\begin{aligned} \tau &\sim \mathcal{C}^+ \left(0, \tau_0^2\right) \ , \\ \tau_0 &= \frac{p_0}{D - p_0} \frac{\sigma}{\sqrt{N}} \ , \end{aligned}$$

where p_0 denotes the number of relevant predictor variables. On the basis of the dimensionality of the data set and an assumed number of relevant predictors, the global shrinkage parameter can be defined systematically. Second, sparsity information enters the reformulated local shrinkage parameter $\tilde{\lambda}_j$ through τ, which prevents the vanishing mean problem that arises with the horseshoe, as well as the aforementioned issues attributed to weakly identified parameters.

3.4 The Dirichlet-Horseshoe Prior

To enhance precision and offer more adaptivity, we combine the Dirichlet-Laplace with the regularized horseshoe, such that we integrate the notion of relative levels of predictor importance and local shrinkage, using information on the overall sparsity. The implied hierarchical structure of

the Dirichlet-horseshoe then can be expressed as:

$$
\begin{aligned}
\beta_j \mid \lambda_j, \phi_j, \tau, c &\sim \mathcal{N}\left(0, \tau^2 \tilde{\lambda}_j^2 \tilde{\phi}_j^2\right) \ , \\
\tau &\sim \mathcal{C}^+\left(0, \tau_0^2\right) \ , \\
\tilde{\lambda}_j^2 &= \frac{c^2 \lambda_j^2}{c^2 + \tau^2 \lambda_j^2} \ , \\
\tilde{\boldsymbol{\phi}} &= \boldsymbol{\phi} D \ , \\
\lambda_j &\sim \mathcal{C}^+(0, \gamma) \ , \\
\boldsymbol{\phi} &\sim \text{Dir}(a, \ldots, a) \ .
\end{aligned}
$$

The proposed Dirichlet-horseshoe unifies global (τ), local (λ_j), and joint ($\boldsymbol{\phi}$) regularization. The global and local regularization trade off in $\tilde{\lambda}_j$, with the objective of also shrinking large coefficients when they are weakly identified, to prevent flat posteriors.

The peculiarity of the Dirichlet distribution maps all components in $\boldsymbol{\phi}$ between 0 and 1; it guarantees that the sum of these components is equal to 1. Rescaling the components by D allows the Dirichlet-horseshoe to nest the regularized horseshoe estimator as a special case, as long as the posterior for $\tilde{\boldsymbol{\phi}} = \mathbf{1}$. In practice, this special case is unlikely; it implies equally important predictor variables. However, the Dirichlet-horseshoe might increasingly come to resemble the regularized horseshoe when the predictor variables grow more homogeneous in their explanatory power for the target variable.

For $\tilde{\boldsymbol{\phi}} \neq \mathbf{1}$, the Dirichlet-horseshoe shrinks all predictors that are less important than an average to be stronger than the regularized horseshoe does, such that for $\tilde{\phi}_j < 1$, it holds that $\tau^2 \tilde{\lambda}_j^2 \tilde{\phi}_j^2 < \tau^2 \tilde{\lambda}_j^2$. Analogously, any predictor of greater importance than the average is subjected to a smaller degree of shrinkage. In turn, the Dirichlet-horseshoe functions as a regularizer with greater selectivity capacity between important and unimportant signals.

The slab width can be adjusted by means of c. In line with Piironen and Vehtari (2017), we suggest establishing a weakly informative prior for c. A reasonable choice is the inverse-Gamma distribution,

$$
c^2 \sim \text{InvGamma}(\alpha, \beta) \ ,
$$

which has a heavy right and light left tail, so it does not concentrate excessive probability mass near 0, a feature that is desirable for coefficients that already are considered far from 0 (Piironen and Vehtari 2017). Incorporating global parameters in the model, estimated from available data, also provides a high degree of adaptivity to different sparsity patterns and can control for Type I errors (Berry 1987; Scott and Berger 2006).

The local regularization component λ_j, drawn from a half-Cauchy distribution, enables our prior to accommodate even the largest coefficients. The component's impact on the estimator can be controlled by the scale parameter γ. For the optimal choice of τ_0, we follow Piironen and

Vehtari (2017) and use:

$$\tau_0 = \frac{p_0}{D - p_0} \frac{\sigma}{\sqrt{N}} .$$

Then we define hyperpriors for both the concentration parameter a and the number of relevant predictors p_0 to reduce the dependence of the results on these parameters when little or no prior information is available. To induce sparsity through ϕ, the concentration rates must lie between 0 and 1. To increase the degree of sparsity, modelers should choose smaller values. The number of relevant predictors p_0 controls the scale τ_0 of the half-Cauchy distribution from which the global hyperparameter τ is drawn; p_0 is inversely related to the degree of global regularization. In practice, p_0 is often unknown and needs to be sampled.

For statistical reasons, we recommend drawing the fraction of relevant predictors η rather than the number of relevant predictors p_0, according to

$$p_0 = \eta D ,$$

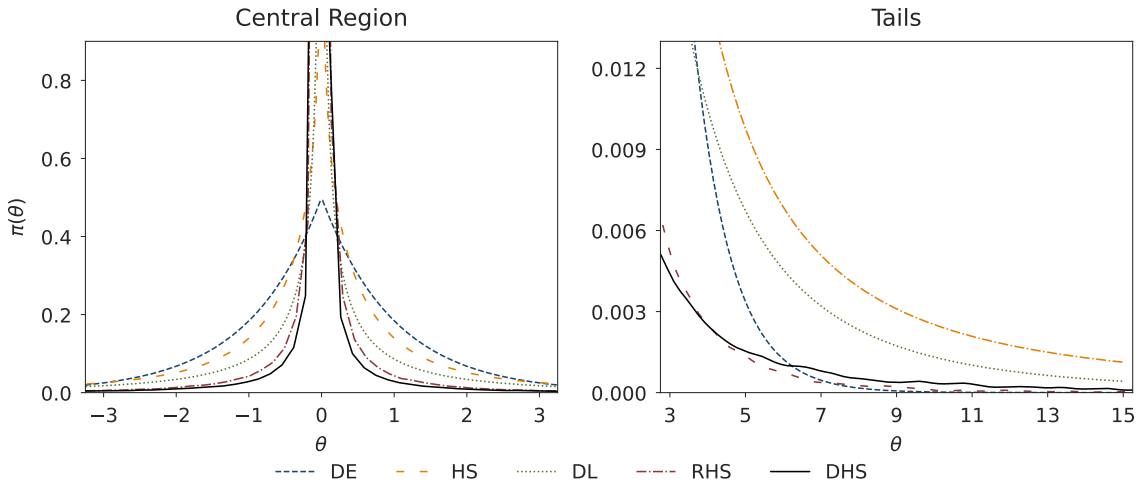
because η is a continuous quantity bounded between 0 and 1, which makes it more convenient to sample using numerical methods, such as Markov chain Monte Carlo or Hamiltonian Monte Carlo.

On the basis of empirical experience, Bhattacharya et al. (2015) determine that choosing $a = 1/D$ may induce over-shrinkage if some relatively small coefficients exist. They also emphasize that $a = 1/D$ can create numerical issues when D is large. In addition to fixing $a = 0.5$, they use a uniform prior for the concentration rate a, but uniform distributions can lead to slow convergence and high autocorrelation in the chain.

Therefore, we recommend drawing a and η from informative Beta distributions, skewed toward 0 (relative to the presumed sparsity ratio), to allocate more prior probability near 0 than near 1. This effort can help the sampler explore the posterior distribution more efficiently. For example, Beta$(2, 8)$ has a mean equal to 0.2, which may be appropriate if the researcher assumes around 20% of the predictors are relevant, a priori. According to our findings, and in an effort to restrain computational expense, it may be reasonable to define one joint parameter ν for both the concentration a and the fraction of relevant predictors η, such that $\eta = a = \nu$, with $\nu \sim$ Beta(α, β).

Figure 3.1 shows the marginal density of the Dirichlet-horseshoe compared with the other priors. We use the double exponential prior in place of the Bayesian adaptive lasso, because in a Bayesian lasso regression, each regression coefficient follows a conditional double exponential prior (i.e., Laplace prior) (Tibshirani 1996; Park and Casella 2008). We observe that the Dirichlet-horseshoe allocates substantial probability mass close to 0 and has a relatively light tail. The tails are large enough to accommodate large coefficients though, and they are more pronounced than those of the double exponential prior. These key features enable the Dirichlet-horseshoe to perform well in relation to the problems of sparse Bayesian regularization and prediction (Polson

Figure 3.1
Marginal Densities



Marginal density of the Dirichlet-horseshoe (DHS) in comparison with those of the double exponential (DE), horseshoe (HS), Dirichlet-Laplace (DL), and regularized horseshoe (RHS) priors.

and Scott 2010).

The density functions provided by horseshoe, regularized horseshoe, and Dirichlet-horseshoe all lack a closed-form representation. We approximate the horseshoe density using the elementary functions, as detailed in Theorem 1 offered by Carvalho, Polson, and Scott (2010). The density functions for the latter two priors are sampled using 10,000,000 draws. We assume $N = 100$, $D = 100$, $a = 0.5$, $\eta = 0.5$, and $\sigma = 1$, and we smooth the tails of the regularized and Dirichlet horseshoes by scaling the kernel bandwidth, for illustration.

3.5 Simulation Studies

We illustrate the finite sample performance of the proposed Dirichlet-horseshoe prior, compared with the other priors, using the normal means problem (Van Der Pas, Kleijn, and Van Der Vaart 2014; Bhattacharya et al. 2015), and then a multiple regression problem with correlated predictors under sparsity.

3.5.1 Simulation Study I: Normal Means Problem

In the normal means problem, we strive to estimate a D-dimensional mean $\boldsymbol{\theta_0}$ based on a D-dimensional vector of observations $\boldsymbol{y}$ corrupted with Gaussian white noise,

$$y_j = \theta_j + \varepsilon_j \ ,$$
$$\varepsilon_j \sim \mathcal{N}(0, 1) \ ,$$

for various sparsity ratios q_D and mean sizes defined by the signal strength A. Particularly, we sample $\Omega = 100$ replicates of a $D = 400$ dimensional vector $\mathbf{y}$ from a $\mathrm{N}_D(\boldsymbol{\theta_0}, \boldsymbol{I_D})$ distribution, where $\boldsymbol{\theta_0}$ has $q_D \in \{0.05, 0.1\}$ nonzero entries, all set to a constant value $A \in \{4, 5, 6\}$. In turn, we conduct six simulation settings per prior candidate in each replication, for 3,600 simulations in total. Note that $A = \sqrt{2\ln(400)} \approx 3.5$ is the "universal" threshold for this problem (Johnstone and Silverman 2004; Van Der Pas, Kleijn, and Van Der Vaart 2014). Below this threshold, the nonzero components in $\boldsymbol{\theta_0}$ are too small to be detected (Piironen and Vehtari 2017). Therefore, we expect the most imprecise estimates to manifest for $A = 4$.

We also seek to illustrate the impact of the hyperparameter choice ν, which controls both the concentration rate and the fraction of relevant parameters in the Dirichlet-horseshoe prior, by presenting the results for different hyperparameter selections in the simulation study: $D = 400$, $q_D = 0.05$, and A varying over $A \in \{4, 5, 6\}$. The hyperprior choices include $\nu \sim \text{Beta}(2, 8)$, $\nu = 0.01$, $\nu = 0.1$, and $\nu = 0.5$.

3.5.2 Simulation Study II: Multiple Regression Problem

In the multiple regression problem, we aim to estimate a D-dimensional parameter vector $\boldsymbol{\beta}$ drawn from a log-normal distribution based on a $N \times D$ dimensional design matrix $\mathbf{X}$, which in turn is drawn from a multivariate normal distribution, such that

$$
\begin{aligned}
y_i &= \beta_0 + \beta_j x_{ij} + \varepsilon_i \ , \\
\beta_0 &\sim \mathcal{N}(0, 1) \ , \\
\beta_j &\sim \mathcal{LN}(0, 0.5) \ , \\
\boldsymbol{X} &\sim \mathcal{N}_D(\mathbf{0}, \boldsymbol{\Sigma}) \ , \\
\varepsilon_i &\sim \mathcal{N}(0, 1) \ .
\end{aligned}
$$

The variance-covariance matrix $\boldsymbol{\Sigma}$ contains unit standard deviation on its main diagonal and varying, pairwise correlations ρ on its off-diagonals. We set the number of predictors $D = 100$, let the sample size be $N \in \{80, 200\}$ ($N < D$, $N > D$), and vary the correlations ρ over $\rho \in \{0.25, 0.5, 0.75\}$ (weak, moderate, strong). The parameter vector $\boldsymbol{\beta}$ has a sparsity ratio of 20%, such that we set 80 of the 100 predictors to 0. Again, we conduct $\Omega = 100$ replications per setting and prior candidate, which yields a total of 3,600 simulations.

3.5.3 Estimation

For the estimation, we use the probabilistic programming package PyMC (Salvatier, Wiecki, and Fonnesbeck 2016). In particular, we apply the No-U-Turn (NUTS) sampler (Hoffman and Gelman 2014), a recursive algorithm for continuous variables based on Hamiltonian mechanics. It extends the Hamiltonian Monte Carlo by eliminating the need to set a number of steps, because an automatic stoppage occurs once the sampler starts to make a U-turn. We set a large

acceptance probability of 0.99 (the default value is 0.8), which implies a small step size, so that we can achieve non-divergent trajectories for the samples. For each specification, we run three chains with 3,000 iterations and an additional 1,000 burn-in samples that we discard. We monitor convergence with the Gelman-Rubin statistic $\hat{R}$ (Gelman and Rubin 1992; Brooks and Gelman 1998).

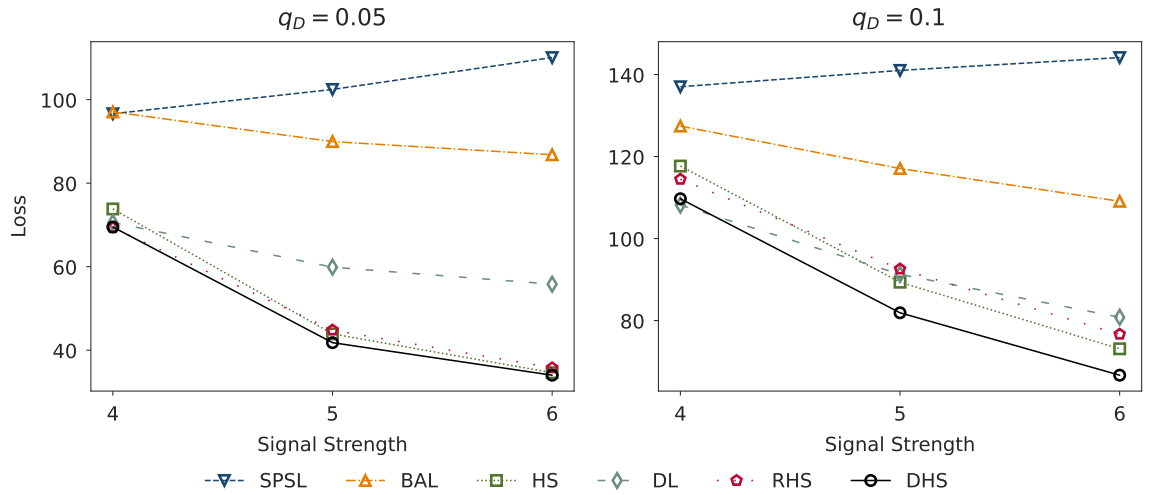
3.6 Results

This section presents the results of both simulation studies. We illustrate the performance of the Dirichlet-horseshoe compared with the alternative priors. Figures 3.2 and 3.4 depict the squared error loss corresponding to the posterior median. Tables B.1 to B.4 in the appendix present these results in tabular form, along with several additional measures, such as coverage based on the 95% highest density intervals (HDIs), standard deviation of loss, standard deviation of the median, Type I error, $\hat{R}$, runtime, and number of divergences. These measures are averaged across replications, and all values converge. We present the detailed hierarchical structure of all priors, together with the parameter values used in the simulation studies, in Section B.2.

3.6.1 Simulation Study I: Normal Means Problem

Figure 3.2 depicts the loss for varying signal strengths $A \in \{4, 5, 6\}$ based on sparsity ratios of $q_D = 5\%$ (left panel) and $q_D = 10\%$ (right panel) for the normal means problem with dimensionality $D = 400$. Specifically, for a sparsity ratio of 5% and a signal strength of 4, the lowest

Figure 3.2
Results of Simulation Study I



Squared error loss corresponding to the posterior median derived from the normal means problem (Simulation Study I), with a $D = 400$ dimensional parameter vector for varying signal strengths $A \in \{4, 5, 6\}$ and sparsity ratios $q_D = 5\%$ (left panel) and $q_D = 10\%$ (right panel). We abbreviate the considered priors SPSL (spike-and-slab), BAL (Bayesian adaptive lasso), HS (horseshoe), DL (Dirichlet-Laplace), RHS (regularized horseshoe), and DHS (Dirichlet-horseshoe).

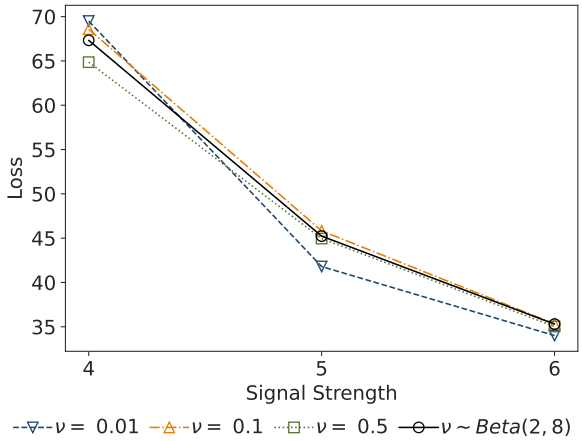
loss occurs for the regularized horseshoe (69.3), followed closely by Dirichlet-horseshoe (69.5), Dirichlet-Laplace (70.6), and horseshoe (73.8) priors. Bayesian adaptive lasso and spike-and-slab priors reveal the most imprecise results (97.1 and 96.6), which persists at larger signal strengths too. As the signals strengthen, the losses associated with all priors except spike-and-slab decrease, as is to be expected, because the signal-to-noise ratio grows.

For $A = 5$ and $A = 6$, the Dirichlet-horseshoe achieves the lowest loss (41.8 and 34.0), followed by the horseshoe (43.9 and 34.6) and regularized horseshoe (44.8 and 35.8). The Dirichlet-Laplace prior performance deteriorates (59.9 and 55.8). At a sparsity ratio of 10%, we find similar patterns. For $A = 4$, the loss for the Dirichlet-horseshoe (110.0) is slightly higher than that for the Dirichlet-Laplace (108.0) but considerably lower than those of the regularized horseshoe (114.0), horseshoe (118.0), Bayesian adaptive lasso (127.0), and spike-and-slab (137.0).

For the signal strengths $A = 5$ and $A = 6$, the loss of the Dirichlet-horseshoe (81.9 and 66.7) is lowest, and the differences with the other priors grow more pronounced. That is, the horseshoe (89.4 and 73.1) comes second, closely followed by the Dirichlet-Laplace (91.3 and 80.8) and regularized horseshoe (92.6 and 76.7). Overall, the proposed, novel prior consistently demonstrates superior or competitive accuracy on all tested specifications.

Figure 3.3 illustrates the sensitivity of the Dirichlet-horseshoe to a selection of hyperparameters that constitute the proposed hyperprior $\nu \sim \text{Beta}(2, 8)$, together with fixed values of $\nu = 0.01$, $\nu = 0.1$, and $\nu = 0.5$. We evaluate the results on the basis of the loss calculated

Figure 3.3
Results of Hyperparameter Sensitivity Analysis



Squared error loss corresponding to the posterior median for different hyperparameter choices ν, derived from the normal means problem with a $D = 400$ dimensional parameter vector, sparsity ratio $q_D = 5\%$, and varying signal strengths $A \in \{4, 5, 6\}$.

from the normal means problem with dimensionality $D = 400$, sparsity ratio $q_D = 5\%$, and varying signal strengths $A \in \{4, 5, 6\}$.

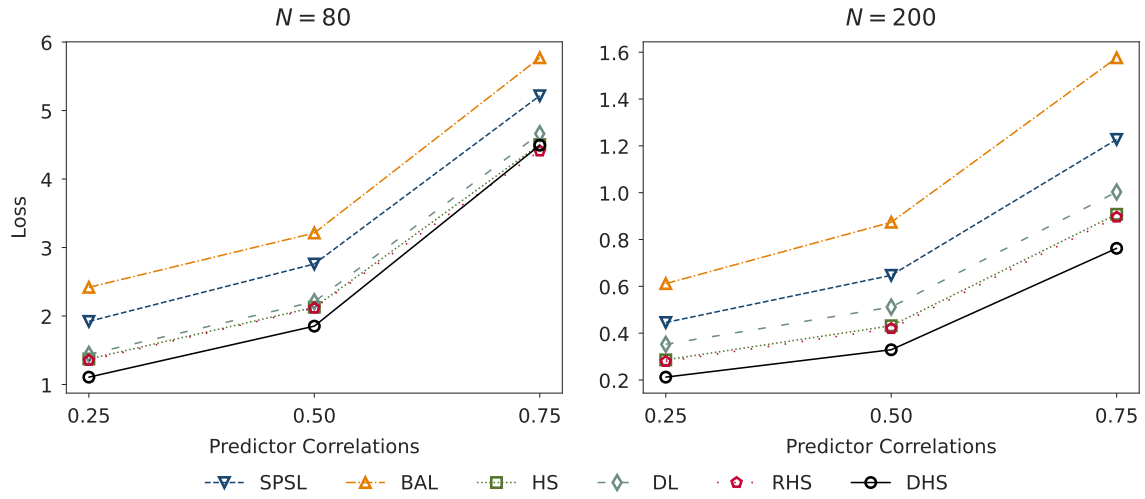
For the smallest signal strength $A = 4$, we observe minor differences in losses across the hyperprior choices $\nu \sim \text{Beta}(2, 8)$, $\nu = 0.1$, and $\nu = 0.5$, but the largest loss emerges for $\nu = 0.01$.

As Bhattacharya et al. (2015) point out, choosing overly small values for the concentration parameter can lead to overshrinkage and poor performance in the presence of relatively small signals. For $A = 5$ and $A = 6$, we note the weak sensitivity of the Dirichlet-horseshoe to the hyperprior choice. The differences in loss are rather small, indicating that the choice of hyperparameter has only a minor impact in these settings.

3.6.2 Simulation Study II: Multiple Regression Problem

Figure 3.4 depicts the loss estimated from the multiple regression problem with $D = 100$ predictors for varying correlation strengths $\rho \in \{0.25, 0.5, 0.75\}$ (weak, moderate, strong) in samples with $N = 80$ observations (left panel) and $N = 200$ observations (right panel). For the small sample size of 80 and weak to moderate predictor correlations, the Dirichlet-horseshoe achieves the lowest loss (1.11 and 1.85), followed by the regularized horseshoe (1.35 and 2.12), horseshoe (1.37 and 2.13), and Dirichlet-Laplace (1.44 and 2.21).

Figure 3.4
Results of Simulation Study II



Squared error loss corresponding to the posterior median derived from the multiple regression setting (Simulation Study II) with a $D = 100$ dimensional parameter vector drawn from a log-normal distribution for varying correlations $\rho \in \{0.25, 0.5, 0.75\}$ and observations $N = 80$ (left panel) and $N = 200$ (right panel). The design matrix is generated by a multivariate normal distribution with correlated predictors. The parameter vector $\boldsymbol{\beta}$ has a sparsity ratio of 20%, such that we set 80 of the 100 predictors to 0. We abbreviate the priors SPSL (spike-and-slab), BAL (Bayesian adaptive lasso), HS (horseshoe), DL (Dirichlet-Laplace), RHS (regularized horseshoe), and DHS (Dirichlet-horseshoe).

Similar to the normal means problem in Simulation Study I, the Bayesian adaptive lasso and spike-and-slab priors produce the most imprecise results, as is true for all the predictor correlations we examine. With strong correlations, the Dirichlet-horseshoe reveals a loss of 4.49—slightly larger than the values for the regularized horseshoe (4.41) and slightly lower than the horseshoe (4.50), and Dirichlet-Laplace (4.66). With a larger sample size (200), our proposed Dirichlet-horseshoe prior exhibits the smallest loss for all predictor correlations

$\rho \in \{0.25, 0.5, 0.75\}$, with values equal to 0.21, 0.33, and 0.76, respectively. Thus, it also achieves the best outcome for strong correlations, followed by regularized horseshoe (0.28, 0.42, 0.90), horseshoe (0.29, 0.43, 0.91), and Dirichlet-Laplace (0.35, 0.51, 1.00).

Generally, the loss of all priors diminishes with the larger sample size but grows with increasing predictor correlations, which is reasonable, in that it becomes more challenging for the algorithms to disentangle predictor variables. As in the normal means problem, our proposed Dirichlet-horseshoe prior achieves superior or highly competitive accuracy for all the tested specifications.

3.7 Real-Data Example: Alcohol Use and Emotional Cues

In this section, we extend the performance illustration with a real-data example. Similar to Fischer et al. (2023), we estimate the effect of the reference point-dependent emotional cues *Surprise* and *Suspense* (Ely, Frankel, and Kamenica 2015), with the same data set and using their proposed hierarchical state-switching model. However, we apply the Dirichlet-horseshoe instead of the spike-and-slab prior used by those authors. The data set consists of minute-by-minute recorded beer sales in a soccer stadium, as well as data related to in-play betting odds, in-play match events, and match-day information. The design matrix contains 10 predictors at the within-level and 49 predictors at the between-level. All data are observed over six soccer seasons, from 2013/14 to 2018/19, which entail 98 matches, leading to $98 \times 90 = 8,820$ observations in total (a match lasts 90 minutes). Away-team stands are excluded from this sample.

Fischer et al. (2023) follow Buraimo et al. (2020) and use betting odds to estimate outcome probabilities p for all potential outcomes of the match, namely, home win (H), draw (D), or away win (A), in each match minute t. Using these estimated probabilities, they calculate the constructs *Surprise* and *Suspense*. Fischer et al. (2023) also determine (emotional) states $S = 0$ and $S = 1$ on basis of these estimates, such that *Surprise*, calculated as,

$$\textit{Surprise}_t = u\left(\sum_{m \in H,D,A} \left[p_t^m - p_{t-1}^m\right]^2\right) ,$$

is high if the current outcome probabilities contradict the anterior beliefs (backward-looking). In contrast, *Suspense*,

$$\textit{Suspense}_t = u\left[\sum_{m \in H,D,A} p_{t+1}^{HG}\left\{\left(p_{t+1}^m \mid p_{t+1}^{HG}\right) - p_t^m\right\}^2 + p_{t+1}^{AG}\left\{\left(p_{t+1}^m \mid p_{t+1}^{AG}\right) - p_t^m\right\}^2\right] ,$$

is a forward-looking measure that increases when the variance related to possible outcomes in the next minute is high. Here, p_{t+1}^{HG} is the probability that the home team scores in the next minute. The superscript "AG" instead refers to the away team's probability of scoring. The utility function is $u(\cdot) = \sqrt{\cdot}$ for both measures.

The two emotional states then are given as

$$\begin{aligned} p_t^H < p_{t-1}^H \quad &\Rightarrow \quad S = 0 \ , \text{ and} \\ p_t^H > p_{t-1}^H \quad &\Rightarrow \quad S = 1 \ , \end{aligned}$$

where $S = 0$ denotes the negative state, because the outcome probabilities for the home team have decreased from the last period to the current period. The positive state is indicated by $S = 1$. These states are well-justified, because consumption by home and away team fans is separated by stands. This model also allows for a random intercept at the between-level.

The within-level is given by,

$$\begin{aligned} [\ln(Y_{1it}) \mid S_{it} = s] = \nu_{1i} &+ \sum_{\text{l=1}}^{L} \ln\left(Y_{1i(t-l)}\right) \phi_{1sl} \\ &+ \sum_{l=0}^{L} \left(\boldsymbol{X_{1i(t-l)}\beta_{1sl}} + \boldsymbol{Z_{1i(t-l)}\gamma_{1sl}}\right) \\ &+ \psi_{1ts} + \varepsilon_{1it} \ , \end{aligned}$$

where,

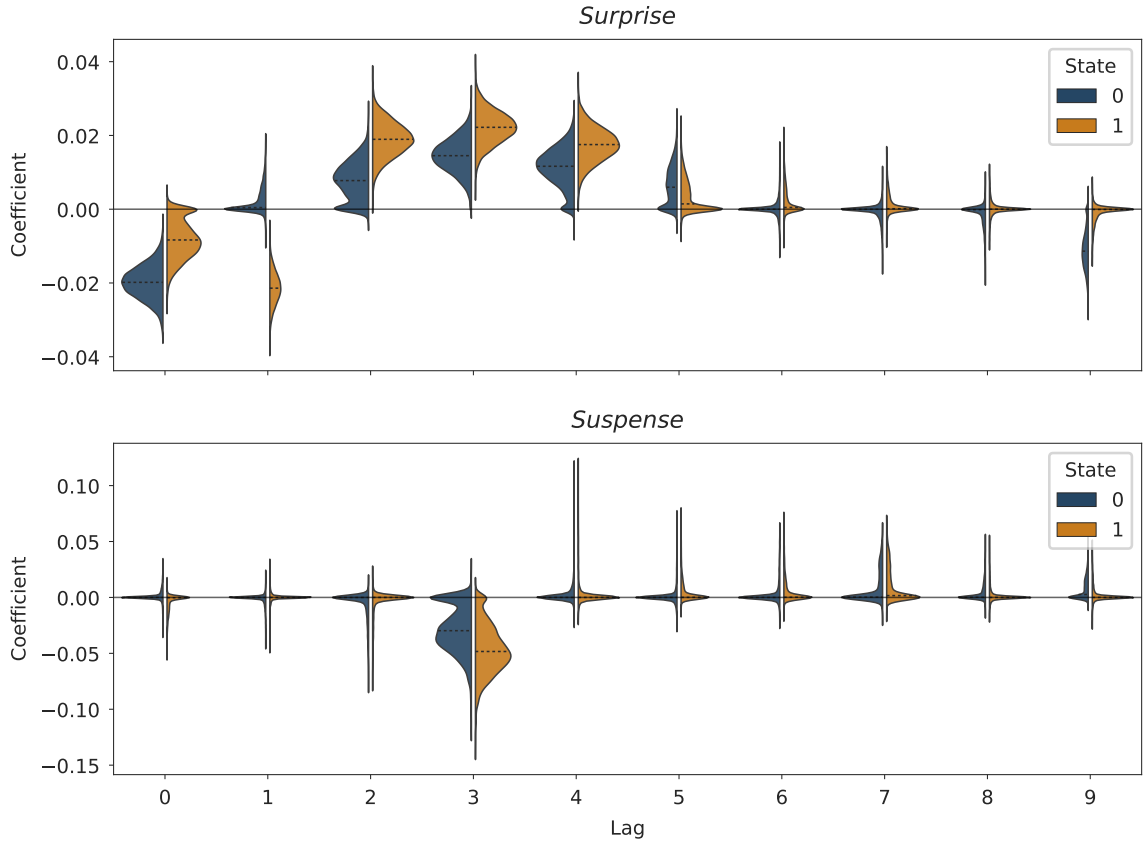
$$\nu_{1i} = \boldsymbol{Z_{2i}\gamma_2} + \upsilon_{2i} \ ,$$

at the between-level. Beer sales are collected in $\boldsymbol{Y}$, and $\boldsymbol{X_{1i(t-l)}}$ contains the key regressors, *Surprise* and *Suspense*. Covariates in $\boldsymbol{Z_{1i(t-l)}}$, capturing weather conditions and match key events, among other variables, control for potential confounds at the within-level. Covariates related to match-specific information, like the number of spectators, beer price, weekday, and so on, and gathered by $\boldsymbol{Z_{2i}}$, control for confounds at the between-level. We use the estimation procedure described for Simulation Studies I and II.

Figure 3.5 presents the results based on Dirichlet-horseshoe. The regressors are standardized, and the dependent variable contains inverse hyperbolic sine-transformed sales. Therefore, the marginal effects can be interpreted in terms of a one standard deviation increase. Our results resonate with the main findings of Fischer et al. (2023). We find predominantly positive effects for *Surprise* and negative effects for *Suspense* on the number of beer sales in both states. Moreover, the effects are stronger for the positive state $S = 1$ overall.

We observe a greater level of decisiveness (sharpness) for the Dirichlet-horseshoe compared with the spike-and-slab prior when it comes to variable selection. Accordingly, we observe a higher density allocated around 0 for small-coefficient posteriors exposed at lags 6-8, for example, as well as slightly weaker effects for *Suspense* when $l = 3$. For some nonzero posteriors, the probability mass locally peaks close to 0 though. The resulting tendency for multimodal posteriors is typical of horseshoe priors, due to the global and local shrinkage aspects.

Figure 3.5
Results of Real-Data Example



Posterior distributions of the effects of the emotional cues *Surprise* and *Suspense* on beer sales, derived from the real-data example. The figure shows posteriors for all lags in both states (negative $S = 0$ and positive $S = 1$). The dotted lines indicate the estimated median effect.

3.8 Discussion

We present the Dirichlet-horseshoe, a novel prior designed for sparse estimations in high-dimensional settings, which integrates the concept of Dirichlet-Laplace priors with the regularized horseshoe. In two simulation studies and a real-data example, we demonstrate that the proposed prior yields more accurate estimates than both well-established and recently developed methods, namely, the spike-and-slab prior, Bayesian adaptive lasso, horseshoe, Dirichlet-Laplace, and regularized horseshoe. Our approach offers a reliable default choice for addressing problems related to varied levels of sparsity (i.e., identifying needles and straws in haystacks), and it exhibits favorable convergence properties.

Considering the normal means problem with sparsity ratios of 5% and 10%, our proposed prior achieves significantly better performance than the spike-and-slab and Bayesian adaptive lasso; it performs similarly to the horseshoe, Dirichlet-Laplace, and regularized horseshoe priors at a signal strength of 4. At higher signal strengths (5 or 6), the proposed Dirichlet-horseshoe exhibits a slight advantage over all these priors, showcasing its superior performance.

The Dirichlet-horseshoe also is well suited for small N, large D problems, and it can deal with correlated predictor variables. Specifically, for weak to moderate correlations (i.e., 0.25 to 0.5), the Dirichlet-horseshoe provides more accurate estimates than competing priors. Furthermore, even in the presence of strong correlations, the prior continues to be more accurate, as long as the sample sizes are sufficiently large.

Our novel prior further offers a significant degree of flexibility, allowing for easy customization to accommodate specific settings, as demonstrated in the real-data example. Therefore, we consider the Dirichlet-horseshoe prior relevant for a wide range of application fields. It also can address various common challenges in modern statistics, including high-parametric, complex settings, such as network models (Epskamp, Rhemtulla, and Borsboom 2017) or dynamic latent variable models (Kelava and Brandt 2019), as well as common regression and classification tasks, function estimation, and covariance matrix regularization.

Unless substantial knowledge about the hyperpriors of the Dirichlet-horseshoe exists, we advocate for using the parameters in Appendix B.2.6 as a good starting point. Furthermore, we emphasize that this novel approach extends beyond this particular option; it may be possible to test alternative hyperpriors as well. We use full Bayesian inference for the hyperpriors since opting for empirical Bayes estimation of hyperparameters (as highlighted in the advantages of the empirical Bayes method by Petrone, Rousseau, and Scricciolo (2014)) or using cross-validation would pose computational difficulties and would not allow accounting for the characteristics of the posterior distributions.

When using numerical methods for posterior estimation, such as Markov chain Monte Carlo, Hamiltonian Monte Carlo, or NUTS, we recommend a noncentered representation of the densities, combined with small step sizes, to achieve fewer divergences in the trajectories for the samples. Continued research might investigate reducing the computational expense required for Dirichlet-horseshoe, without sacrificing its performance. Tests of the prior in various applications also are required to confirm its robustness and further specify both its strengths and its weaknesses.

B Appendix

B.1 Further Results

Table B.1
Results (I) of Simulation Study I

		$q_D = 0.05$				$q_D = 0.1$			
		Loss	StdLoss	StdMedian	Coverage	Loss	StdLoss	StdMedian	Coverage
$A = 4$	SPSL	96.6	0.1431	0.6510	0.9812	137.0	0.1348	0.7284	0.9747
	BAL	97.1	0.1689	0.7146	0.9923	127.0	0.1496	0.7334	0.9857
	HS	73.8	0.2784	0.4539	0.9892	118.0	0.1957	0.5862	0.9815
	DL	70.6	0.2244	0.6309	0.9919	108.0	0.1840	0.6546	0.9845
	RHS	69.3	0.2673	0.5014	0.9889	114.0	0.1822	0.6330	0.9814
	DHS	69.5	0.2939	0.4494	0.9900	110.0	0.2057	0.5875	0.9826
$A = 5$	SPSL	102.0	0.1411	0.6951	0.9852	141.0	0.1273	0.7587	0.9859
	BAL	90.0	0.1739	0.7137	0.9955	117.0	0.1581	0.7310	0.9913
	HS	43.9	0.3480	0.4728	0.9935	89.4	0.2312	0.5997	0.9889
	DL	59.9	0.2387	0.6292	0.9956	91.3	0.2032	0.6510	0.9909
	RHS	44.8	0.3369	0.5020	0.9931	92.6	0.2126	0.6310	0.9905
	DHS	41.8	0.3582	0.4535	0.9938	81.9	0.2471	0.5772	0.9893
$A = 6$	SPSL	110.0	0.1436	0.7243	0.9878	144.0	0.1187	0.7788	0.9871
	BAL	86.8	0.1711	0.7128	0.9957	109.0	0.1431	0.7295	0.9922
	HS	34.6	0.3198	0.4743	0.9971	73.1	0.1925	0.6025	0.9927
	DL	55.8	0.2292	0.6277	0.9962	80.8	0.1732	0.6490	0.9922
	RHS	35.8	0.3181	0.4934	0.9968	76.7	0.1892	0.6239	0.9925
	DHS	34.0	0.3303	0.4470	0.9971	66.7	0.2099	0.5653	0.9931

This table presents performance metrics based on 95% HDIs, averaged across simulation replicates. The measures are derived from the normal means problem (Simulation Study I) with a $D = 400$ dimensional parameter vector. We perform $\Omega = 100$ replications for all combinations of signals strengths $A \in \{4, 5, 6\}$ and sparsity ratios $q_D \in \{0.05, 0.1\}$. We consider the spike-and-slab (SPSL), Bayesian adaptive lasso (BAL), horseshoe (HS), Dirichlet-Laplace (DL), regularized horseshoe (RHS), and Dirichlet-horseshoe (DHS) priors.

Table B.2
Results (II) of Simulation Study I

		$q_D = 0.05$				$q_D = 0.1$			
		TypeIErr	Rhat	Minutes	Divergences	TypeIErr	Rhat	Minutes	Divergences
$A = 4$	SPSL	0.0004	1.00	0.4216	0	0.0012	1.00	0.4864	0
	BAL	0.0012	1.00	2.00	21	0.0015	1.00	1.90	31
	HS	0.0001	1.00	0.9389	118	0.0002	1.00	1.11	105
	DL	0.0005	1.00	2.54	0	0.0005	1.00	2.42	0
	RHS	0.0001	1.00	0.4583	0	0.0005	1.00	0.4244	0
	DHS	0.0001	1.00	4.68	7.49	0.0003	1.00	4.29	3.38
$A = 5$	SPSL	0.0009	1.00	0.5204	0	0.0017	1.00	0.4934	0
	BAL	0.0014	1.00	1.97	25.4	0.0012	1.00	1.88	30.9
	HS	0.0001	1.00	0.9861	131	0.0003	1.00	1.34	95
	DL	0.0006	1.00	2.40	0	0.0006	1.00	2.42	0
	RHS	0.0001	1.00	0.4315	0.0400	0.0005	1.00	0.4438	0
	DHS	0.0002	1.01	4.47	4.60	0.0003	1.00	4.44	1.83
$A = 6$	SPSL	0.0014	1.00	0.4924	0	0.0024	1.00	0.5030	0
	BAL	0.0014	1.00	1.89	23	0.0013	1.00	1.88	31.8
	HS	0.0001	1.00	1.09	116	0.0003	1.00	1.36	109
	DL	0.0005	1.00	2.42	0	0.0005	1.00	2.41	0
	RHS	0.0002	1.00	0.4359	0	0.0005	1.00	0.4540	0
	DHS	0.0001	1.00	4.45	1.40	0.0003	1.00	4.47	1.25

See the descriptive notes to Table B.1.

Table B.3
Results (I) of Simulation Study II

		$N = 80$				$N = 200$			
		Loss	StdLoss	StdMedian	Coverage	Loss	StdLoss	StdMedian	Coverage
$\rho = 0.25$	SPSL	1.92	0.4223	0.2256	0.9948	0.4456	0.2645	0.0869	0.9885
	BAL	2.42	0.4099	0.2744	0.9972	0.6116	0.2233	0.0946	0.9841
	HS	1.37	0.4449	0.1579	0.9878	0.2860	0.3186	0.0721	0.9884
	DL	1.44	0.4377	0.1917	0.9939	0.3525	0.2883	0.0804	0.9898
	RHS	1.35	0.4422	0.1586	0.9884	0.2791	0.3235	0.0715	0.9879
	DHS	1.11	0.5348	0.1280	0.9868	0.2121	0.3872	0.0594	0.9881
$\rho = 0.50$	SPSL	2.76	0.3267	0.2600	0.9946	0.6471	0.2536	0.1057	0.9898
	BAL	3.21	0.2827	0.3114	0.9981	0.8737	0.2278	0.1144	0.9870
	HS	2.13	0.3693	0.1903	0.9864	0.4316	0.2859	0.0879	0.9887
	DL	2.21	0.3544	0.2291	0.9939	0.5116	0.2752	0.0977	0.9907
	RHS	2.12	0.3599	0.1944	0.9868	0.4198	0.2910	0.0868	0.9885
	DHS	1.85	0.4337	0.1624	0.9850	0.3291	0.3506	0.0741	0.9888
$\rho = 0.75$	SPSL	5.22	0.3433	0.3254	0.9904	1.23	0.2698	0.1467	0.9885
	BAL	5.77	0.3284	0.3875	0.9954	1.58	0.2455	0.1574	0.9866
	HS	4.50	0.3840	0.2512	0.9805	0.9078	0.2818	0.1221	0.9860
	DL	4.66	0.4054	0.3030	0.9903	1.00	0.2915	0.1356	0.9884
	RHS	4.41	0.3748	0.2590	0.9826	0.8957	0.2844	0.1216	0.9856
	DHS	4.49	0.4476	0.2254	0.9768	0.7623	0.3345	0.1064	0.9857

This table presents performance metrics based on 95% HDIs, averaged across simulation replicates. The measures are derived from the multiple regression setting (Simulation Study II) with a $D = 100$ dimensional parameter vector drawn from a log-normal distribution. The design matrix is generated by a multivariate normal distribution with correlated predictors. We perform $\Omega = 100$ replications for all combinations of observations $N \in \{80, 200\}$ and correlations $\rho \in \{0.25, 0.5, 0.75\}$. We consider the spike-and-slab (SPSL), Bayesian adaptive lasso (BAL), horseshoe (HS), Dirichlet-Laplace (DL), regularized horseshoe (RHS), and Dirichlet-horseshoe (DHS) priors.

Table B.4
Results (II) of Simulation Study II

		$N = 80$				$N = 200$			
		TypeIErr	Rhat	Minutes	Divergences	TypeIErr	Rhat	Minutes	Divergences
$\rho = 0.25$	SPSL	0.0000	1.00	0.3974	0	0.0025	1.00	0.5614	0
	BAL	0.0000	1.00	0.8220	120	0.0076	1.00	0.9750	94.5
	HS	0.0003	1.00	0.7104	64.9	0.0005	1.00	0.9305	76.8
	DL	0.0001	1.00	0.9277	0.3800	0.0014	1.00	1.06	2.99
	RHS	0.0003	1.00	0.5096	0.1500	0.0005	1.01	0.6999	0.4700
	DHS	0.0003	1.00	1.58	11	0.0003	1.01	1.66	24.7
$\rho = 0.50$	SPSL	0.0004	1.00	0.5328	0	0.0022	1.00	0.6584	0
	BAL	0.0001	1.00	0.8209	82.4	0.0063	1.00	0.9727	79.8
	HS	0.0003	1.00	0.7402	69.1	0.0006	1.00	0.9357	58.6
	DL	0.0001	1.00	0.9274	1.17	0.0011	1.00	1.06	6.06
	RHS	0.0003	1.01	0.5613	0.0700	0.0005	1.00	0.9318	0.3900
	DHS	0.0001	1.00	1.58	13.4	0.0003	1.00	1.66	23
$\rho = 0.75$	SPSL	0.0006	1.00	0.6019	0	0.0022	1.00	0.8126	0
	BAL	0.0004	1.00	0.8207	68.8	0.0046	1.00	0.9755	88
	HS	0.0006	1.00	0.7695	71.2	0.0004	1.00	0.9376	79.6
	DL	0.0006	1.00	0.9262	1.36	0.0013	1.00	1.06	6.39
	RHS	0.0009	1.00	0.6668	0.0400	0.0006	1.00	1.07	0.1300
	DHS	0.0009	1.00	1.58	10.5	0.0004	1.00	1.66	16.3

See the descriptions for Table B.3.

B.2 Priors Simulation Studies

B.2.1 Spike-and-Slab

$$\begin{aligned}\beta_j \mid \lambda_j, c, &\sim \lambda_j \mathcal{N}\left(0, c^2\right) \\ \lambda_j &\sim \text{Beta}\left(.5, .5\right) \\ c^2 &\sim \text{InvGamma}(1, 4)\end{aligned}$$

B.2.2 Bayesian Adaptive Lasso

$$\begin{aligned}\beta_j \mid \lambda_j, &\sim \text{DE}\left(0, \frac{1}{\lambda_j}\right) \\ \lambda_j, &\sim \mathcal{C}^{+}\left(0, 1\right)\end{aligned}$$

B.2.3 Horseshoe

$$\begin{aligned}\beta_j \mid \lambda_j, \tau &\sim \mathcal{N}\left(0, \lambda_j^2 \tau^2\right) \\ \lambda_j &\sim \mathcal{C}^{+}\left(0, 1\right) \\ \tau &\sim \mathcal{C}^{+}\left(0, 1\right)\end{aligned}$$

B.2.4 Dirichlet-Laplace

$$\begin{aligned}\beta_j \mid \phi_j, \tau &\sim \text{DE}\left(0, \phi_j \tau\right) \\ \boldsymbol{\phi} &\sim \text{Dir}\left(.5, \ldots, .5\right) \\ \tau &\sim \text{Gamma}\left(\frac{D}{2}, \frac{1}{2}\right)\end{aligned}$$

B.2.5 Regularized Horseshoe

$$
\begin{aligned}
\beta_j \mid \lambda_j, \tau, c &\sim \mathcal{N}\left(0, \tau^2 \tilde{\lambda}_j^2\right) \\
\tau &\sim \mathcal{C}^+\left(0, \tau_0^2\right) \\
\tilde{\lambda}_j^2 &= \frac{c^2 \lambda_j^2}{c^2 + \tau^2 \lambda_j^2} \\
\lambda_j &\sim \mathcal{C}^+(0, 1) \\
c^2 &\sim \text{InvGamma}(1, 4) \\
\tau_0 &= \frac{p_0}{D - p_0} \frac{\sigma}{\sqrt{N}} \\
p_0 &= \eta D, \\
\eta &\sim \text{Beta}(2, 8)
\end{aligned}
$$

B.2.6 Dirichlet-Horseshoe

$$
\begin{aligned}
\beta_j \mid \lambda_j, \psi_j, \tau, c &\sim \mathcal{N}\left(0, \tau^2 \tilde{\lambda}_j^2 \tilde{\phi}_j^2\right) \\
\tau &\sim \mathcal{C}^+\left(0, \tau_0^2\right) \\
\tilde{\lambda}_j^2 &= \frac{c^2 \lambda_j^2}{c^2 + \tau^2 \lambda_j^2} \\
\tilde{\boldsymbol{\phi}} &= \boldsymbol{\phi} D \\
\lambda_j &\sim \mathcal{C}^+(0, 1) \\
\boldsymbol{\phi} &\sim \text{Dir}(\nu, \ldots, \nu) \\
c^2 &\sim \text{InvGamma}(1, 4) \\
\tau_0 &= \frac{p_0}{D - p_0} \frac{\sigma}{\sqrt{N}} \\
p_0 &= \nu D \\
\nu &\sim \text{Beta}(2, 8)
\end{aligned}
$$

Chapter 4

Price Formation Dynamics and Learning in the Tennis Sports Betting Market[22]

4.1 Motivation

Understanding the mechanisms behind price formation and equilibrium dynamics is a central concern in financial markets research. The concept of tâtonnement, introduced by Walras (1889), describes a process where agents adjust their buy and sell offers until supply and demand reach equilibrium—a fundamental aspect of price discovery. Biais, Hillion, and Spatt (1999) apply this concept to the preopening period in the Paris Bourse, revealing that while early preopening prices are noisy, they become more informative and efficient as the market nears opening. This convergence towards an equilibrium valuation is key to understanding price formation.

Although the existing literature extensively explores learning behavior in broader financial markets, studies focusing on this process during the pre-match phase of sports betting markets remain notably scarce. This paper addresses this gap by investigating the information content of pre-match price movements in the tennis betting market, focusing on understanding the price formation process, identifying potential mispricing, and evaluating the market's learning rate.

Research on betting market efficiency consistently finds evidence that betting markets may not be efficient, revealing inefficiencies in how new information is incorporated into prices, i.e., learning. In an efficient market, the price at any given time t reflects the conditional expectation of a betting contract, thereby eliminating opportunities for profitable strategies based on mispricing. Early work by Thaler and Ziemba (1988) highlights market inefficiencies in the horse racing market, where the expected returns from betting on favorites significantly exceed those from betting on underdogs—a phenomenon known as the favorite-longshot bias.

Subsequent studies uncover similar inefficiencies in other sports betting contexts. For example, Gray and Gray (1997) find that simple strategies like betting on home-team underdogs in the NFL can yield positive returns, often driven by market overreactions to recent team performances. Forrest and Simmons (2008) show that sentiment biases odds in Spanish and Scottish soccer matches, with bookmakers exploiting bettors' preferences by pricing popular bets higher.

In contrast, Croxson and James Reade (2014) analyze goal arrivals in soccer using high-frequency in-play data from the betting exchange Betfair and find that prices respond swiftly and fully to new information related to goals scored. However, the focus on goals scored near half-time limits the generalizability of their findings to events occurring during regular play. A more recent study by Hegarty and Whelan (2024) finds that betting markets for soccer and tennis

[22] This chapter is a revised version of Nagel (2024), available at SSRN.

do not fully adhere to the strong-form definition of market efficiency, as expected returns on favorites tend to be higher than those on underdogs, exhibiting a favorite-longshot bias. However, they argue that this bias does not lead to consistently profitable betting strategies, suggesting that weak-form market efficiency cannot be entirely rejected.

Further exploring the betting market, Gandar et al. (1998) investigate how changes in betting lines from opening to closing reflect the information content of these lines and improve the accuracy of game outcome forecasts. Their findings suggest that informed traders play a significant role in the market, as the shifts in betting lines tend to enhance predictive accuracy, indicating that the market assimilates new information effectively as it becomes available.

Moskowitz (2021) reveals significant evidence of momentum by analyzing a wide array of betting contracts, aligning with the idea of delayed overreaction and contradicting the theories of underreaction and rational pricing. Ramirez, Reade, and Singleton (2023) analyze the relationship between Wikipedia page views of professional tennis players and bookmakers' pricing, finding that a "buzz factor"—measured by the difference in pre-match page views relative to the previous year—significantly predicts mispricing. They demonstrate that betting on players with higher pre-match buzz generates substantial profits, suggesting that sportsbooks could improve pricing efficiency by considering this buzz factor.

Most existing studies that investigate the pre-match phase focus on aggregated effects from the opening to the closing using cross-sectional data, often from a single bookmaker (e.g., Gandar et al. 1998). This approach overlooks the learning dynamics within the market and the valuable insights embedded in intermediate price changes. To capture these dynamics, it is crucial to analyze the entire time series of prices throughout the market's betting period (i.e., from opening to closing) across multiple bookmakers. To the best of our knowledge, this paper is the first to explore these pre-match learning dynamics within the sports betting market using these longitudinal data[23].

Moskowitz (2021) highlights the complexity of financial markets as empirical laboratories, which are influenced by various factors. The joint hypothesis problem, as described by Fama (1970), complicates hypothesis testing because it simultaneously evaluates both the hypothesis and the validity of the underlying model. This interdependence makes it challenging to distinguish between flawed hypotheses and incorrect models. However, the sports betting market offers a unique setting for research compared to traditional financial market securities. Sports bets are isolated from systematic economic risks and have exogenous terminal values, allowing for a clearer analysis of bettor behavior and pricing anomalies without the confounding effects of broader market risks (Moskowitz 2021; Palacios-Huerta 2023).

Additionally, Moskowitz (2021) emphasizes that the short, predetermined termination dates of sports betting contracts, determined by the outcomes of events independent of investor ac-

[23]We specifically refer to learning during the pre-match phase, defined by the first and last prices set by bookmakers before kickoff. In contrast, in-play learning, such as market responses to major events like goals during a match, has been more frequently studied (e.g., Croxson and James Reade 2014; Angelini, De Angelis, and Singleton 2022).

tions, provide a clean basis for identifying mispricing and testing theories about price corrections. Using the sports betting market as a controlled laboratory allows for the testing of theories and hypotheses that could be extrapolated and generalized to broader financial markets. Both markets share parallels, including high transaction volumes, liquidity, and the availability of information, all of which influence decisions and outcomes. Market makers provide liquidity and facilitate trades or bets, while professional analysts and bettors engage in arbitrage, mirroring financial market activities.

Behavioral patterns, such as those explained by psychological theories (e.g., Krieger, Davis, and Strode 2021; Goodell et al. 2023), are observed in both markets and indicate similar underlying decision-making processes. While entertainment motives are more apparent in sports betting, they are also present in financial markets, where investors often derive non-monetary satisfaction from their investments (Gao and Lin 2015).

Thus, the findings of this study may not only have significant implications for understanding price discovery and learning in the sports betting market, but also for providing potentially relevant insights for the broader financial market.

4.1.1 Market Structure

Tennis is either played between two players (singles) or between two teams of two players each (doubles). To ensure a homogeneous sample, we focus exclusively on singles matches. Unlike many team sports, tennis has fewer variables that can influence match outcomes right before the start, such as lineup changes in soccer. Additionally, there are only two potential match outcomes: a player can either win or lose. With only one player per side, the number of uncontrollable exogenous factors is minimized, making singles tennis an attractive choice for professional bettors.

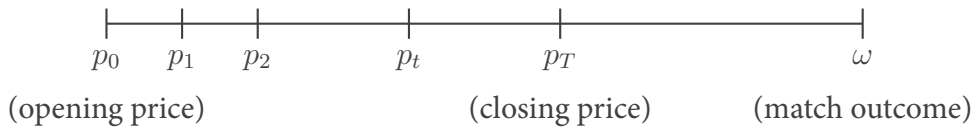
In this paper, we focus on the fixed-odds market for tennis, a popular betting format in which odds are fixed at the time bets are placed, regardless of subsequent changes. Specifically, we examine moneyline betting, where bettors wager on which player will win the match. The fixed odds we consider are expressed as decimal odds, also known as European odds, and presented as a decimal number (e.g., 2.50). This decimal number represents the total payout a bettor will receive for a winning bet per unit staked. Therefore, the total amount returned for a winning bet, including the initial stake, is the product of the stake and the odds. If a bettor loses a bet, they lose their stake.

While odds represent the potential payout of a bet, their reciprocals—known as implied probabilities, which can be viewed as the price p of taking the bet—reflect the likelihood of winning. For example, if the decimal odds for Player 1 to win are 2.50, the implied probability is $1/2.50 = 0.40$. Higher implied probabilities imply lower odds and a lower potential payout, reflecting a lower risk and thus a higher price. Conversely, lower implied probabilities indicate higher odds and a higher payout, reflecting a higher risk and thus a lower price.

Implied probabilities reflect the sum of the estimated probabilities and the margin the bookmaker incorporates to make profit. As the outcome is binary and the probabilities sum to 1, the implied probability for Player 2 is the difference between 1 and the implied probabilities for Player 1, deducting the margin. Therefore, the implied probabilities for Player 1 and Player 2 are almost perfectly negatively correlated: an increase in Player 1's implied probabilities by a certain amount necessitates an equal decrease in Player 2's implied probabilities, if the margin remains constant. However, since the margin may vary, the correlation is not perfectly negative. In tennis, the assignment of a particular player to Player 1 or Player 2 is not systematically related to their winning prospects.

Bookmakers dynamically place and adjust their prices for both potential match outcomes from the market open to the market close. According to Gandar et al. (1998) and Moskowitz (2021), bookmakers set an initial price on each contract at the open with the goal of maximizing risk-adjusted profits. By trying to either balance the total stakes on both sides of the contract or its outcome probabilities, bookmakers ensure that they receive the margin with minimal risk exposure. If the bookmaker's predictions are more accurate than the bettors' predictions, they might also decide to take some risk exposure to earn higher profits.

Figure 4.1
Timeline of a Betting Contract



This figure presents the typical timeline of a betting contract, showing the price series from the initial opening price (p_0) to the final closing price (p_T), along with the terminal value, representing the match outcome (ω).

Figure 4.1 depicts the timeline of a betting contract, showing the prices from the opening to the closing, along with the terminal value. Once a bookmaker sets the opening price p_0 for a given match, bettors can place their bets until the start of the match. During this period, betting can change the price $(p_1, p_2, \ldots, p_T)$ as bookmakers attempt to balance the betting volume on both sides of the contract. Prices can also shift without changes in volume if new information, such as a player's injury, becomes available (Moskowitz 2021).

4.2 Methods

We begin our analysis by examining the cross-section of opening and closing prices before proceeding to the total time series. Throughout this article, we use N to denote the number of matches, with $i = 1, \ldots, N$, and K to specify the number of bookmakers, where $j = 1, \ldots, K$. Time runs from $t = 1, \ldots, T$. We represent the match outcome by ω, where $\omega = 0$ if Player 1 loses and $\omega = 1$ if Player 1 wins the match. Vectors and matrices are highlighted in bold.

4.2.1 Forecast Accuracy of Individual Bookmakers

To assess the effectiveness of bookmakers in predicting match outcomes, we calculate the root mean squared error (RMSE) for each bookmaker. This metric measures the forecast error of their opening prices in predicting match results. By analyzing these RMSE values, we evaluate the forecast accuracy of each bookmaker's opening prices.

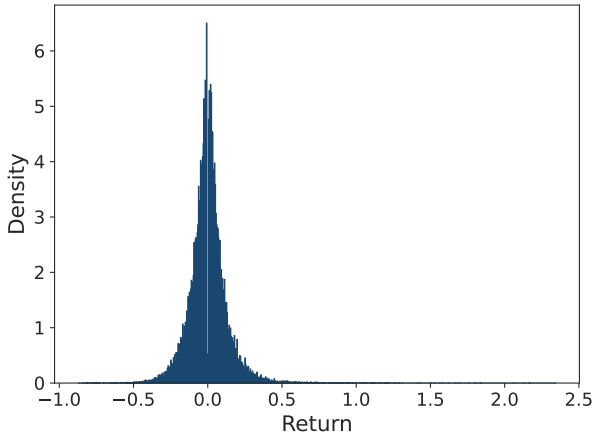
It is important to note that accuracy alone does not necessarily imply well-calibrated probabilities. For example, if a model predicts a probability of 0.8, it should be correct 80% of the time when making such predictions. Evaluating the overall reliability of predicted probabilities requires considering metrics like the expected calibration error, which is not the primary focus of this analysis.

4.2.2 Price Change Mechanisms: Reasons and Implications

Price movements from the open to the close convert bookmakers' forecasts of match outcomes into market forecasts (Gandar et al. 1998). Since the terminal values of the contracts (i.e., actual match outcomes) are observable, we can examine whether these price movements contain information relevant to predict the match outcome.

Figure 4.2 presents the distribution of the open to close returns, calculated as $r_T = p_T/p_0 - 1$, for all matches and bookmakers. The returns are approximately normally distributed and,

Figure 4.2
Distribution of Open-to-Close Returns



This figure presents the distribution of open-to-close returns based on all bookmakers and matches. The open-to-close returns are calculated as $\boldsymbol{r}_T = \boldsymbol{p}_T/\boldsymbol{p}_0 - 1$. The distribution is approximately normal and ranges from -0.5 to 0.5, with some outliers.

excluding some outliers, range from -0.5 to 0.5 with an interquartile range of 0.1237. Note that, unlike in the stock market, these returns are not realizable through strategies such as going short at the opening and long at the closing, as it is not possible to sell odds at the current price and later back them at a different price.

Gandar et al. (1998) offer three explanations for betting imbalances at the opening stages that give rise to price changes: the release of new information, randomness of order flow, and

prediction errors by the bookmakers. Given the relatively short betting windows with a median of 19.57 hours, the arrival of new information is unlikely to be responsible for all price changes.

For the NFL betting market, Gandar et al. (1988) show that the release of new information during short betting periods can explain only a small fraction of all price movements. Furthermore, in singles tennis, there are no short-term changes in lineups that can distort expectations, as often happens in team sports like soccer. While randomness of order flow could explain some of the smaller price changes, it is an unlikely explanation for the medium-sized and larger-sized price changes depicted in Figure 4.2.

4.2.3 Price Movements on New Information

If prices move based on information (e.g., relevant information about the winning prospects is announced after the open but before the match starts) and the market reacts rationally to the news, the closing price should better predict the match outcome than the opening price (Moskowitz 2021). In this scenario, the closing price becomes the conditional expectation of the match outcome, meaning there is no return predictability from the close to the end of the match, and changes from the open to the close have no explanatory power for the return from the close to the end of the match.

To analyze the response to information, we adjust the linear regression proposed by Moskowitz (2021) by incorporating bookmaker random effects and controlling for a vector of covariates. The hierarchical model is specified as follows:

$$\begin{aligned} r_{ij\omega} &= \nu_{0j} + \nu_{1j} r_{ijT} + \boldsymbol{\theta} \boldsymbol{X}_{ij} + \varepsilon_{ij}, \\ \nu_{0j} &= \nu_0 + u_{0j}, \\ \nu_{1j} &= \nu_1 + u_{1j}, \end{aligned} \tag{4.1}$$

where $r_{ij\omega}$ represents the match and bookmaker-specific close-to-end return, calculated as $r_{ij\omega} = \omega_i / p_{ijT} - 1$. The fixed effects matrix $\boldsymbol{X}$ comprises the time from open to close and competition indicators (ATP, WTA, ...), with $\boldsymbol{\theta}$ representing the corresponding coefficients. We denote the bookmaker random effects for the intercept and slope by u_{0j} and u_{1j}, respectively, and ε_{ij} is a Gaussian white noise innovation term.

Moskowitz (2021) demonstrate that $\nu_1 = 0$ if prices move based on information and markets respond rationally. If prices move for non-informational reasons, such as investor sentiment or noise, $\nu_1 = -1$ because biased closing prices are corrected by the exogenous match outcomes. Finally, prices can move based on information, but markets may respond irrationally. For example, there could be over- or underreactions to information, which are two of the leading behavioral mechanisms in the asset pricing literature (see, e.g., Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Moskowitz (2021)). In these cases, $\nu_1 > 0$ if the market underreacts and $\nu_1 < 0$ if the market overreacts to information.

4.2.4 Relative Forecast Accuracy of Opening and Closing Prices

If bookmakers introduce bias into initial prices, we can consider two scenarios (Gandar et al. 1998). In the first, opening prices are biased, either intentionally or due to prediction errors, while subsequent price movements result from noise. Bookmakers may intentionally bias opening prices to maximize profits or achieve a balanced book, as observed by Dare and MacDonald (1996) in the wagering market on Super Bowl games. The "hot hand" phenomenon explored by Camerer (1989) and Brown and Sauer (1993) illustrates irrational betting behavior that bookmakers might account for when setting initial prices. If biased opening prices remain uncorrected by the bettors, closing prices would have equal or less predictive power for match outcomes compared to opening prices. In the second scenario, bettors identify and exploit the biased opening prices. This would lead to closing prices becoming superior predictors of match outcomes.

To investigate the statistical difference between the relative predictive abilities of the two forecasts, we use the AGS test as applied by Gandar et al. (1998) and developed by Ashley, Granger, and Schmalensee (1980). The test procedure involves calculating the forecast errors implied by both the opening and closing prices:

$$\underbrace{e_{ij0} = |\omega_i - p_{ij0}|}_{\text{forecast error opening}} \quad \text{and} \quad \underbrace{e_{ijT} = |\omega_i - p_{ijT}|}_{\text{forecast error closing}},$$

where e_{ij0} and e_{ijT} are the match and bookmaker-specific forecast errors of the opening and closing prices, respectively. Letting

$$\Delta_{ij} = e_{ij0} - e_{ijT} \quad \text{and} \quad \Sigma_{ij} = e_{ij0} + e_{ijT},$$

with Δ_{ij} being the difference and Σ_{ij} the sum of the forecast errors, we can estimate the regression implied by the AGS test. As before, we use a random effects model to account for the nested data structure and control by a vector of fixed effects:

$$\begin{aligned} \Delta_{ij} &= \lambda_{0j} + \lambda_{1j} \left(\Sigma_{ij} - \overline{\Sigma}_j\right) + \boldsymbol{\theta} \boldsymbol{X}_{ij} + \varepsilon_{ij}, \\ \lambda_{0j} &= \lambda_0 + u_{0j}, \\ \lambda_{1j} &= \lambda_1 + u_{1j}, \end{aligned} \tag{4.2}$$

with $\overline{\Sigma}_j$ being the mean forecast error of bookmaker j. The individual estimates for the intercept and slope allow us to draw conclusions about the relative forecast accuracy of opening and closing prices: the intercept λ_0 provides information about the difference in the mean absolute forecast errors, while the slope λ_1 tests for a significant difference in the forecast error variances (Gandar et al. 1998). Specifically, an intercept estimate significantly larger (smaller) than zero indicates a smaller (larger) mean forecast error for closing prices. Similarly, a slope estimate significantly larger (smaller) than zero indicates a smaller (larger) mean forecast error variance for closing prices.

4.2.5 Magnitude and Direction of Price Movements

If price movements from the open to the close indicate that the market perceives a mispricing at the open, this suggests that a player is either undervalued or overvalued. For instance, an increasing price implies that the market believes a player was initially undervalued, leading bettors to strategically place their bets on that player to exploit the mispricing.

When rational bettors dominate irrational bettors in the market in terms of aggregated betting volume, the undervalued player attracts higher stakes relative to their initial price. As a result, bookmakers need to adjust the price upward to balance the book by the closing. The magnitude of the price change can also indicate the market's confidence in how much a player was over- or undervalued at the opening (Gandar et al. 1998). Furthermore, the winning rate of players who become more favored during betting (i.e., those experiencing positive price changes) should exceed 0.5. Conversely, the winning rate of players who become less favored (i.e., those experiencing negative price changes) should fall below 0.5. These winning rates are expected to deviate more significantly from 0.5 for larger price changes, with the magnitude of winning rates increasing (or decreasing) proportionately with the extent of the positive (or negative) price changes.

If we organize the matches by the extent of price changes from opening to closing, using intervals $\boldsymbol{\Delta}(p_T, p_0)$ that capture different magnitudes of price changes, and then sort these intervals from the largest negative to the largest positive changes, we would expect to observe winning rates $\overline{\boldsymbol{\omega}}$ within these bins that reflect the price movements. Specifically, winning rates should be significantly lower than 0.5 for intervals indicating negative price changes and significantly higher than 0.5 for intervals with positive price changes, with a monotonic increase in winning rates as we move from negative to positive intervals.

We further examine this pattern by regressing the winning rates on the intervals' average price change magnitudes $\overline{\boldsymbol{\Delta}}(p_T, p_0)$ (Gandar et al. 1998). We use a random effects model to incorporate bookmaker-specific effects and control for the number of price changes:

$$\begin{aligned} \overline{\boldsymbol{\omega}} &= \eta_{0j} + \eta_{1j}\, \overline{\boldsymbol{\Delta}}_j(p_T, p_0) + \kappa C + \varepsilon_j, \\ \eta_{0j} &= \eta_0 + u_{0j}, \\ \eta_{1j} &= \eta_1 + u_{1j}, \end{aligned} \tag{4.3}$$

where κ is the coefficient on the number of price changes.

Since we estimate a regression model with a generated regressor and dependent variable, $\overline{\boldsymbol{\Delta}}_j(p_T, p_0)$ and $\overline{\boldsymbol{\omega}}$, where we replace expected values with sample averages, we must adjust the standard error of the associated regression coefficient. This adjustment is necessary because the generated variables are estimated rather than directly observed, and their variability affects the precision of the coefficient estimates. To address this, we use 1,000 bootstrap samples to approximate the regressor coefficient's distribution and calculate its standard error and confidence intervals from the bootstrap distribution.

If price changes are noise or irrational betting moves prices, the magnitude and direction of price changes cannot explain actual winning rates. In this case, $\eta_0 = 0.5$ and $\eta_1 = 0$. On the other hand, if rational bettors dominate irrational bettors in the market, price changes are proportionately related to winning rates. In this case, $\eta_0 = 0.5$ and $\eta_1 > 0$ (Gandar et al. 1998).

4.2.6 Alternative Hypotheses

Not all price changes are equally informative; instead, we expect to observe a visible learning process. To examine this learning process, we formulate the following two alternative hypotheses, adapted from Biais, Hillion, and Spatt (1999) using the information content embedded in intermediate price changes (i.e., price changes between opening and closing). These hypotheses describe the two extreme cases: price changes being only noisy and price changes being completely informative.

Under the noise hypothesis,

$$H_0 : p_{ijt} = \mathrm{E}\left[\omega_i \mid \mathrm{I}_0\right] + \varepsilon_{ijt},$$

where $\mathrm{E}\left[\omega_i \mid \mathrm{I}_0\right]$ is the expected match outcome conditional on an information set at $t = 0$, the price at time t does not reflect any information learned or processed since the market open. In this case, the placed bets are not informative about the match outcome; their explanatory power is no greater than that of the bookmakers' initial expectations. The information set at $t = 0$ includes all the information available up to and including the market open that bookmakers use to make predictions.

Contrary, under the learning hypothesis,

$$H_A : p_{ijt} = \mathrm{E}\left[\omega_i \mid \mathrm{I}_t\right], \tag{4.4}$$

where $\mathrm{E}\left[\omega_i \mid \mathrm{I}_t\right]$ is the expected match outcome conditional on an information set at time t, bettors act competitively and drive the price to the conditional expectation of the match outcome by exploiting immediate betting opportunities that the pricing offers. Consequently, the price at time t reflects the conditional expectation for the match outcome, implying that the predictive power of these prices increases over the period from the open to the close.

To test the alternative hypotheses, we need to regress the returns of the prices from the open to the end onto the returns of the prices from the open to each time t. Following Hodrick (1987) and Biais, Hillion, and Spatt (1999), we employ so-called unbiasedness regressions to control for a potentially time-varying distribution of the price series, which may be induced through the learning process itself. Specifically, we estimate one unbiasedness regression across matches for each t between the beginning and the end of the betting window, controlling for bookmaker random effects. Thus, we analyze the distribution of prices for each point in time t.

Under the learning hypothesis, the slope coefficient β_1 in the regression

$$\begin{aligned} \omega_i - \mathrm{E}[\omega_i \mid \mathrm{I}_0] &= \beta_{0jt} + \beta_{1jt}\left(p_{ijt} - \mathrm{E}[\omega_i \mid \mathrm{I}_0]\right) + \varepsilon_{ijt}, \\ \beta_{0jt} &= \beta_{0t} + u_{0jt}, \\ \beta_{1jt} &= \beta_{1t} + u_{1jt}, \end{aligned} \tag{4.5}$$

is equal to one for all t. Additionally, the residual variances measuring the remaining uncertainty in the predictions should decrease as the kickoff gets closer. Under the noise hypothesis, the slope coefficient β_1 is equal to zero for all t because intermediate price changes at t have no informational content and do not help to predict the deviation between the actual match outcome and its ex ante expectation, such that

$$\mathrm{cov}\left(\omega_i - \mathrm{E}\left[\omega_i \mid I_0\right], p_{ijt} - \mathrm{E}\left[\omega_i \mid I_0\right]\right) = \mathrm{cov}\left(\omega_i - \mathrm{E}\left[\omega_i \mid I_0\right], \boldsymbol{\varepsilon}_{ijt}\right) = 0.$$

Additionally, the residual variances of the regressions should remain at the same level.

Positive slope coefficients that deviate from zero indicate improving, yet still biased, intermediate prices, with the bias increasing as the coefficient diverges further from one. In such cases, the learning hypothesis remains rejected. Conversely, coefficients smaller than zero imply that intermediate price changes not only lack informational value but also deteriorate the accuracy of forecasts.

4.2.7 Learning Rate

To examine how fast information is processed into prices, we estimate the learning rate using the methodology outlined by Biais, Hillion, and Spatt (1999) for financial market securities. The procedure builds on the work of Vives (1995), who characterizes the asymptotic speed of learning by

$$t^{\gamma}\left(p_t - \omega\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right), \quad t \to \infty, \tag{4.6}$$

where γ is the convergence, or learning rate, and σ represents the standard deviation of a normal distribution. Because Equation (4.6) only holds asymptotically, we consider large values of t (i.e., the end of the period) and approximate the equation by

$$p_t - \omega = \frac{\varepsilon}{t^{\gamma}}\sigma,$$

where ε is normally distributed with zero conditional mean $\mathrm{E}\left[\varepsilon \mid \mathrm{I}_t\right] = 0$ and unit conditional variance $var(\varepsilon \mid \mathrm{I}_t) = 1$. Taking squares and expectations, we get:

$$\mathrm{E}\left[\left(p_t - \omega\right)^2 \mid \mathrm{I}_t\right] = \frac{\sigma^2}{t^{2\gamma}}. \tag{4.7}$$

Repeating the procedure for $t-1$ yields:

$$\mathrm{E}\left[(p_{t-1}-\omega)^2 \mid \mathrm{I}_{t-1}\right] = \frac{\sigma^2}{(t-1)^{2\gamma}}. \tag{4.8}$$

Dividing Equation (4.7) by Equation (4.8), we get:

$$\frac{\mathrm{E}\left[(p_t-\omega)^2 \mid \mathrm{I}_t\right]}{\mathrm{E}\left[(p_{t-1}-\omega)^2 \mid \mathrm{I}_{t-1}\right]} = \left(\frac{t-1}{t}\right)^{2\gamma}.$$

Implementing the same steps for the comparison between p_{t-1} and p_{t-2}, we obtain the two moment conditions

$$\boldsymbol{m}(\boldsymbol{p},\gamma) = \begin{bmatrix} \mathrm{E}\left[(p_t-\omega)^2 - \left(\frac{t-1}{t}\right)^{2\gamma}(p_{t-1}-\omega)^2 \mid \mathrm{I}_{t-1}\right] = 0 \\ \mathrm{E}\left[(p_{t-1}-\omega)^2 - \left(\frac{t-2}{t-1}\right)^{2\gamma}(p_{t-2}-\omega)^2 \mid \mathrm{I}_{t-2}\right] = 0 \end{bmatrix}, \tag{4.9}$$

with γ being the learning rate parameter to be estimated. Biais, Hillion, and Spatt (1999) point out the following two convenient statistical properties of the moment conditions. First, they are robust to heteroscedasticity reflected by changes in σ across time and matches because σ^2 cancels out. Second, the moment conditions are robust to temporal aggregation. If the true learning rate is not t but scaled by a constant nt, Equation (4.7) becomes:

$$\mathrm{E}\left[(p_t-\omega)^2 \mid \mathrm{I}_t\right] = \frac{\sigma^2}{(nt)^{2\gamma}}, \tag{4.10}$$

and Equation (4.8) becomes:

$$\mathrm{E}\left[(p_{t-1}-\omega)^2 \mid \mathrm{I}_{t-1}\right] = \frac{\sigma^2}{[n(t-1)]^{2\gamma}}, \tag{4.11}$$

which still yields the same moment conditions, because n is eliminated when we divide Equation (4.10) by Equation (4.11).

4.2.7.1 Generalized Methods of Moments Estimation

To estimate the learning rate, we utilize the Generalized Methods of Moments (GMM) as proposed by Hansen (1982). The GMM estimator is consistent, asymptotically normal, and efficient. The estimation process involves solving the minimization problem

$$\arg\min_{\gamma} \left[\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{g}_N(\boldsymbol{p},\gamma)\right]' \boldsymbol{W}_N \left[\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{g}_N(\boldsymbol{p},\gamma)\right], \tag{4.12}$$

where $\boldsymbol{g}_N(\boldsymbol{p},\gamma)$ is the vector of moment conditions, and $\boldsymbol{W}_N$ is the weighting matrix.

Despite the system already being overidentified (2 moment conditions, 1 parameter), the

inclusion of instruments can enhance the precision of the estimates. An instrument z needs to be any variable that belongs to the information set at $t-2$, i.e., $z \in I_{t-2}$. Following Biais, Hillion, and Spatt (1999), we employ the following seven instruments:

$$\boldsymbol{z} = \begin{bmatrix} 1 \\ p_{t-2} - p_{t-3} \\ (p_{t-2} - p_{t-3})^2 \\ p_{t-3} - p_{t-4} \\ (p_{t-3} - p_{t-4})^2 \\ p_{t-4} - \mathrm{E}\left[\omega \mid \mathrm{I}_0\right] \\ (p_{t-4} - \mathrm{E}\left[\omega \mid \mathrm{I}_0\right])^2 \end{bmatrix}.$$

Including these instruments increases the number of moment conditions to 14, such that $\boldsymbol{g}_N(\boldsymbol{p}, \gamma)$ is of dimension (14×1).

First, we derive the estimates by minimizing the objective function (4.12) using first-stage GMM. Next, we apply the continuous-updating estimator (CUE), as proposed by Hansen, Heaton, and Yaron (1996), to improve the precision of our estimates. In the CUE approach, the weighting matrix $\boldsymbol{W}_N$ is replaced by the optimal weighting matrix $\boldsymbol{W}(\gamma)$, which is continuously updated during the minimization process. This involves re-estimating the matrix at each iteration based on the current parameter estimates.

The optimal weighting matrix $\boldsymbol{W}$ is a consistent estimator of the inverse of the variance-covariance matrix of the moment conditions, denoted as $\boldsymbol{W} = \boldsymbol{S}^{-1}$. It assigns greater weight to more precise moment conditions (i.e., those with lower variance) and less weight to moment conditions with higher variance, thus improving the efficiency of the estimator.

For the choice of times $t, t-1, \ldots, t-4$, we follow the recommendation of Biais, Hillion, and Spatt (1999) by selecting every fifth time increment to mitigate numerical instability. Specifically, we consider the final observation for t, the fifth last observation for $t-1$, and so on. This approach balances the need for these times to be close to the end of the period with the goal of reducing instability in parameter estimation.

The GMM estimator requires specifying a starting value for the parameter to be estimated. We use $\gamma = 0.01$. However, the starting value can significantly impact the estimates. Thus, we perform the estimation for nine additional randomly selected starting values (drawn from a uniform distribution with intervals $\gamma \in [0, 1]$) to verify that the estimates are not highly sensitive to the choice of starting values. We conduct the estimation separately for each bookmaker. To evaluate the validity of the model's moment conditions, we use the Hansen J-test. This test is asymptotically χ^2-distributed with degrees of freedom equal to the number of moment conditions (14) minus the number of parameters (1). The J-statistic is computed as:

$$J_N = N\boldsymbol{g}_N(\boldsymbol{p}, \hat{\gamma})' \left[\widehat{Avar}(\boldsymbol{g}_N(\boldsymbol{p}, \hat{\gamma}))\right]^+ \boldsymbol{g}_N(\boldsymbol{p}, \hat{\gamma}) \underset{d}{\rightarrow} \chi^2 \quad (13)$$

4.2.7.2 Bayesian Estimation

We use GMM to estimate bookmaker-specific learning rates by analyzing each bookmaker's data separately. However, our comprehensive dataset, which includes longitudinal data across multiple bookmakers, also enables us to estimate the overall average learning rate and specific deviations for each bookmaker using the entire dataset, including all bookmakers, through the moment conditions (4.9). This approach effectively accounts for the nested structure of the data and enhances the accuracy and precision of our estimates.

Bayesian estimation is advantageous in this context, as it facilitates the formulation of such models while also providing the full posterior distribution of the estimated parameters. Furthermore, this estimation procedure helps confirm the reliability of the results.

Using our prior knowledge about the domain and distribution of the parameters—for example, recognizing that the learning rate, as a measure of convergence, should be non-negative—we apply the following weakly informative prior distributions:

$$
\begin{aligned}
\gamma_j &\sim \mathcal{N}^T\left(\mu_\gamma, \sigma_\gamma^2; 0, \infty\right), \\
\mu_\gamma &\sim \mathcal{N}^T\left(0, 1; 0, \infty\right), \\
\sigma_\gamma &\sim \text{Exp}\left(2.5\right), \\
\sigma_\varepsilon &\sim \mathcal{C}^+\left(0, 0.01\right).
\end{aligned}
$$

Here, $\mathcal{N}^T\left(\mu_\gamma, \sigma_\gamma^2; 0, \infty\right)$ denotes a truncated normal distribution with a lower bound of zero. We use the truncated normal distribution for the bookmaker-specific learning rates to maintain the shape of the normal distribution while ensuring non-negativity. For the mean μ_γ, we use a truncated standard normal distribution to model the average learning rate of all bookmakers in the sample, and an exponential distribution for standard deviation σ_γ. For the noise variance σ_ε, we use the more heavily-tailed half-Cauchy distribution. These priors provide some structure and regularization to the model without being too restrictive, thereby enabling the data to effectively guide the shaping of the posterior distributions.

Furthermore, we estimate the average learning rates for favorites and longshots. Matches where Player 1's opening price is higher than Player 2's opening price are categorized as favorites, while the remaining matches are classified as longshots. If the opening prices are equal, the classification is determined by comparing the closing prices using the same logic. Differences in learning rates between these groups could provide valuable insights into the mechanisms driving the favorite-longshot bias.

We extend and refine this analysis by dividing the matches into 10 opening price percentile intervals: 0-10, 10-20, ..., 90-100. This approach allows us to differentiate between more extreme favorites and longshots and their less extreme counterparts, providing insights into how learning rates vary across different price ranges.

Additionally, we investigate whether learning rates vary based on the level of competition.

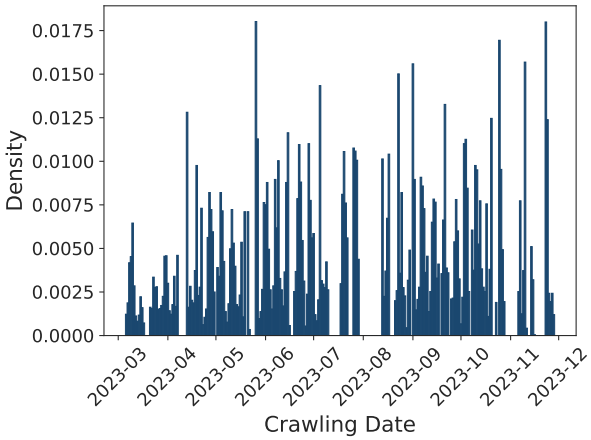
To do this, we classify all matches in the sample as either professional or amateur, designating ATP and WTA competitions as professional and Challenger and ITF competitions as amateur. The share of professionals is equal to 0.3672, while the share of amateurs is 0.6328. This comparison allows us to observe the differences in the estimated parameters between the two groups, providing insights into how learning is driven by the level of competition.

We estimate the parameters using the No-U-Turn sampler (NUTS) and using automatic differentiation variational inference (ADVI). For detailed information on these algorithms and the procedures used, please refer to Appendix C.5.

4.3 Data

Our dataset comprises 39,955 tennis matches, totaling 2,952,877 observations, web-scraped from Oddsportal[24] using Selenium in Python. The data collection spanned approximately 38 weeks, from March 2023 to November 2023, with the scraper running continuously, aside from occasional interruptions due to server and technical issues. Figure 4.3 illustrates the density of the crawling process, highlighting the relative frequency of matches crawled per day.

Figure 4.3
Probability Distribution of Crawled Matches Over Time



This figure presents the relative frequencies of crawled matches per day over the total cralwing period from March 2023 to November 2023. Zero probability mass indicate either that no matches were available at this date or that the crawling process was interrupted due to server or other technical errors.

The dataset comprises detailed information about each match, including match ID, involved players, the country in which the match took place, name of the competition or tour, tournament name, list of bookmakers that offered odds on the match, match outcome, and final score of the match. Most importantly, the dataset includes comprehensive odds information provided by up to 31 bookmakers per match. These odds cover both opening and closing odds (i.e., the first and last odds for a given match) and updates to the odds from open to close of the market.

[24]https://www.oddsportal.com/

4.3.1 Data Preparation

Because the odds for Player 1 are nearly perfectly negatively correlated with those for Player 2, we consider only the odds for Player 1. We exclude matches where the winner is determined by withdrawal, walkover, retirement, default, or disqualification. Including these matches could introduce selection bias, as such incidents may be systematically related to a player's winning prospects and odds. Only if bookmakers correctly account for such cases, the conditional expectation of the match outcome, given the opening odds, would be unbiased.

To eliminate idiosyncrasies related to very small bookmakers, we calculate the number of observations each bookmaker contributes to the sample and discard those in the lowest 25% quantile. This increases the stability, reliability, and interpretability of the estimates, particularly as it removes groups that are too small to provide meaningful estimates or that might introduce excessive noise or bias into the analysis. Additionally, we discard observations with implausible margins, specifically those that are negative or greater than 15%, as well as time series with zero variance.

The timing of when bookmakers first offer their odds varies, meaning that not all bookmakers place and update their odds simultaneously. Likewise, the betting periods varies by match, ranging from several hours to up to a few days. To obtain a homogeneous sample, we remove time series with lengths shorter than 12 hours and longer than 72 hours. Note that Oddsportal provides only the opening odds and the last 20 updates before kickoff. Hence, there are at most 21 observations per match and bookmaker.

Moreover, we use implied probabilities because they have better statistical properties than odds. For example, they range between zero and one if no-arbitrage conditions hold, and their empirical distribution is approximately normal (see Figure C.1 in the appendix).

For the time series analysis, we generate homogeneous time intervals for each match, determined by the timestamp of the first opening odds of all bookmakers and the kickoff time of each match. We resample the data into 1-minute intervals and impute the missing values caused by irregular time intervals using last tick interpolation: for each missing data point, we use the most recent available data point to fill the gap, because the last known observation remains valid until a new observation is available.

Then, we read out the implied probabilities (i.e., prices) at every second time percentile increment of the series, such that each time series has a length of 51 observations (i.e., 50 plus the opening price). This choice is driven by the objective of finding the optimal balance between the length of the time series and their variability, which inherently involves a trade-off.

Subsequently, we impute the remaining missing observations at the beginning of some time series resulting from different opening timestamps using a multivariate imputation strategy. This strategy employs a Bayesian Ridge regression model that incorporates prior distributions over the parameters to predict missing values by accounting for the relationships between multiple features. A detailed description of the imputation strategy is provided in Appendix C.3.

We investigate the statistical properties of the data by applying various time series analyses, testing for a unit root, autocorrelation, and heterogeneity. Appendix C.4 presents the detailed results from these tests.

4.3.2 Descriptive Statistics

Table 4.1 provides a detailed summary of the numerical variables used in this study. Match out-

Table 4.1
Descriptive Statistics for Numerical Variables

	count (in mil.)	mean	std	min	25%	50%	75%	max
Match Outcome	1.65	0.5094	0.4999	0	0	1	1	1
Price	1.65	0.5438	0.1925	0.0272	0.4000	0.5435	0.6849	0.9901
Op. Price	1.65	0.5441	0.1896	0.0307	0.4065	0.5405	0.6803	0.9901
Cl. Price	1.65	0.5434	0.1947	0.0292	0.4000	0.5464	0.6897	0.9901
Time	1.65	26.7	14.8	12	15.1	19.6	37.4	72
No. Price Changes	1.65	11.2	5.58	1	7	11	16	20

This table presents descriptive statistics for the numerical variables used in the study. The columns display the number of observations (*count*) in millions, sample mean (*mean*), standard deviation (*std*), minimum value (*min*), first quartile (*25%*), median (*50%*), third quartile (*75%*), and maximum value (*max*).

come averages 0.5094, implying that the players assigned to position 1 have an unconditional winning rate of 50.94%, demonstrating a nearly equal distribution of match outcomes between Player 1 and Player 2.

Prices range from 0.0272 and 0.9901, with an average price across matches and time of 0.5438. This number is slightly larger than 0.5 due to the margin bookmakers add to the fair prices to make a profit. The corresponding standard deviation is 0.1925, reflecting moderate variability in the prices. The opening and closing prices average 0.5441 and 0.5434, respectively, with standard deviations of 0.1896 and 0.1947. These values indicate that the distribution of prices does not change strongly from the opening to the closing stages of betting. For a comparison, see Figure C.2 in the appendix.

On average, matches are available for betting for 26.7 hours, with a standard deviation of 14.8 hours. Figure C.3 in the appendix illustrates the distribution of the betting periods in hours from the opening prices to the kickoff for all bookmakers and matches in the sample.

The number of price changes per match and bookmaker averages 11.2, with a standard deviation of 5.58, suggesting active adjustment of prices over time. Figure C.4 in the appendix displays the distribution of the number of price movements.

Table 4.2 provides a detailed summary of the categorical variables. There are 24 different bookmakers left in the dataset (*unique*), with Marathonbet being the most frequently represented bookmaker (*top*), appearing 130,432 times (*freq*). Figure C.5 in the appendix illustrates the number of matches offered by each bookmaker. The data covers matches played in 81 different countries, with France being the most common. The most frequent competition is the Challenger

Men. The dataset also includes a wide range of tournaments, with the ATP French Open being the most represented.

Table 4.2
Descriptive Statistics for Categorical Variables

	count (in mil.)	unique	top	freq
Bookmaker	1.65	24	Marathonbet	130,432
Country	1.65	81	France	166,203
Competition	1.65	5	Challenger Men	527,542
Tournament	1.65	1,781	ATP French Open	30,090

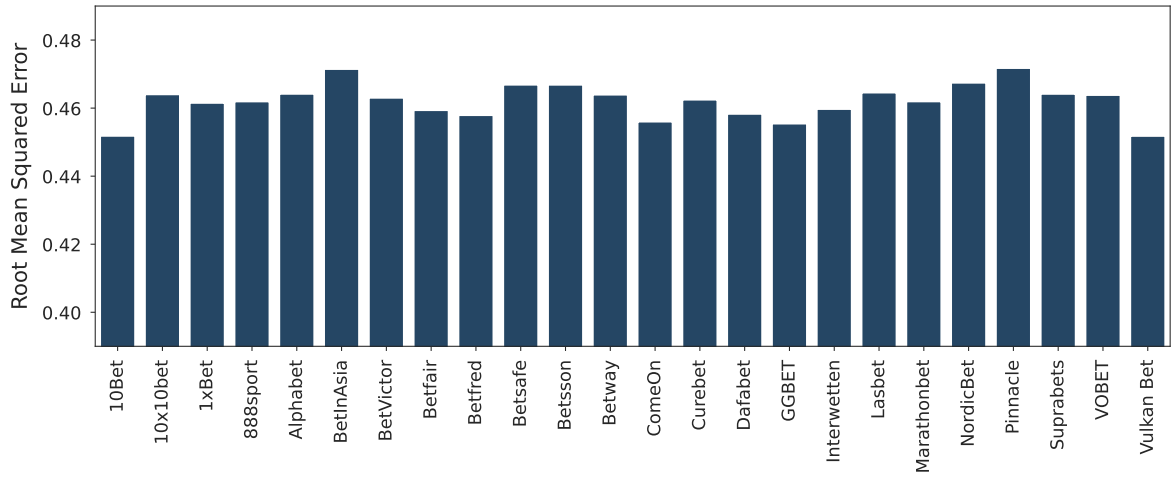
This table presents descriptive statistics for the categorical variables used in the study. The columns display the number of observations (*count*) in millions, number of unique values (*unique*), the most frequent value (*top*), and the number of occurrences of the most frequent value (*freq*).

4.4 Results

4.4.1 Forecast Accuracy of Individual Bookmakers

Figure 4.4 illustrates the bookmaker-specific RMSE, assessing the effectiveness of each bookmaker's opening prices in predicting match outcomes. The bookmakers Vulkan Bet, 10Bet, and

Figure 4.4
Accuracy of Opening Prices for Predicting Match Outcomes



This figure presents the RMSE of opening prices for predicting match outcomes for each bookmaker.

GGBET demonstrate the best performance, with the lowest RMSE values. In contrast, Pinnacle, BetInAsia, and NordicBet exhibit the highest RMSE values, indicating less accurate predictions. Overall, all bookmakers show similar prediction accuracy based on opening prices, with RMSE values clustering around 0.45. These values suggest that the bookmakers' models make fairly accurate predictions. For context, if the predicted probabilities align exactly with the true class

labels, the RMSE would be 0. In contrast, random predictions (e.g., 50/50 probabilities in binary classification) result in an RMSE of 0.5.

Importantly, these individual-specific forecast errors do not account for differences in the time stamps of opening prices. For instance, some bookmakers may post their initial prices later than others, potentially leading to higher accuracy if their prices benefit from a learning process and thus better capture information relevant to predicting match outcomes.

4.4.2 Price Movements on New Information

Table 4.3 presents the regression results implied by Equation (4.1), where we regress close-to-end returns on open-to-close returns to evaluate the market's response to new information that emerges between the opening and closing stages of betting.

Table 4.3
Predictability of Close to End Returns

	Coef.	Std.Err.	z	P> |z|	[0.025	0.975]
Intercept	-0.050	0.006	-7.989	0.000	-0.063	-0.038
RtrnOpnCls	0.023	0.019	1.188	0.235	-0.015	0.061
TsDur	-0.006	0.003	-1.831	0.067	-0.012	0.000
Compet_Challenger_Men	-0.024	0.008	-3.126	0.002	-0.039	-0.009
Compet_ITF_Men	-0.072	0.008	-8.820	0.000	-0.088	-0.056
Compet_Misc	0.005	0.047	0.104	0.917	-0.088	0.097
Compet_WTA	-0.047	0.009	-5.106	0.000	-0.064	-0.029
Bookies Var	0.000	0.000				
Bookies x RtrnOpnCls Cov	0.000					
RtrnOpnCls Var	0.000					

This table presents the results from the mixed linear regression model of the close-to-end returns on open-to-close returns. The fixed effects slope parameter *RtrnOpnCls* is not statistically different from zero, providing evidence for rational price movements based on new information. We estimate the regression model using maximum likelihood and the optimizer fully converged.

The results suggest that the null hypothesis—that prices move based on information and markets respond rationally—cannot be rejected at conventional significance levels. The 95% confidence interval for the slope of *RtrnOpnCls* ranges from -0.015 to 0.061, suggesting that the closing prices represent the conditional expectation of the match outcome. The constant term (*Intercept*) indicates the estimate for the average bookmaker margin allocated to Player 1, which is expected to be equal to that of Player 2, given that the assignment of players to 1 or 2 is not systematic. In contrast, separating the sample into favorites and longshots is expected to yield a non-uniform allocation, as suggested by the favorite-longshot bias, providing an avenue for further exploration of this topic.

The inter-group variation in the intercepts and slopes (*Bookies Var* and *RtrnOpnCls Var*) is very small, indicating that the results do not differ strongly among bookmakers. These results contradict those of Moskowitz (2021), who identifies a tendency for overreactions in the market based on a different dataset covering various kinds of sports and betting contracts.

4.4.3 Relative Forecast Accuracy of Opening and Closing Prices

Table 4.4 presents the regression results from the AGS test Equation (4.2), which examines the statistical difference between the relative predictive abilities of opening and closing prices in forecasting match outcomes. This test investigates whether price movements from the open to the close correct biased opening prices or are merely noise. Table C.1 in the appendix provides the AGS test results evaluated separately for each bookmaker.

Table 4.4
Relative Forecast Accuracy of Opening and Closing Prices

	Coef.	Std.Err.	z	P> |z|	[0.025	0.975]
Intercept	0.003	0.000	5.376	0.000	0.002	0.003
Exog	-0.004	0.001	-7.062	0.000	-0.005	-0.003
TsDur	0.001	0.000	4.015	0.000	0.000	0.001
Compet_Challenger_Men	0.002	0.000	4.385	0.000	0.001	0.003
Compet_ITF_Men	0.003	0.000	7.719	0.000	0.003	0.004
Compet_Misc	-0.008	0.003	-3.320	0.001	-0.013	-0.003
Compet_WTA	0.003	0.000	6.168	0.000	0.002	0.004
Bookies Var	0.000	0.000				
Bookies x Exog Cov	-0.000	0.000				
Exog Var	0.000	0.000				

This table presents the results from the mixed linear regression model implied by the AGS test to estimate the predictive abilities of the opening and closing price forecasts on match outcomes. We estimate the regression model using restricted maximum likelihood and the optimizer fully converged. The intercept β_0 (*Intercept*) indicates a smaller mean forecast error for closing prices, while the slope β_1 (*Exog*) suggests a larger forecast error variance for closing prices.

The results indicate that the mean forecast error of the closing prices is significantly smaller than the mean forecast errors of the opening prices because β_0 (*Intercept*) is significantly larger than zero, with confidence intervals ranging from 0.002 to 0.003. However, the slope β_1 (*Exog*) is significantly smaller than zero, with confidence intervals ranging from -0.005 to -0.003, suggesting a larger forecast error variance for closing prices.

These findings suggest that price movements from the open to the close, driven by aggregated bettors, reduce bias in the betting process. However, this process also increases variance. The inter-group variation in the intercepts and slopes (*Bookies Var* and *Exog Var*) is very small, indicating that the results do not differ significantly across bookmakers.

4.4.4 Magnitude and Direction of Price Movements

Table 4.5 presents the winning rates at different price change intervals, ordered by the magnitude of the change. The interval with the largest negative price changes from the open to the close, ranging from -1 to -0.15, has an average price change of -0.192 calculated from 1,484 matches. This bin exhibits a winning rate of 0.3369, which is significantly smaller than 0.5. The winning rates monotonically increase with the magnitude of prices changes. For the interval with the largest positive price changes, ranging from 0.15 to 1, with an average change of 0.1927 calculated from 1,554 matches, the winning rate is 0.6345, which is significantly larger than 0.5.

Table 4.5
Winning Rates at Different Price Change Magnitudes

Interval	Avg. Change	Avg. Moves	No. Matches	Winning Rate	Z-statistic	p-value
$]-1,-0.15]$	-0.1920	13	1,484	0.3369	-13.3	0.0000
$]-0.15,-0.12]$	-0.1322	11.6	2,165	0.4125	-8.27	0.0000
$]-0.12,-0.09]$	-0.1028	10.8	5,089	0.4429	-8.20	0.0000
$]-0.09,-0.06]$	-0.0730	9.43	11,422	0.4520	-10.3	0.0000
$]-0.06,-0.03]$	-0.0434	8.31	24,443	0.4780	-6.88	0.0000
$]-0.03,0[$	-0.0153	6.87	41,330	0.4853	-5.99	0.0000
$]0,0.03[$	0.0152	6.89	40,595	0.5279	11.2	0.0000
$[0.03,0.06[$	0.0434	8.37	23,742	0.5462	14.3	0.0000
$[0.06,0.09[$	0.0727	9.63	11,062	0.5629	13.3	0.0000
$[0.09,0.12[$	0.1026	10.7	4,644	0.5896	12.4	0.0000
$[0.12,0.15[$	0.1330	11.5	2,044	0.5866	7.95	0.0000
$[0.15,1[$	0.1927	13.1	1,554	0.6345	11	0.0000

This table presents winning rates for different price changes sorted by magnitudes. Specifically, we illustrate the price change interval, the average change within an interval, the average number of price movements, and the number of matches that fall in each interval. Further, we present the winning rates, the test-statistic testing if the winning rate is equal to 0.5, and the corresponding p-value.

Table 4.6 presents the results of the regression of the winning rates on average price change magnitudes per interval Equation (4.3). The intercept estimate β_0 (*Intercept*) is not significantly different from one-half, suggesting that the unconditional winning rate is around 50% and consistent with the expectation that bets are placed on equally likely outcomes. The slope estimate β_1 (*AvgChange*) is significantly greater than zero, indicating a positive relationship between the magnitude of price changes and actual winning rates.

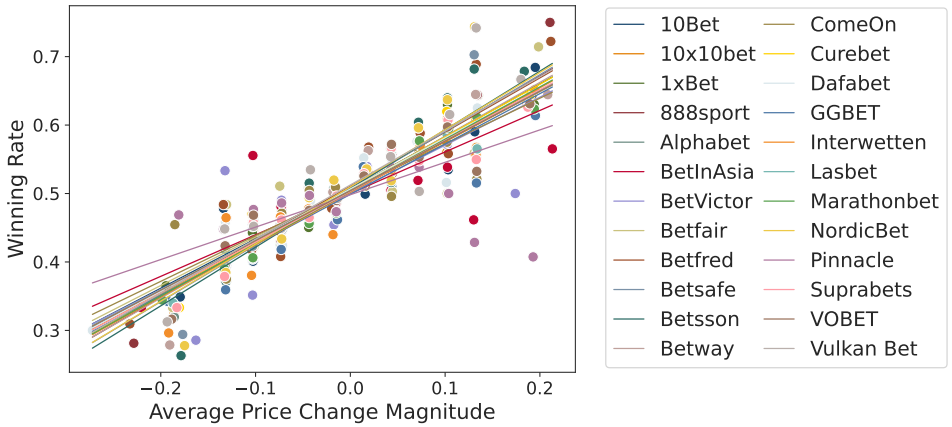
Table 4.6
Relationship Between Winning Rates and Price Change Magnitudes

	Coef.	Std.Err.	z	P> |z|	[0.025	0.975]
Intercept	0.505	0.003	159.179	0.000	0.499	0.511
AvgChange	0.741	0.026	28.106	0.000	0.691	0.792
NumMatches	0.000	0.000	0.478	0.632	-0.000	0.000
Bookies Var	0.000					
Bookies x AvgChange Cov	-0.000					
AvgChange Var	0.006	0.060				

This table presents the results of the mixed linear regression model of actual winning rates on average price change magnitudes per interval. We estimate the model using restricted maximum likelihood and the optimizer fully converged.

Figure 4.5 graphically illustrates the estimated random intercepts and slopes for the bookmakers, highlighting the common intercept around 0.5 and similar, yet slightly varying, slopes. These results imply that larger price changes from the open to the close are associated with higher winning rates for players whose prices increase, and lower winning rates for players whose prices decrease. If prices do not change, winning rates are approximately 0.5. The slight variations in slopes across bookmakers suggest that the price movements of some bookmakers are more informative for explaining match outcomes, with steeper slopes indicating greater explanatory power.

Figure 4.5
Estimated Random Intercepts and Slopes



This figure presents the estimated random intercepts and slopes of the bookmakers from the regression of actual winning rates on average price changes magnitudes from the market open to close.

It appears that the market identifies both under- and overvalued players, as well as the degree of the incorrect assessment. The market corrects this mispricing by forcing the bookmakers to adjust their prices in the appropriate direction and magnitude. These findings indicate that the influence of informed, strategic bettors is substantial and significantly influences market movements while reducing the impact of less informed and casual bettors.

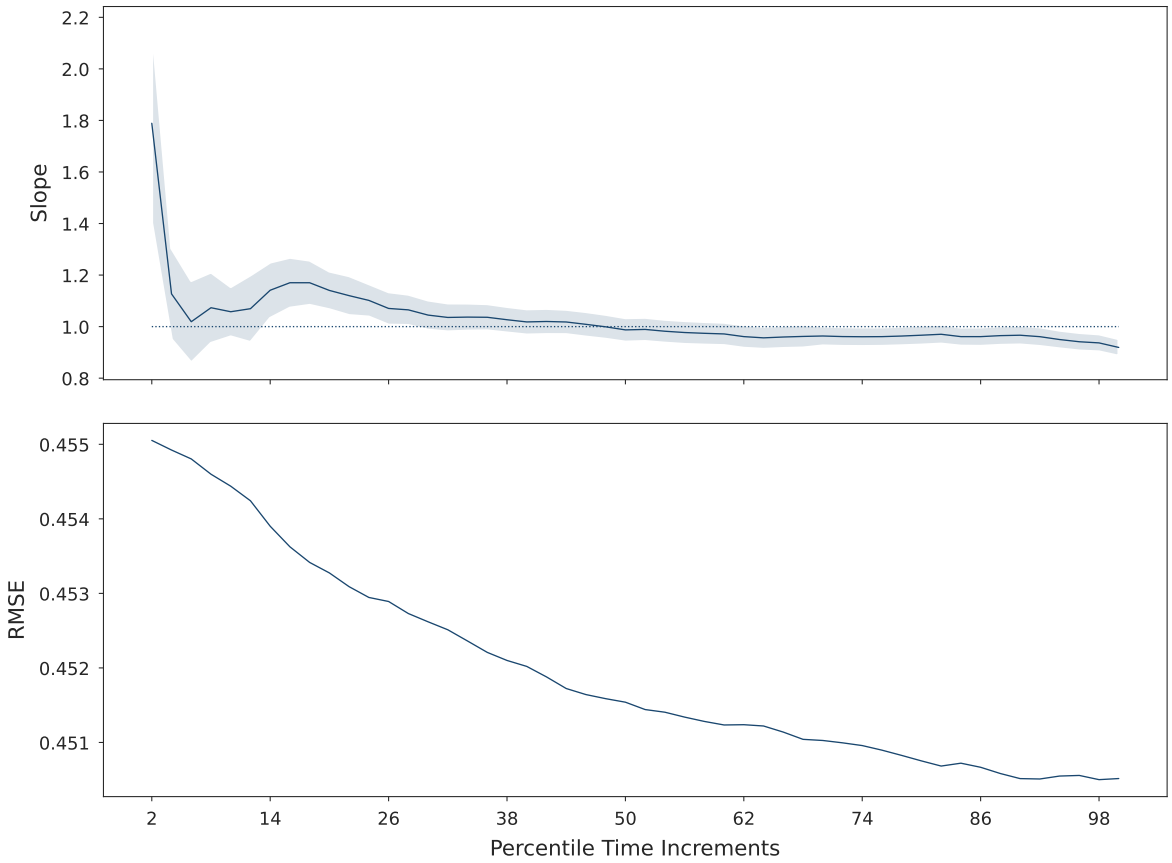
4.4.5 Alternative Hypotheses: Unbiasedness Regressions

So far, we solely focus on opening and closing prices. For the following analyses, we examine the entire price series from opening to closing. This allows us to examine how information is incorporated into prices over time and to identify a learning process. For this analysis, we include only groups with time series featuring fewer than 20 price changes, as Oddsportal provides only the opening price and the last 20 updates before kickoff. For time series with 20 updates, we cannot be certain whether additional price changes occurred between the opening price and the first displayed update.

Figure 4.6 presents the results of the unbiasedness regressions (Equation (4.5)), where we regress the difference in prices from the opening to the end onto the difference from the open to each time t, in order to examine the learning process throughout the betting period. The upper panel of the figure displays the slope estimates, with the blue-shaded area representing the 95% confidence intervals, while the lower panel illustrates the RMSE of the regressions over time.

Right after the open, the slope coefficients are significantly larger than 1, suggesting that there is neither evidence for learning, nor for noise. At this time, intermediate prices start improving forecast accuracy but still exhibit bias, such that the learning hypothesis remains rejected. Then, the slope starts to approximate 1. At the 4th percentile time increment, the confidence bounds begin overlapping with the value of 1, providing evidence of learning up to the 12th percentile.

Figure 4.6
Unbiasedness Regression Results



This figure presents the results of the unbiasedness regressions. The upper panel depicts the slope estimates over time, with the blue-shaded area displaying the 95% confidence intervals. The lower panel depicts the RMSE for each point in time.

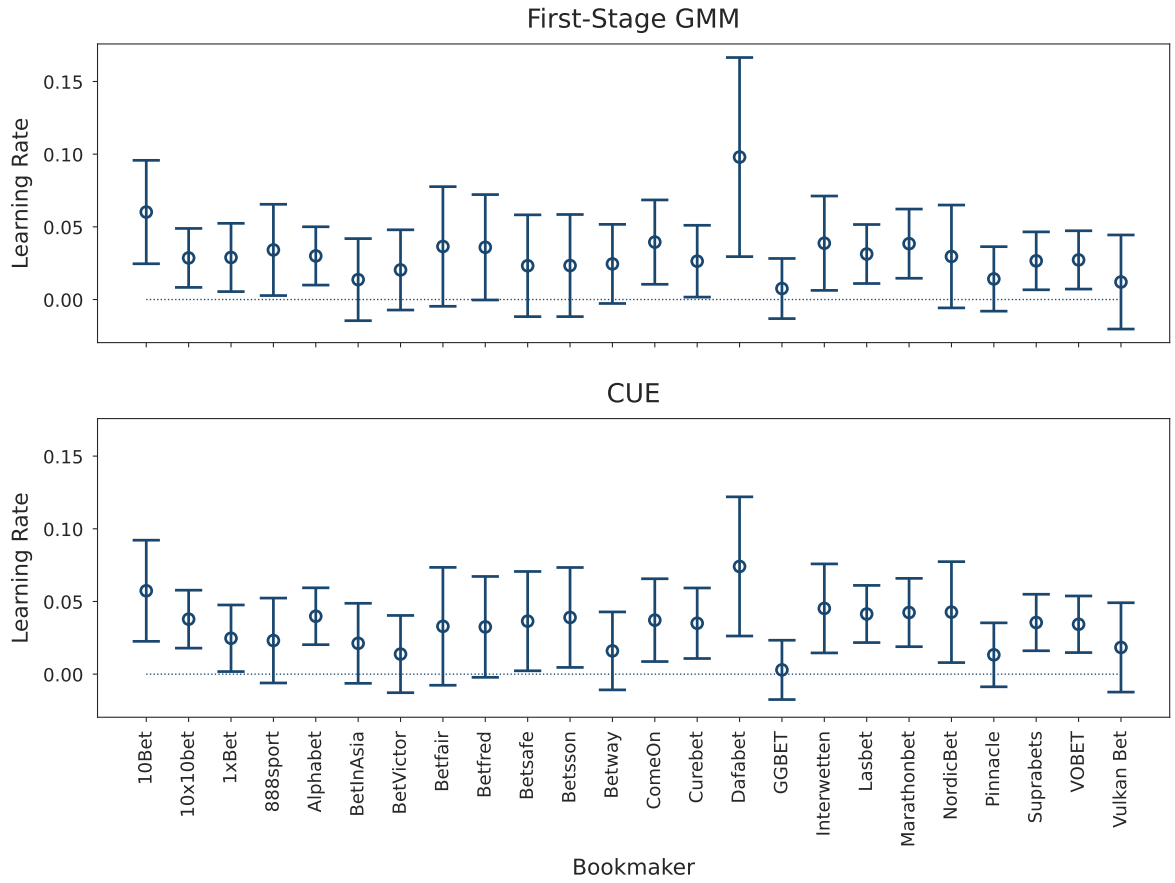
Between the 12th and 30th percentile time increments, the confidence bounds no longer overlap with the value of 1, indicating a pause in learning. However, from the 30th to the 90th percentile, the confidence bounds once again overlap with 1, suggesting a resumption of learning. Just before the closing, the slope coefficients slightly dip below the threshold. Furthermore, the RMSE continuously declines from the open to the close, supporting the learning hypothesis.

4.4.6 Learning Rate

4.4.6.1 Generalized Methods of Moments Estimation

Figure 4.7 presents the bookmaker-specific GMM estimates of the learning rate γ, with the upper panel showing estimates based on first-stage GMM and the lower panel based on the CUE. Figures C.6 and C.7 in the appendix illustrate the J-statistics and the corresponding p-values for first-stage GMM and the CUE, respectively. The estimated learning rates (CUE) for the individual bookmakers range between 0.0029 (GGBET) and 0.0741 (Dafabet), with an average learning rate across bookmakers of 0.0332. The standard error for some bookmakers is relatively large,

Figure 4.7
GMM Parameter Estimates



This figure presents the bookmaker-specific GMM estimates of the learning rate, with the upper panel showing estimates based on first-stage GMM and the lower panel based on the CUE.

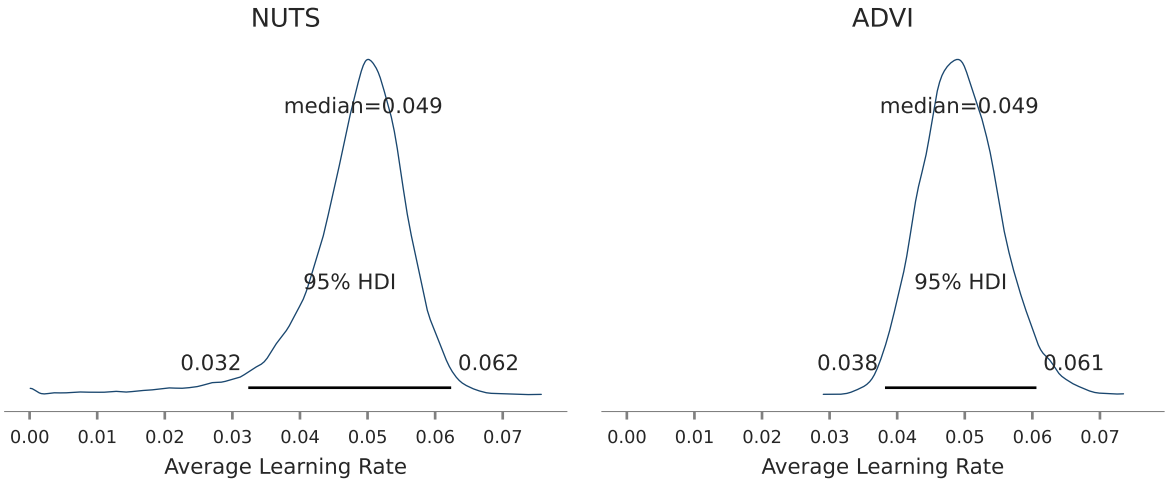
resulting in confidence intervals that overlap with zero.

Our estimated learning rates for the tennis sports betting market are inconsistent with the theoretical prediction of 0.5 by Vives (1995), who examine how quickly private information about the value of a risky financial asset is incorporated into prices, and with the extension to 1.5 by Germain, Meddahi, and Renault (1996), who consider a scenario where agents continuously receive new private signals. Furthermore, Biais, Hillion, and Spatt (1999) find an even higher learning rate of 2.7 with a relatively large standard error of 0.86 during the preopening period of the Paris Bourse. However, these studies are constrained by the lack of a fundamental or terminal value for their assets, requiring the use of proxies—often the last available price—which could introduce bias. Disregarding this limitation, the findings suggest that financial markets may process information significantly faster than sports betting markets, likely because they are more strongly driven by institutional traders, such as large firms, banks, and hedge funds, which diminishes the cumulative impact of noise traders. However, direct comparisons remain challenging due to the differing contexts and settings of the two markets.

4.4.6.2 *Bayesian Estimation*

Figure 4.8 presents the posterior distributions for the overall average learning rate across bookmakers μ_γ obtained from the estimation procedures NUTS (left panel) and ADVI (right panel).

Figure 4.8
Posterior for the Average Learning Rate



This figure presents the posterior distributions for the average learning rate μ_γ obtained from NUTS (left panel) and ADVI (right panel). The black bar indicates the 95% highest density intervals (HDI).

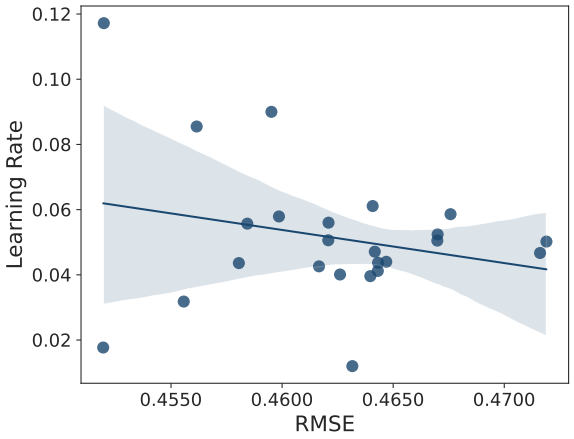
The posterior distributions closely resemble each other and align with the estimates obtained from GMM, supporting the validity of these results. Based on NUTS, the median is 0.0494, and the 95% highest density intervals (HDI) range from 0.0324 to 0.0624. We monitor convergence and stability using trace and density plots for NUTS (Figures C.8 and C.9 in the appendix) and track the mean, standard deviation, and ELBO for ADVI (Figure C.10 in the appendix).

Figure C.11 in the appendix displays the bookmaker-specific posterior distributions for the learning rate γ, obtained from NUTS, while Figure C.12 in the appendix depicts them based on ADVI. In line with the GMM results, the bookmaker-specific posterior distributions indicate significant heterogeneity, indicating that the individual bookmakers process information at different rates. We find a negative correlation (-0.2414) between bookmaker-specific learning rates and their RMSE for predicting match outcomes based on opening prices. This suggests that bookmakers with more accurate opening prices tend to have higher learning rates, indicating they also adapt more quickly to information released by the market between the open and the close of the market.

Figure 4.9 visually demonstrates this relationship in a regression plot, where we regress these individual learning rates on the RMSE. It is important to emphasize that we do not intend to suggest a structural relationship between these two metrics, as such an analysis would exceed the scope of this study. Figure 4.10 illustrates the average learning rates' posterior distributions for favorites and longshots. The orange vertical bar displays the median learning rate estimated

Figure 4.9

Regression Plot Learning Rates and RMSE

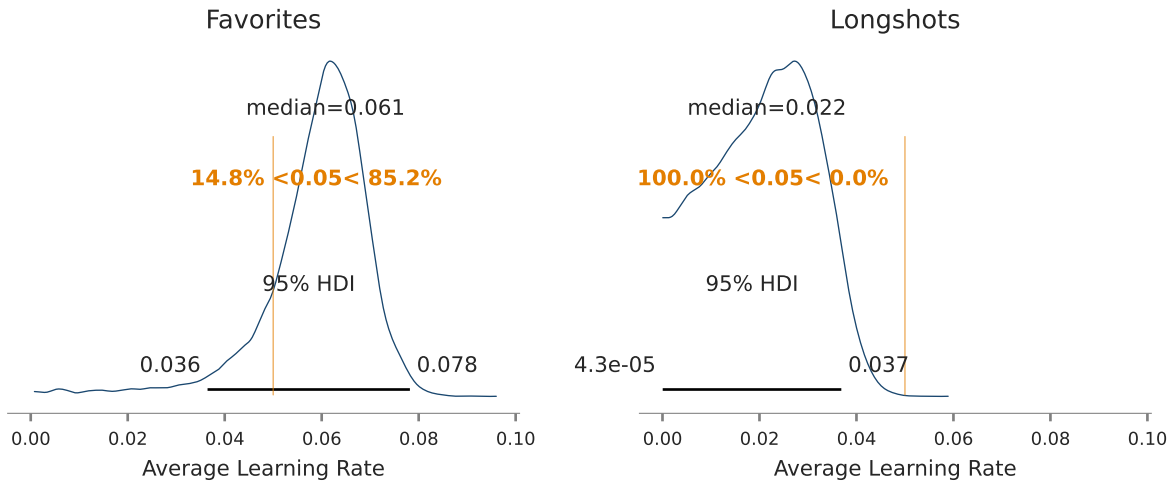


This figure presents the relationship between bookmaker-specific learning rates and their RMSE for predicting match outcomes based on opening prices.

based on the total sample (NUTS) and shows the percentage below and above this reference value. The results reveal a distinct difference in learning rates between these two groups, with favorites showing a significantly higher median learning rate (0.0608) compared to longshots (0.0215). The difference may be attributed to the favorite-longshot bias, where bettors inherently prefer long-shot bets and overestimate the chances of longshots. This behavioral tendency can slow the adjustment of prices for longshots compared to favorites.

Figure 4.10

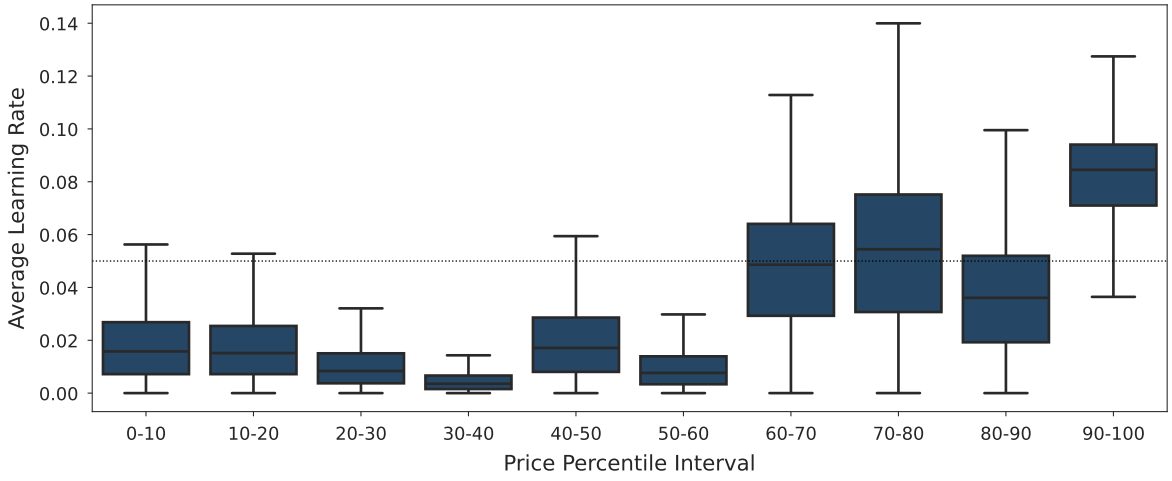
Posteriors for Learning Rates of Favorites vs. Longshots



This figure shows the posterior distributions for the learning rates of favorites (left panel) and longshots (right panel). The horizontal black bar indicates the 95% highest density intervals (HDI), while the vertical orange bar shows the reference value represented by the median learning rate estimated from the total sample. Additionally, the percentage below and above this reference value is displayed.

Figure 4.11 presents a further refinement of this analysis by dividing the matches into ten opening price percentile intervals: 0-10, 10-20, ..., 90-100, with the dotted line showing the reference value represented by the median learning rate estimated from the total sample (NUTS).

Figure 4.11
Posteriors for Average Learning Rates Across Different Price Percentile Intervals



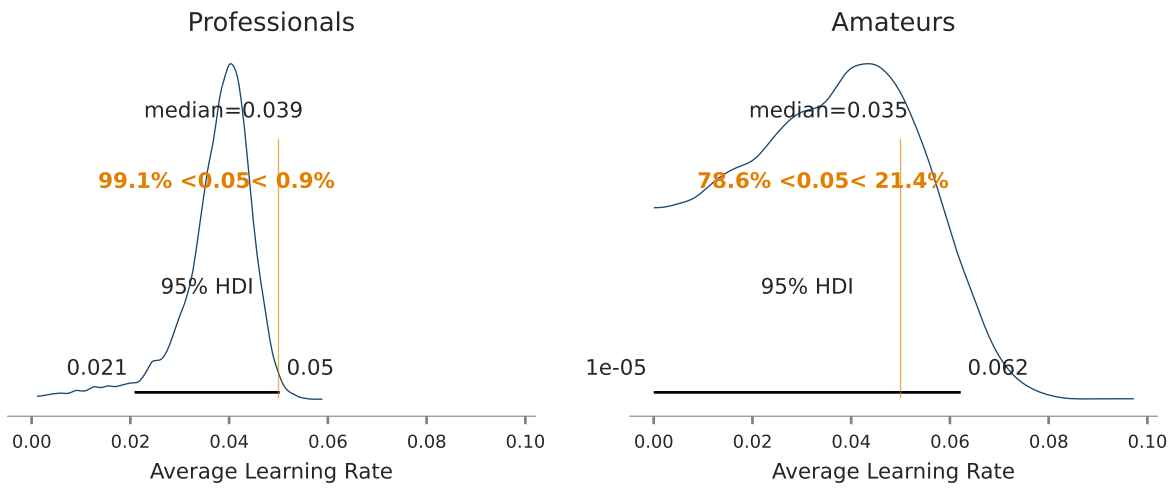
This figure presents the posterior distributions of learning rates across ten opening price percentile intervals: 0-10, 10-20, ..., 90-100. The dotted line shows the reference value represented by the median learning rate estimated from the total sample.

The results indicate that learning rates are very low and similar across the lower price percentile intervals (0-10 to 50-60), which rather represent the longshots. However, beyond these percentiles (i.e., those rather representing favorites), learning rates increase significantly, with the highest rates observed in the top percentile interval (90-100).

Learning rates appear to vary across different price ranges, with significantly lower rates observed for longshots, regardless of the extent of their longshot status. In contrast, for favorites, the intensity of their status appears to have a more pronounced impact, with stronger favorites tending to exhibit higher learning rates.

Finally, Figure 4.12 displays the posterior distributions for the average learning rates obtained from separate estimations for the two competition levels: professionals and amateurs. The median learning rate for professionals (0.0389) is slightly higher than that for amateurs (0.0354), who exhibit greater dispersion. This suggests that bookmakers may integrate information more quickly and precisely in professional-level matches, which could be attributed to the higher unpredictability, lower availability of information, and smaller betting volumes in amateur matches. Another reason for the deviating learning rates might be a different share of rational bettors in each market.

Figure 4.12
Posterior for Average Learning Rates by Competition Level



This figure presents the bookmaker-specific posteriors for μ_γ using professionals (left panel) and amateurs (right panel). The horizontal black bar indicates the 95% highest density intervals (HDI), while the vertical orange bar shows the reference value represented by the median learning rate estimated from the total sample. Additionally, the percentage below and above this reference value is displayed.

4.5 Discussion

The findings of this study provide a comprehensive view of the learning dynamics within the tennis sports betting market, revealing how prices evolve from opening to closing and the factors that influence this process. Learning, in this context, refers to how the market assimilates new information over time, gradually correcting initial biases and moving towards more accurate pricing.

One of the central contributions of this study is the detailed examination of pre-match learning throughout the betting period using longitudinal data. The results indicate that while prices at market opening are often noisy and biased, they become more accurate as the market approaches closing. By the closing, the market identifies and corrects both undervalued and overvalued players, appropriately adjusting for the magnitude of mispricing. This pattern suggests that the market undergoes a significant learning process, driven primarily by the actions of rational bettors who identify and exploit mispricings. These bettors force bookmakers to adjust their prices, thereby reducing biases and improving the accuracy of closing prices as unbiased estimates of match outcomes.

The rate of learning observed in this market is slower compared to traditional financial markets, likely due to the nature of the sports betting market, where information flow is less continuous and more event-driven. In contrast, financial markets are more heavily influenced by institutional traders, such as large firms, banks, and hedge funds, which contributes to faster information processing and a reduced impact from noise traders. However, direct comparisons

are challenging. Learning rates in financial markets are often estimated using proxies, such as the last available price, due to the absence of a clear fundamental or terminal value for assets, which can introduce bias.

An important aspect of the learning process is the heterogeneity in learning rates across different market segments. The results show that favorites exhibit significantly higher learning rates compared to longshots, suggesting that the market assimilates information about favorites more quickly, likely due to the higher volume of stakes and greater attention these players receive. In contrast, longshots, experience much slower learning rates, indicating that the market is less responsive to new information about these players.

Taking a more granular perspective, we find that these varying learning rates differ not only between favorites and longshots but also across different price ranges. Learning rates are significantly lower for longshots, as indicated by their lower prices, irrespective of the degree of their longshot status. Conversely, for favorites, the intensity of their status (i.e., the level of prices) has a more pronounced effect on learning rates, with stronger favorites tending to exhibit higher learning rates.

Finally, the observed differences in learning also extends to the competition among bookmakers. Bookmakers operating in more competitive environments tend to adjust their prices more rapidly, reflecting faster learning processes. This variation suggests that the market's ability to learn and correct mispricings also depends on factors such as the level of competition.

For market participants, the implications of this research are manifold. For bookmakers, understanding the learning dynamics in the market can offer strategic advantages. For instance, reducing the implementation of large corrections as well as improving their risk management and their profitability. For bettors, the insights can enhance their betting strategies, enabling them to make more informed decisions.

However, we also acknowledge certain limitations. The focus on the tennis betting market, while providing a clear and controlled setting, may limit the generalizability of the findings to other markets with different structures and patterns. While our study focuses on the tennis betting market, which offers a controlled and individualized setting, it is important to reflect on whether these findings can be generalized to other sports, particularly team sports like soccer, or even broader financial markets. While generalization is possible and expected to some extent, it ultimately depends on various underlying conditions.

A key factor influencing the generalizability to other sports is popularity. Soccer, for instance, is far more globally popular than tennis, drawing a larger and more diverse audience of both casual and professional bettors. This diversity could lead to different patterns of learning and information processing. For higher proportions of rational bettors, we expect bookmakers to adjust their opening prices more rapidly and accurately, resulting in faster convergence between opening and closing prices compared to tennis.

Another critical driver is the role of chance in determining match outcomes. Tennis is an individual sport, where the outcome largely depends on the skill level and performance of two

players. There are fewer external factors that can affect the match, leading to a more predictable outcome based on rankings, form, and head-to-head records. While chance still plays a role (e.g., injuries, unforced errors), the better player tends to win more often.

In contrast, team sports like soccer inherently introduces more variability, where a multitude of variables—team strategies, lineup, weather conditions, and more—can affect outcomes. In low-scoring sports like soccer, where the outcome is more influenced by chance, we may observe lower learning rates, as bookmakers are less able to rely on past patterns to adjust their prices. This increased uncertainty might lead to divergences from the patterns observed in tennis. Future research could expand this analysis to include other betting formats or team sports to validate and extend the insights gained.

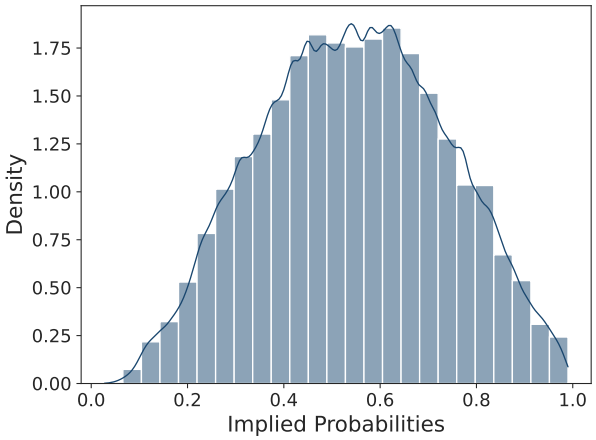
4.6 Summary of Findings

This study examines price formation and learning dynamics in the pre-match phase of the tennis sports betting market using longitudinal data from various bookmakers. The findings reveal that the market undergoes a significant learning process from opening to closing, driven by the actions of rational bettors and influenced by competition among bookmakers. Learning emerges not immediately at the beginning of the betting period but increases toward the middle, continuing almost until closing. Learning rates show substantial variation across bookmakers and market segments. This heterogeneity suggests that the market's ability to assimilate information and correct mispricings varies depending on the level of competition, match prominence, and price ranges. While learning occurs more slowly in this market compared to traditional financial markets, the mechanisms by which prices adjust and biases are corrected are nonetheless effective. These insights not only enhance our understanding of price formation in sports betting but also offer broader implications for the study of learning and information assimilation in other markets.

C Appendix

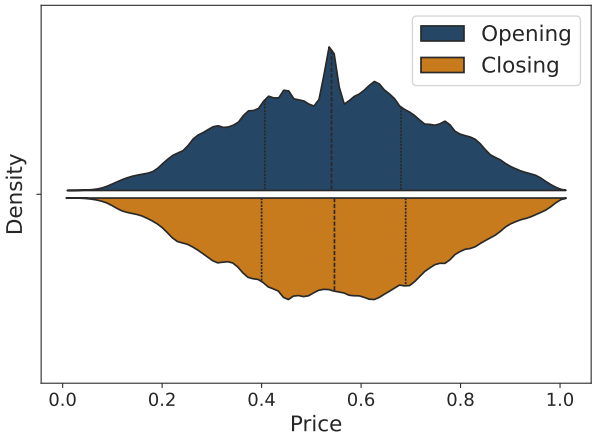
C.1 Figures

Figure C.1
Distribution of Implied Probabilities



This figure presents the distribution of implied probabilities across all matches, bookmakers, and time points.

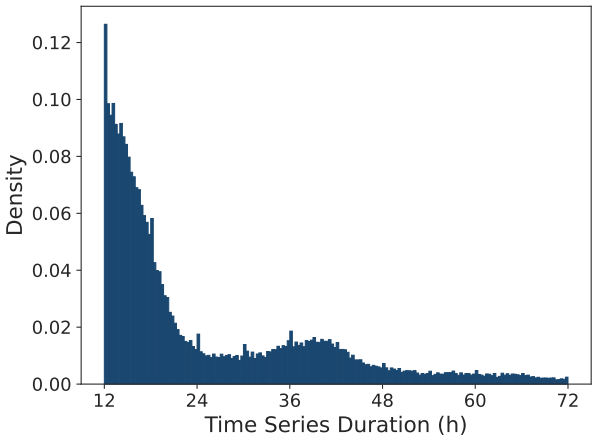
Figure C.2
Distribution of Opening and Closing Implied Probabilities



This figure presents the distribution of opening (blue) and closing (orange) prices across matches and bookmakers. The dashed lines indicate the first quartile, median, and third quartile, respectively.

Figure C.3

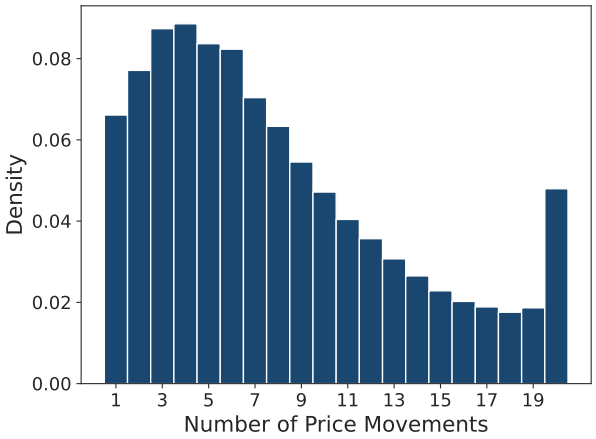
Distribution of Betting Periods (Open to Close)



This figure presents the distribution of the betting periods from open to close based on all matches and bookmakers. Note that we only consider periods between 12 and 72 hours.

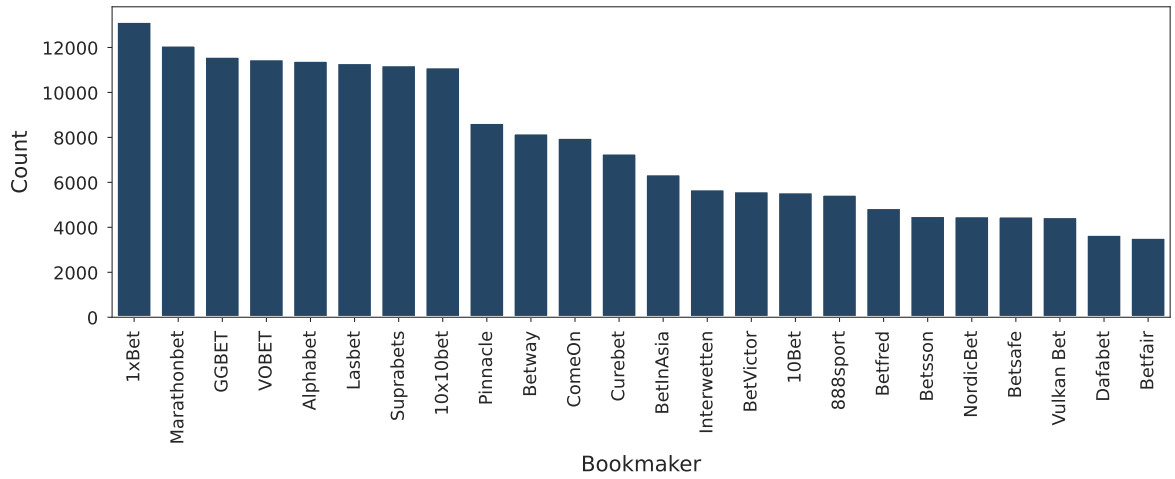
Figure C.4

Distribution of Number of Price Movements



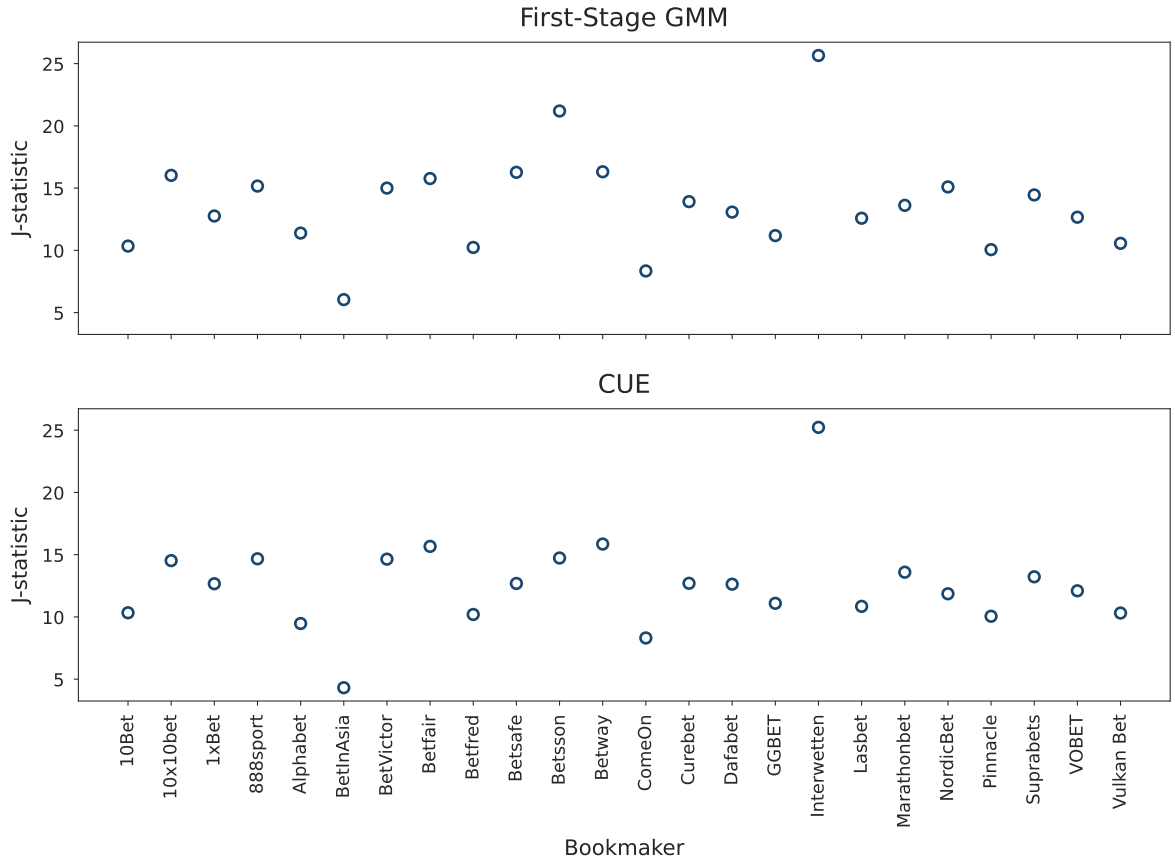
This figure presents the number of price changes per match and bookmaker. Note that Oddsportal provides only the opening prices and the last 20 updates before kickoff. Hence, there are at most 21 observations per bookmaker. The limitation of 20 price changes at most explains the heaping at 20.

Figure C.5
Number of Matches Offered by Each Bookmaker



This figure presents the number of matches offered by each bookmaker, indicating how frequently each bookmaker provides odds.

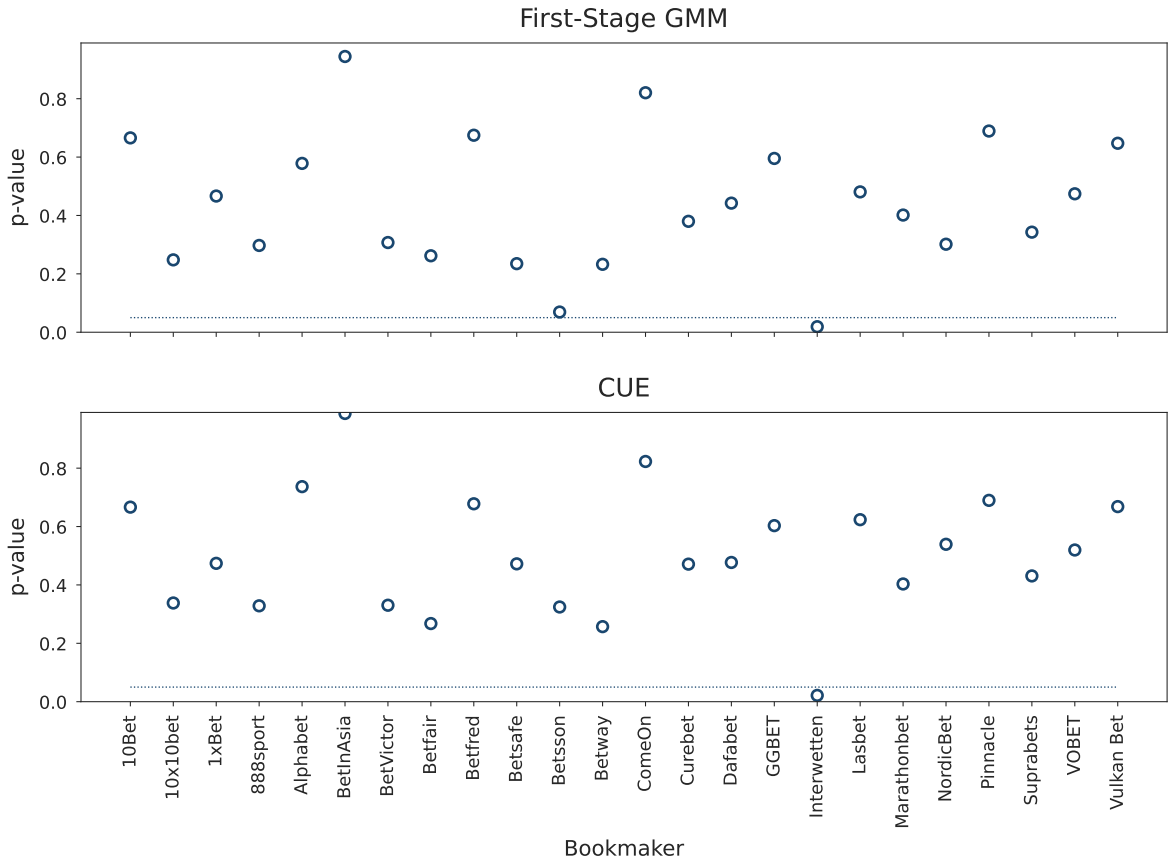
Figure C.6
Bookmaker-Specific J-Statistics



This figure presents the bookmaker-specific J-statisitc from first-stage GMM (upper panel) and the CUE (lower panel).

Figure C.7

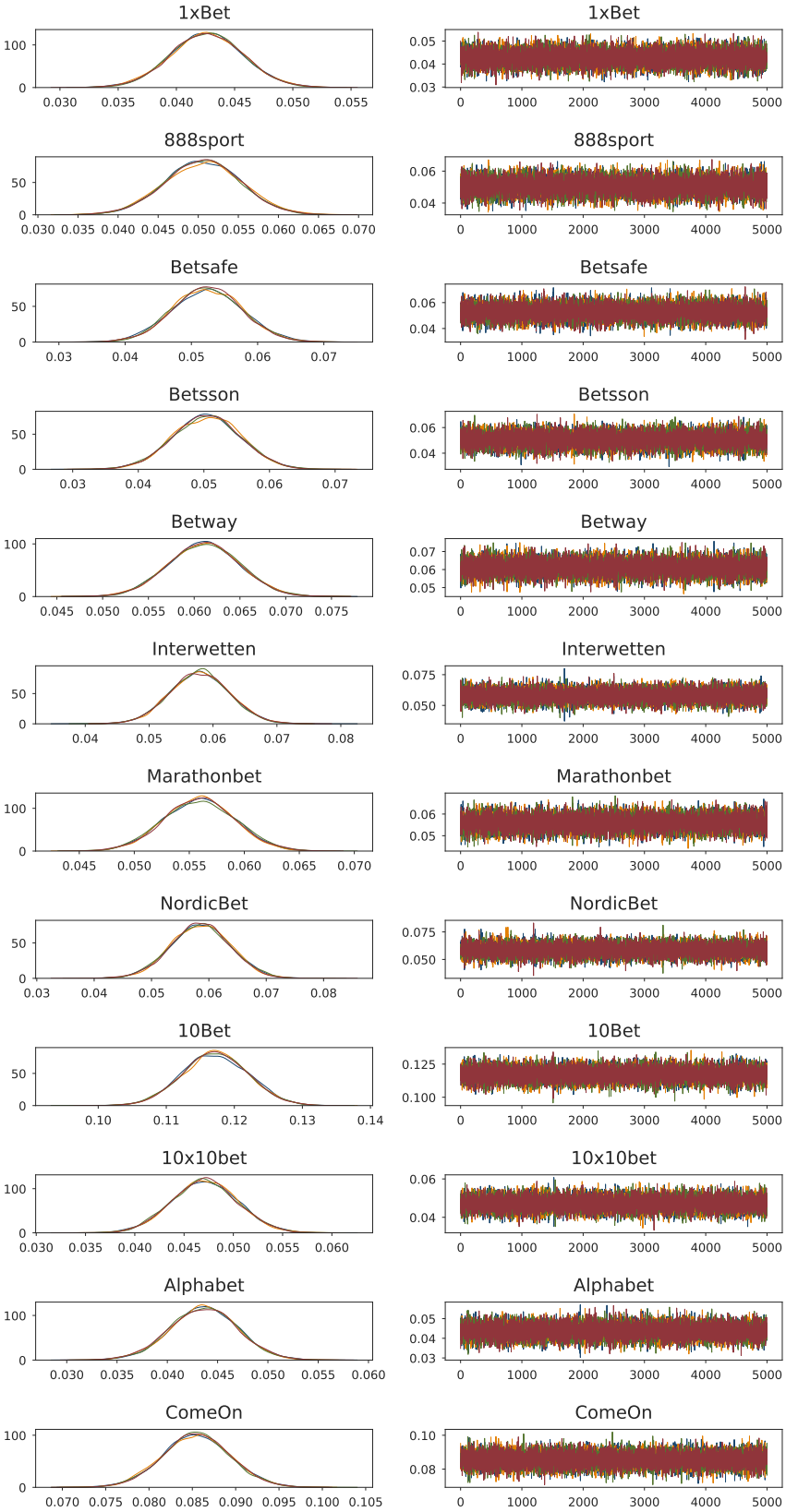
Bookmaker-Specific p-Values of J-Statistics



This figure presents the p-values of the individual J-statisitcs from first-stage GMM (upper panel) and the CUE (lower panel). The dotted line intersects the y-axis at 0.05 to highlight the rejection area. For all bookmakers except Interwetten, the model (moment conditions) cannot be rejected at the 5% significance level.

Figure C.8

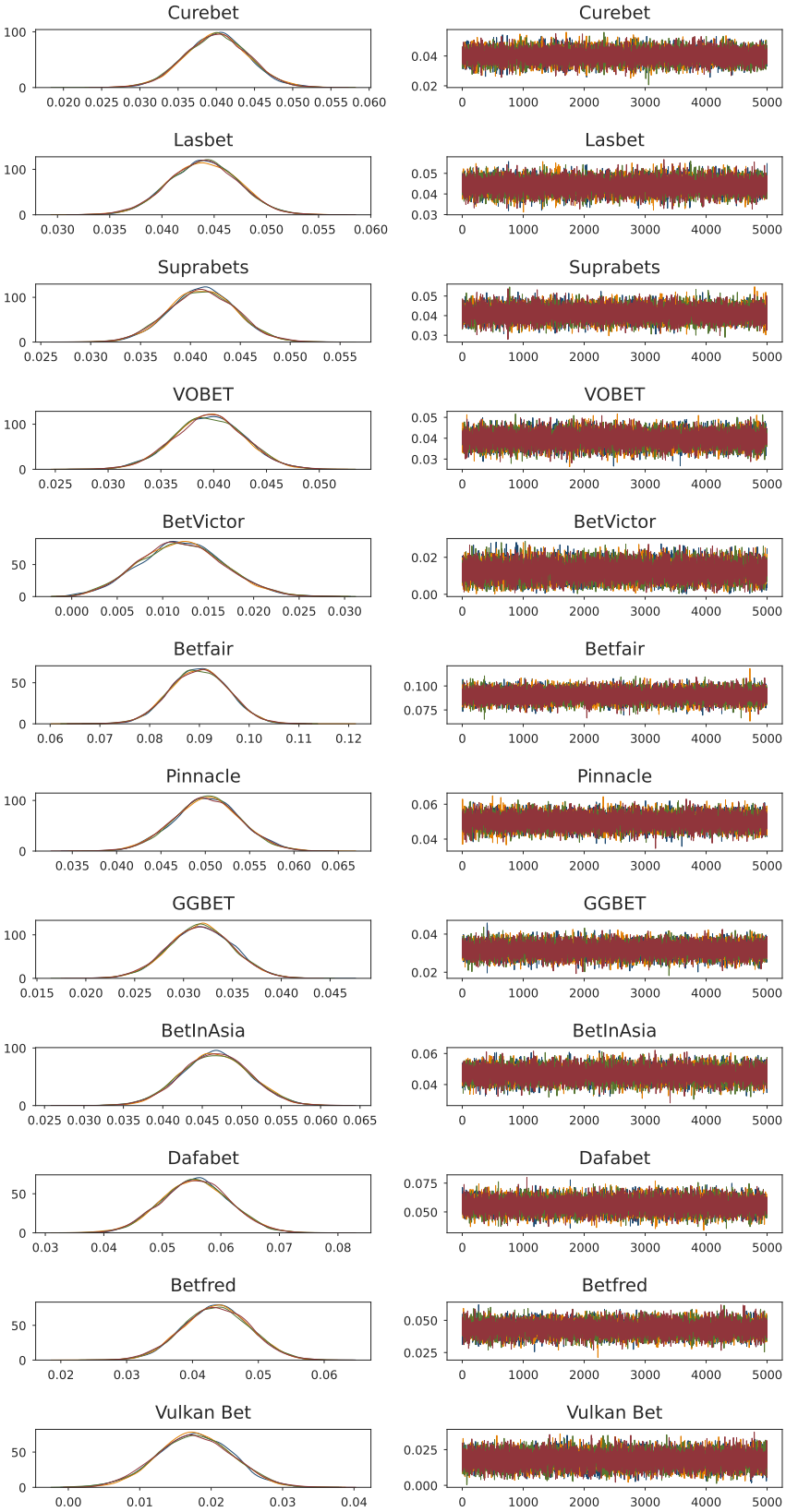
Bookmaker-Specific Densities and Traces (I)



This figure presents the first half of the bookmaker-specific posterior distributions (left panel) and traces (right panel) to monitor convergence and stability. Each color represents an individual chain.

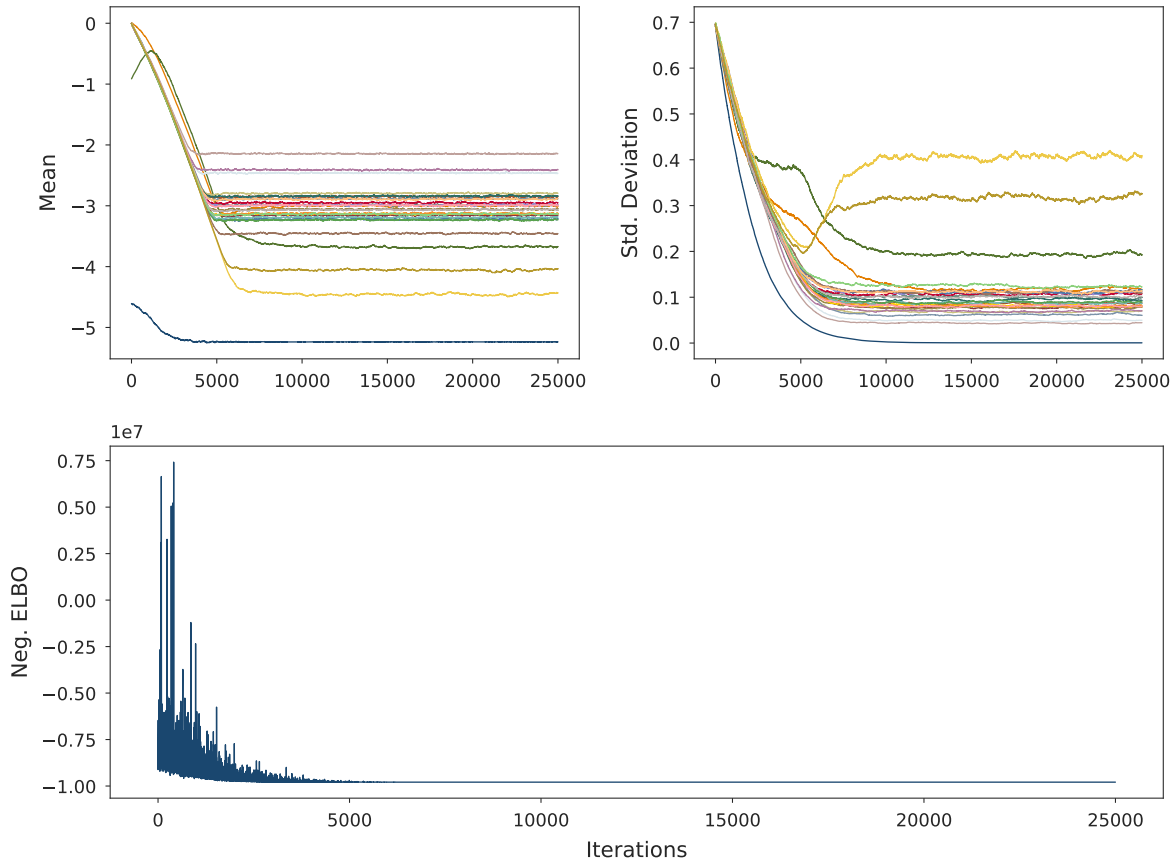
Figure C.9

Bookmaker-Specific Densities and Traces (II)



This figure presents the second half of the bookmaker-specific posterior distributions (left panel) and traces (right panel) to monitor convergence and stability. Each color represents an individual chain.

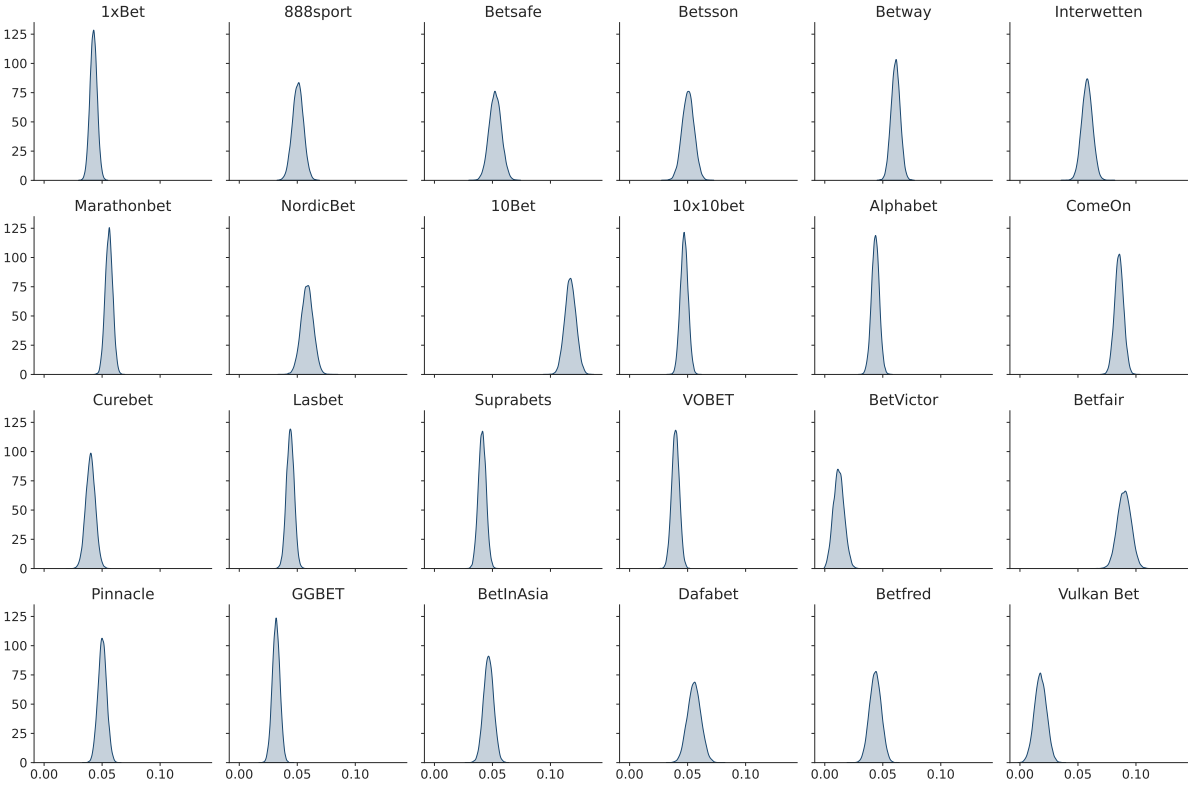
Figure C.10
ADVI Tracker for the Variational Parameters



This figure presents the ADVI tracker tracking the variational parameters to gain insights into the convergence and stability of the variational inference process.

Figure C.11

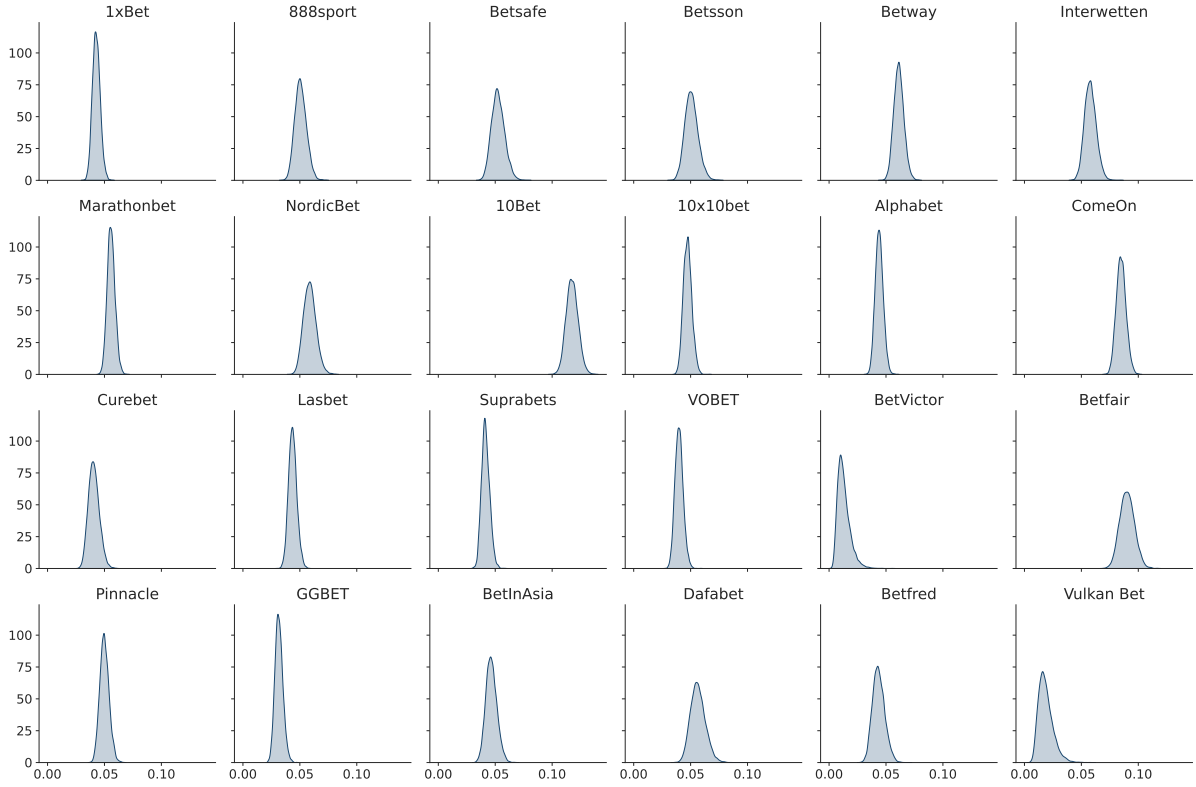
Bookmaker-Specific Posteriors for Learning Rates (NUTS)



This figure presents the bookmaker-specific posteriors for the learning rate obtained from NUTS.

Figure C.12

Bookmaker-Specific Posteriors for Learning Rates (ADVI)



This figure presents the bookmaker-specific posteriors for the learning rate obtained from ADVI.

C.2 Tables

Table C.1

Bookmaker-Specific Relative Forecast Accuracy of Opening and Closing Prices

	N	RMSE(e_0)	RMSE(e_T)	β_0	β_1	$p(\beta_0)$	$p(\beta_1)$
10Bet	5,129	0.4499	0.4466	0.0009	0.0020	0.5732	0.0040
10x10bet	10,238	0.4557	0.4529	0.0035	-0.0042	0.0103	0.0000
1xBet	12,515	0.4501	0.4462	0.0038	-0.0064	0.0150	0.0000
888sport	4,792	0.4605	0.4579	0.0018	-0.0067	0.2750	0.0000
Alphabet	10,647	0.4562	0.4537	0.0034	-0.0044	0.0098	0.0000
BetInAsia	6,124	0.4645	0.4630	0.0013	0.0002	0.3838	0.7426
BetVictor	4,789	0.4645	0.4632	0.0025	-0.0050	0.1028	0.0000
Betfair	3,093	0.4589	0.4574	-0.0008	-0.0011	0.6493	0.2040
Betfred	4,083	0.4597	0.4549	0.0048	-0.0069	0.0064	0.0000
Betsafe	4,078	0.4607	0.4580	0.0032	-0.0047	0.0348	0.0000
Betsson	4,072	0.4607	0.4581	0.0034	-0.0048	0.0319	0.0000
Betway	6,787	0.4599	0.4563	0.0022	-0.0035	0.1449	0.0000
ComeOn	7,375	0.4520	0.4490	-0.0006	0.0016	0.6688	0.0055
Curebet	6,775	0.4552	0.4520	0.0041	-0.0043	0.0074	0.0000
Dafabet	3,372	0.4547	0.4526	0.0048	-0.0029	0.0050	0.0000
GGBET	10,940	0.4492	0.4466	0.0010	-0.0044	0.4614	0.0000
Interwetten	4,928	0.4590	0.4546	0.0037	-0.0051	0.1275	0.0000
Lasbet	10,550	0.4560	0.4534	0.0031	-0.0045	0.0194	0.0000
Marathonbet	11,496	0.4511	0.4479	0.0055	-0.0070	0.0002	0.0000
NordicBet	4,077	0.4617	0.4591	0.0036	-0.0048	0.0206	0.0000
Pinnacle	8,357	0.4656	0.4640	0.0011	0.0009	0.3362	0.0639
Suprabets	10,425	0.4565	0.4537	0.0025	-0.0041	0.0558	0.0000
VOBET	10,711	0.4553	0.4529	0.0029	-0.0043	0.0267	0.0000
Vulkan Bet	4,221	0.4508	0.4494	-0.0014	-0.0016	0.2786	0.0093
All	169,574	0.4561	0.4534	0.0025	-0.0036	0.0000	0.0000

This table presents the results from the mixed linear regression model implied by the AGS test for each bookmaker, separately, to estimate the predictive abilities of the opening and closing price forecasts for match outcomes. We estimate the model using restricted maximum likelihood and the optimizer fully converged.

Table C.2
Results of Bayesian Estimation of the Learning Rate

	mean	std_mean	median	std_median	$\text{HDI}^{2.5\%}$	$\text{HDI}^{97.5\%}$	$\hat{R}$
$\hat{\mu}_\gamma$	0.0482	0.0081	0.0494	0.0082	0.0324	0.0624	1.00
$\hat{\sigma}_\gamma$	0.0268	0.0071	0.0253	0.0072	0.0161	0.0407	1.00
$\hat{\gamma}_{1xBet}$	0.0426	0.0031	0.0426	0.0031	0.0365	0.0486	1.00
$\hat{\gamma}_{888sport}$	0.0506	0.0047	0.0506	0.0047	0.0414	0.0598	1.00
$\hat{\gamma}_{Betsafe}$	0.0525	0.0052	0.0524	0.0052	0.0425	0.0629	1.00
$\hat{\gamma}_{Betsson}$	0.0504	0.0051	0.0505	0.0051	0.0406	0.0607	1.00
$\hat{\gamma}_{Betway}$	0.0611	0.0039	0.0611	0.0039	0.0534	0.0685	1.00
$\hat{\gamma}_{Interwetten}$	0.0579	0.0046	0.0579	0.0046	0.0490	0.0670	1.00
$\hat{\gamma}_{Marathonbet}$	0.0559	0.0032	0.0560	0.0032	0.0498	0.0624	1.00
$\hat{\gamma}_{NordicBet}$	0.0585	0.0052	0.0586	0.0052	0.0482	0.0683	1.00
$\hat{\gamma}_{10Bet}$	0.1172	0.0048	0.1172	0.0048	0.1079	0.1265	1.00
$\hat{\gamma}_{10x10bet}$	0.0471	0.0033	0.0471	0.0033	0.0407	0.0537	1.00
$\hat{\gamma}_{Alphabet}$	0.0437	0.0033	0.0437	0.0033	0.0372	0.0504	1.00
$\hat{\gamma}_{ComeOn}$	0.0855	0.0039	0.0855	0.0039	0.0780	0.0932	1.00
$\hat{\gamma}_{Curebet}$	0.0400	0.0041	0.0401	0.0041	0.0320	0.0482	1.00
$\hat{\gamma}_{Lasbet}$	0.0440	0.0033	0.0440	0.0033	0.0375	0.0503	1.00
$\hat{\gamma}_{Suprabets}$	0.0412	0.0033	0.0412	0.0033	0.0347	0.0478	1.00
$\hat{\gamma}_{VOBET}$	0.0396	0.0033	0.0396	0.0033	0.0330	0.0459	1.00
$\hat{\gamma}_{BetVictor}$	0.0121	0.0046	0.0120	0.0046	0.0032	0.0210	1.00
$\hat{\gamma}_{Betfair}$	0.0900	0.0058	0.0900	0.0058	0.0785	0.1009	1.00
$\hat{\gamma}_{Pinnacle}$	0.0502	0.0037	0.0502	0.0037	0.0429	0.0575	1.00
$\hat{\gamma}_{GGBET}$	0.0318	0.0033	0.0318	0.0033	0.0255	0.0383	1.00
$\hat{\gamma}_{BetInAsia}$	0.0467	0.0043	0.0467	0.0043	0.0384	0.0551	1.00
$\hat{\gamma}_{Dafabet}$	0.0557	0.0058	0.0557	0.0058	0.0445	0.0669	1.00
$\hat{\gamma}_{Betfred}$	0.0436	0.0050	0.0436	0.0050	0.0339	0.0533	1.00
$\hat{\gamma}_{VulkanBet}$	0.0177	0.0052	0.0177	0.0052	0.0079	0.0282	1.00
$\hat{\sigma}$	0.0053	0.0000	0.0053	0.0000	0.0053	0.0053	1.00

This table presents the results from the Bayesian estimation of the learning rate parameters based on NUTS.

C.3 Imputation of Initial Prices

For each match, we observe multiple bookmakers offering odds on the outcome, which we convert to implied probabilities, or prices. However, not all bookmakers provide their initial prices at the same time, resulting in different opening price timestamps and thus time series of varying lengths. While the end of the time series is defined by the kickoff time and is consistent across all bookmakers, the start varies. To create homogenous time series, we align the beginning with the earliest opening price timestamp across all bookmakers for that match. This process generates missing values for $K - 1$ bookmakers prior to their initial price offering.

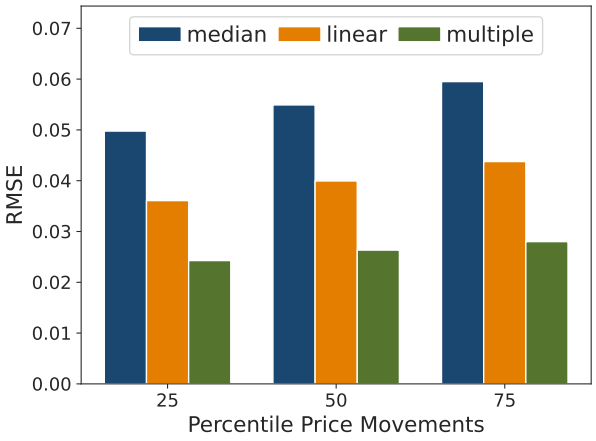
Since a simple backward-filling approach would ignore the dynamic patterns of the data, we impute these missing values—representing a fraction of 0.0784 of the total sample—using a multivariate imputation strategy[25]. The procedure employs a Bayesian Ridge regression model that incorporates prior distributions over the parameters to predict missing values by accounting for

[25]We use scikit-learn's iterative imputer. This strategy assumes that the non-existing values can be consistently estimated based on time series patterns across matches and bookmakers in the dataset. The documentation is available at https://scikit-learn.org/stable/modules/impute.html#iterative-imputer

the relationships between multiple features. These features include the affected bookmaker's implied probabilities for other matches and the implied probabilities of other bookmakers for the same and different matches. In accordance with no-arbitrage conditions, we restrict the algorithm to ensure that the imputed values remain within the range between 0 and 1.

To evaluate the accuracy of the imputed values, we compare the performance of the proposed strategy against two other imputation methods—median imputation and linear extrapolation. We use a simulation design in which a comparable proportion of observations at the start of

Figure C.13
RMSE Implied by Different Imputation Strategies



This figure presents the RMSE implied by different imputation strategies, including median imputation, linear extrapolation, and the proposed multiple imputation.

randomly selected time series are artificially set to None. These missing values are then imputed using each of the imputation strategies. During the simulations, we vary the minimum number of price movements in the affected time series, setting thresholds at the 25th, 50th, and 75th percentiles of the total number of price movements in the dataset. This serves as a proxy for the level of variation within the time series. For each simulation, we compute the RMSE, as shown in Figure C.13.

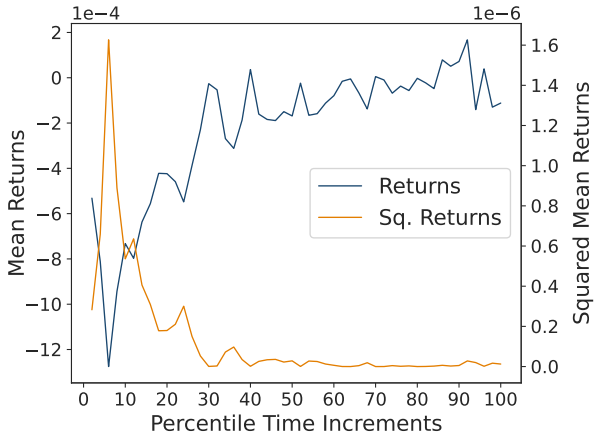
We observe that across all evaluated percentiles, the median imputation strategy consistently produces the highest RMSE, followed by the linear extrapolation strategy. The proposed multiple imputation strategy yields the lowest RMSE for all percentiles.

C.4 Statistical Properties of Time Series Data

To investigate the time series properties of our data, we calculate the returns and the square returns based on the cross-sectional mean from the first observation to the last. This approach avoids the need to test each time series separately and allows us to assess the general temporal pattern while minimizing idiosyncratic effects.

Figure C.14 depicts the cross-sectional mean returns and the squared cross-sectional mean returns from the first percentile time increment to the last. The cross-sectional mean returns

Figure C.14
Cross-sectional Mean Returns and Square Returns Over Time



This figure presents the cross-sectional mean returns and the squared cross-sectional mean returns of the time series. The mean at a specific time t is calculated as the average value across all matches and bookmakers at that particular time point.

exhibit a sharp decline immediately after the opening, followed by a gradual upward trend that declines but persists until the end. The squared cross-sectional mean returns also show high initial variability, which smooths out in the later periods, indicating that the market is stabilizing over time. Using these return series, we perform a variety of time series tests to thoroughly analyze the data's statistical properties.

C.4.1 Augmented Dickey-Fuller Test

First, we test the time series for a unit root by applying an augmented Dickey-Fuller test (Dickey and Fuller 1979), which includes a constant, a linear trend, and a quadratic trend. This involves running the following regression:

$$\Delta y_t = \alpha + \zeta_0 t + \zeta_1 t^2 + \phi y_{t-1} + \sum_{i=1}^{P} \delta_i \Delta y_{t-i} + \varepsilon_t,$$

where y_t represents the cross-sectional mean return at time t, α is a constant term (drift), ζ_0 is the coefficient on the linear time trend, ζ_1 is the coefficient on the quadratic trend, ϕ is the coefficient of the lagged level of the time series (i.e., the process root), δ_j are the coefficients on the lagged differences Δy_{t-i}, which account for higher-order autocorrelation, and P is the number of lagged difference terms included in the model. We determine the number of lags P by using the Bayes-Schwarz information criterion (BIC) as proposed by Schwarz (1978).

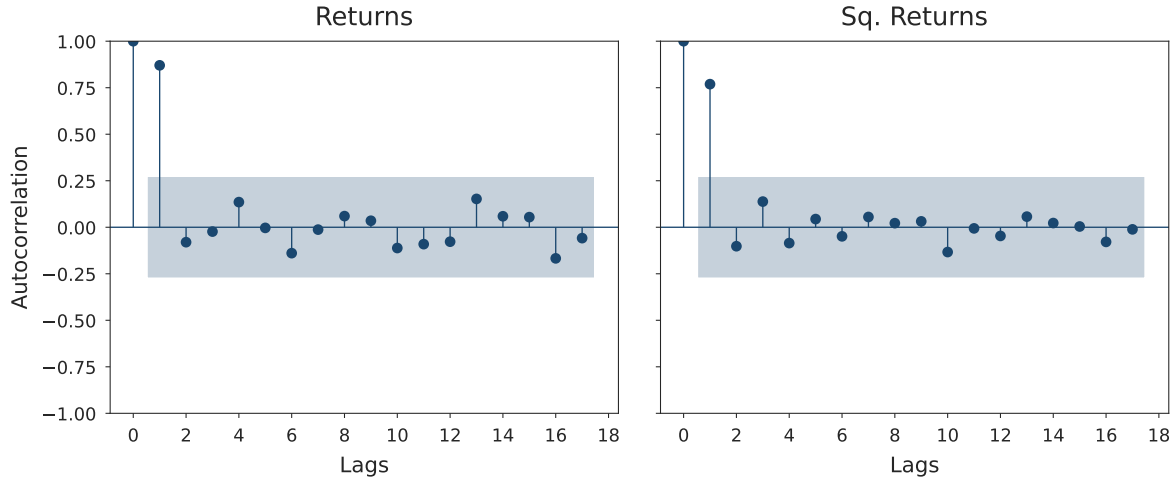
The null hypothesis is that the time series has a unit root, while the alternative hypothesis is that the time series is (weakly) stationary. We obtain a test-statistic value of -5.35 and a corresponding p-value of 0.0002, suggesting that we can reject the null hypothesis that the time series contains a unit root on conventional significance levels.

C.4.2 Partial Autocorrelation Function

Second, we investigate the autocorrelation of the time series using the partial autocorrelation function (PACF). The PACF measures the correlation between two values in the time series while controlling for the influence of intermediate lags, allowing us to identify the direct effect of a past value on the current value.

Figure C.15 illustrates the PACF of the cross-sectional return series (left panel) and squared return series (right panel) for multiple lags. The PACF exhibits a significant effect of past values

Figure C.15
Partial Autocorrelation Function



This figure depicts the PACF of the cross-sectional return series (left panel) and squared return series (right panel) for multiple lags. The blue-shaded area displays the 95% confidence intervals.

on current values up to the first lag for both time series. Moreover, the effect of past values decays over time and becomes non-significant after the first lag, indicating a low degree of persistence in the data.

C.4.3 EGARCH Model

Third, we examine the volatility pattern of the squared return time series. Similar to asset returns in financial markets, price returns in the sports betting market may exhibit time-varying volatility and volatility clustering effects. To estimate the conditional volatility, we employ an exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model, as introduced by Nelson (1991). This model is particularly suited for capturing nonlinearities and asymmetries, allowing it to effectively model scenarios where negative shocks have a different impact on volatility than positive shocks.

For modeling the mean, we use an autoregressive process with a constant term. We determine the number of lags using the significant lags indicated by the PACF for the square returns, which

equals 1. This yields the following EGARCH model:

$$y_t = \alpha + \phi y_{t-1} + \varepsilon_t$$
$$\varepsilon_t = \sigma_t u_t$$
$$\ln\left(\sigma_t^2\right) = \mu + \rho\left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| - \sqrt{2/\pi}\right) + \tau\frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \psi\ln\left(\sigma_{t-1}^2\right),$$

where y_t is the square return series, α is the constant term of the autoregressive process, ϕ is the autoregressive parameter, ε_t the residual composed of the conditional standard deviation (volatility) σ_t and an i.i.d. process u_t with zero mean and unit variance. The intercept μ represents the average long-term volatility, ρ measures the sensitivity of past shocks on current volatility, τ provides insights whether the time series has symmetrical or asymmetrical volatility behavior ($\tau > 0$ indicates that positive shocks bring on increased volatility, while $\tau < 0$ indicates that negative shocks bring on increased volatility), and ψ describes the persistence of volatility over time, with higher values suggesting that past volatility has a lasting effect on current volatility.

Table C.3 presents the estimated parameters of the constant mean EGARCH model. The

Table C.3
Constant Mean EGARCH Model

	coef	std err	t	P> |t|	95.0% Conf. Int.
α	0.0664	1.640e-02	4.049	5.148e-05	[3.426e-02,9.856e-02]
ϕ	0.7111	6.661e-02	10.676	1.315e-26	[0.581, 0.842]
μ	-0.4515	0.162	-2.790	5.277e-03	[-0.769, -0.134]
ρ	0.9344	0.490	1.909	5.628e-02	[-2.502e-02, 1.894]
τ	0.4006	0.188	2.136	3.269e-02	[3.299e-02, 0.768]
ψ	0.8434	5.364e-02	15.724	1.041e-55	[0.738, 0.949]

This table presents the estimated parameters of the EGARCH model. The constant term α represents the predicted average return over the period, while ϕ is the autoregressive parameter estimate. The volatility model intercept μ represents the average long-term volatility, ρ measures the sensitivity of past shocks on current volatility, τ accounts for the asymmetry in the impact of past shocks on volatility, and ψ describes the persistence of volatility over time.

results confirm the presence of conditional heteroscedasticity, suggesting that past volatility has a positive lasting effect on current volatility ($\psi = 0.8434$). Additionally, the results provide evidence for asymmetry in the impact of past shocks on volatility, indicating that positive log returns or shocks may bring on increased volatility ($\tau = 0.4006$).

C.5 Bayesian Estimation - Algorithms and Procedures

The No-U-Turn sampler (NUTS), as proposed by Hoffman and Gelman (2014), is a recursive algorithm for continuous variables based on Hamiltonian mechanics, offering greater computational efficiency compared to other samplers like the Gibbs sampler. We utilize the Rust implementation, Nutpie, which is integrated into the PyMC probabilistic programming package leveraging Numba as its computational backend.

For the NUTS estimations, we increase the acceptance probability slightly to 0.85 from the default 0.8, ensuring smaller step sizes and reducing the likelihood of divergent trajectories in the samples. We run 4 chains with 5,000 iterations each, discarding an additional 2,000 burn-in samples per chain. To monitor convergence, we use the Gelman-Rubin statistic $\hat{R}$ (Gelman and Rubin 1992; Brooks and Gelman 1998) and dispay trace plots for the individual chains.

Automatic differentiation variational inference (ADVI) is a deterministic approximation method that optimizes parameters—typically the mean and standard deviation—by iteratively updating them to minimize the difference between the variational distribution and the posterior. In ADVI, our goal is to find the parameters of the variational distribution that best approximate the posterior distribution by minimizing the Kullback-Leibler (KL) divergence between the true posterior and the variational approximation.

For ADVI estimation, we use 25,000 iterations and 10,000 draws. The number of iterations reflects how many times the optimization algorithm updates the variational parameters during inference, while the number of draws indicates how many samples are taken from the variational distribution during each iteration to estimate the gradients of the objective function. We monitor the convergence of this process using a tracker that records changes in the objective function, such as the evidence lower bound (ELBO).

Chapter 5

Conclusions

This dissertation examines the immediate behavioral responses of the average market participant to new information, providing empirical evidence on consumer behavior and economic decision-making shaped by psychological influences through advanced statistical methods. The present work contributes to the field of behavioral economics by expanding on established theories like Kahneman and Tversky's prospect theory and other key frameworks, which focus on biases such as reference dependence and loss aversion, to enhance our understanding of decision-making processes in economics, particularly in explaining deviations from rational behavior.

The studies in this dissertation leverage sporting events and sports betting markets as a unique setting where economic behavior can be investigated in a natural, real-world environment, yet with controlled conditions characterized by the presence of terminal values and isolation from broader economic market risks. This combination provides a robust foundation for generalizing findings to other markets. Furthermore, this thesis introduces a novel statistical method for regularization within Bayesian frameworks, addressing the challenges of complex, high-dimensional models with sparse data, which are commonly encountered in sports settings.

The first contribution of this dissertation focuses on emotional drinking, specifically examining how the immediate emotions surprise and suspense, as formulized by Ely, Frankel, and Kamenica (2015), that are experienced during soccer matches influence aggregated alcohol consumption as a mass phenomenon. The research contributes to the behavioral economics literature by providing the first field evidence of real-time emotional reactions linked to alcohol use, offering high external validity.

Alcohol use is deeply embedded in society, recognized as a leading risk factor for death and disability (e.g., Griswold et al. 2018) and a major contributor to criminal behaviors such as family violence (Klostermann and Fals-Stewart 2006). However, while most existing studies rely on lab experiments that inherently cannot induce or control intense emotions (Tymula and Glimcher 2018), field evidence on the immediate effects of emotions on alcohol use as a mass phenomenon is largely missing. Unlike previous studies that analyze post-match behavior (e.g., Edmans, Garcia, and Norli 2007; Eren and Mocan 2018; Lindo, Siminski, and Swensen 2018; Cardazzi et al. 2022) or rely on lab experiments, this study utilizes high-frequency transaction data from nearly 100 matches played by a top-division soccer team in Germany, capturing how emotions directly impact decisions during the match.

The findings reveal that surprise leads to increased beer sales, consistent with psychological theories suggesting that people turn to alcohol in emotionally charged situations (e.g., Cooper et al. 1995; Greeley and Oei 1999). Further, the study reveals differences in how positive and negative emotions influence consumption. Although the effects of positive and negative emotional

states are not significantly different, the findings indicate faster responses to negative signals. This suggests that immediate emotional reactions substantially shape consumer behavior. By offering empirical evidence on the impact of emotions during the decision-making process, this study underscores the importance of short-term emotions in driving economic behavior. Additionally, it contributes to the limited empirical research on emotions and alcohol use as a mass phenomenon.

It is crucial to highlight that this study examines both emotions and alcohol consumption at an aggregated level, focusing on the average effects among fans present in the stadium, without drawing conclusions at the individual level. Additionally, from a psychological perspective, it's important to note that these mathematically derived emotions, based on expected match outcomes, may not directly correspond to fans' actual experiences during the match. Another limitation of this study is related to the specific setting chosen. Although it allows for clear observation and isolation of the effects, it may not be fully representative. While it offers a solid foundation for extending findings to other markets, these results should be applied to other contexts with caution. Future research could explore these effects in different sports settings or countries. For instance, soccer, as a low-scoring game, naturally involves fewer moments of surprise but more suspense. Examining how these emotional dynamics affect behavior in high-scoring sports like basketball, where surprises are more frequent, would provide valuable insights.

The second contribution of this thesis builds on the first by addressing methodological challenges in estimating complex, high-dimensional models with sparsity in Bayesian frameworks, including overfitting, effective parameter shrinkage, and multicollinearity. It introduces the Dirichlet-horseshoe, a new Bayesian prior designed to improve the estimation of these complex, sparse models. Traditional methods, like the Bayesian lasso and spike-and-slab priors, often struggle with high-dimensional data, leading to issues like overfitting or inefficient coefficient shrinkage. The Dirichlet-horseshoe prior overcomes these limitations by applying stronger shrinkage to less important predictors and lighter shrinkage to more significant ones, adjusting to varying signal strengths. This approach outperforms existing methods in both simulated and real-world applications, demonstrating its robustness and accuracy across different sparsity levels.

The key contribution lies in this methodological innovation, offering researchers an alternative shrinkage prior for estimating increasingly complex models. However, the broader applicability of this prior outside the case studies presented requires further investigation. Testing this prior across different domains and model structures could reveal additional strengths and limitations, guiding refinements for use in different high-dimensional settings with sparsity.

The third contribution of this thesis examines pre-match price movements in tennis betting contracts to investigate the behavioral responses of average market participants and market makers to new information, particularly when initial prices are biased or imprecise. It further quantifies the speed of price adjustments driven by their actions. By analyzing time series data from multiple bookmakers and matches, this study estimates learning rate parameters to assess the

speed at which information is incorporated, offering empirical insights into price formation processes and equilibrium dynamics within the sports betting market.

Initial prices are often noisy and biased but become more accurate as the market approaches closure, driven primarily by rational bettors who update their beliefs in response to new information. Learning does not occur immediately but intensifies as the betting period progresses. Learning rates are smaller compared to those observed in financial stock markets, with significant heterogeneity across bookmakers and market segments, and faster learning in higher-stakes matches. By using sports betting contracts with known terminal values, the study contributes to understanding price discovery while avoiding biases introduced by proxy values.

However, these findings are specific to the tennis betting market and may not fully generalize to other sports or financial markets with different structures and participants. Similarly, comparisons of estimated learning rates in sports betting market and stock markets should be approached carefully. Further research could expand the analysis to other betting formats or team sports to validate and extend these insights.

In conclusion, this dissertation contributes to the understanding of behavioral economics by offering empirical evidence on how immediate emotional responses and new information influence market behavior. The findings highlight how emotions like surprise and suspense affect consumer decisions, such as alcohol consumption, in real-time, and how bettors drive price discovery and learning in sports betting markets. Additionally, this work presents methodological advancements for regularizing high-dimensional models in Bayesian frameworks, providing a novel shrinkage prior to address complex, sparsity-related issues in contexts where traditional methods fall short.

While the studies provide valuable insights, their context-specific nature limits their generalizability to other markets and domains. Future research could build on these findings to explore similar dynamics in different settings, including broader financial markets, where cognitive biases and emotional reactions also play a crucial role in decision-making. As Camerer, Loewenstein, and Rabin (2006) point out, the " 'franchising' of ideas to application areas (such as finance and labor economics), development of theoretical models, field studies, and including new types of psychology (such as attention, attribution, categorization, and limited memory)" (p. 33) aims to explore key questions in behavioral economics about the implications of bounded rationality, limited willpower, and self-interest. Hopefully, this work encourages further exploration of these topics and contributes to a more nuanced understanding of how emotions and stimuli shape economic behavior in real-world settings.

References

Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*.

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv (2007). "Loss Aversion Under Prospect Theory: A Parameter-Free Measurement". In: *Management science* 53.10, pp. 1659–1674.

Andres, Leander, Marc Fabel, and Helmut Rainer (2023). "How Much Violence Does Football Hooliganism Cause?" In: *Journal of Public Economics* 225, p. 104970.

Angelini, Giovanni, Luca De Angelis, and Carl Singleton (2022). "Informational Efficiency and Behaviour Within In-Play Prediction Markets". In: *International Journal of Forecasting* 38.1, pp. 282–299.

Angrist, Joshua David and Jörn-Steffen Pischke (2010). "The Credibility Revolution in Economics: How Better Research Design is Taking the Con out of Econometrics". In: *Journal of Economic Perspectives* 24 (2), pp. 3–30.

Armagan, Artin, David B. Dunson, and Jaeyong Lee (2013). "Generalized Double Pareto Shrinkage". In: *Statistica Sinica* 23.1, p. 119.

Ashley, Richard, Clive W. J. Granger, and Richard Schmalensee (1980). "Advertising and Aggregate Consumption: An Analysis of Causality". In: *Econometrica: Journal of the Econometric Society* 48.5, pp. 1149–1167.

Asparouhov, Tihomir, Ellen L. Hamaker, and Bengt Muthén (2017). "Dynamic Latent Class Analysis". In: *Structural Equation Modeling: A Multidisciplinary Journal* 24 (2), pp. 257–269.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny (1998). "A Model of Investor Sentiment". In: *Journal of financial economics* 49.3, pp. 307–343.

Barry, Adam, Steven Howell, Trevor Bopp, Michael Stellefson, Beth Chaney, Anna Piazza-Gardner, and Caroline Payne-Purvis (2014). "A Field-Based Community Assessment of Intoxication Levels Across College Football Weekends: Does It Matter Who's Playing?" In: *The Journal of Primary Prevention* 35 (6), pp. 409–416.

Berry, Donald A. (1987). *Multiple Comparisons, Multiple Tests and Data Dredging: a Bayesian Perspective*. Technical Report No. 486.

Bhadra, Anindya, Jyotishka Datta, Nicholas G. Polson, and Brandon Willard (2017). "The Horseshoe+ Estimator of Ultra-Sparse Signals". In: *Bayesian Analysis* 12.4, pp. 1105–1131.

Bhattacharya, Anirban, Debdeep Pati, Natesh S. Pillai, and David B. Dunson (2015). "Dirichlet–Laplace Priors for Optimal Shrinkage". In: *Journal of the American Statistical Association* 110.512, pp. 1479–1490.

Biais, Bruno, Pierre Hillion, and Chester Spatt (1999). "Price Discovery and Learning During the Preopening Period in the Paris Bourse". In: *Journal of Political Economy* 107.6, pp. 1218–1248.

Bizzozero, Paolo, Raphael Flepp, and Egon Franck (2016). "The Importance of Suspense and Surprise in Entertainment Demand: Evidence from Wimbledon". In: *Journal of Economic Behavior & Organization* 130 (C), pp. 47–63.

Brandt, Holger, Jenna Cambria, and Augustin Kelava (2018). "An Adaptive Bayesian Lasso Approach with Spike-and-Slab Priors to Identify Multiple Linear and Nonlinear Effects in Structural Equation Models". In: *Structural Equation Modeling: A Multidisciplinary Journal* 25 (6), pp. 946–960.

Brooks, Stephen and Andrew Gelman (1998). "General Methods for Monitoring Convergence of Iterative Simulations". In: *Journal of Computational and Graphical Statistics* 7, pp. 434–455.

Brown, Philip J. and Jim E. Griffin (2010). "Inference with normal-gamma prior distributions in regression problems". In: *Bayesian Analysis* 5.1, pp. 171–188.

Brown, William O. and Raymond D. Sauer (1993). "Fundamentals or Noise? Evidence From the Professional Basketball Betting Market". In: *The Journal of Finance* 48.4, pp. 1193–1209.

Buraimo, Babatunde, David Forrest, Ian G. McHale, and J. D. Tena (2020). "Unscripted Drama: Soccer Audience Response to Suspense, Surprise, and Shock". In: *Economic Inquiry* 58 (2), pp. 881–896.

Camerer, Colin F. (1989). "Does the Basketball Market Believe in the 'Hot Hand'?" In: *The American Economic Review* 79.5, pp. 1257–1261.

Camerer, Colin F., George Loewenstein, and Matthew Rabin (2006). "Behavioral Economics". In: *Econometric Society Monographs* 42, p. 181.

Card, David and Gordon Dahl (2011). "Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior". In: *The Quarterly Journal of Economics* 126 (1), pp. 103–143.

Cardazzi, Alexander, Bryan C. McCannon, Brad R. Humphreys, and Zachary Rodriguez (2022). "Emotional Cues and Violent Behavior: Unexpected Basketball Losses Increase Incidents of Family Violence". In: *The Journal of Law, Economics, and Organization*.

Carvalho, Carlos M., Nicholas G. Polson, and James G. Scott (2008). "The Horseshoe Estimator for Sparse Signals". Duke University Department of Statistical Science, Discussion Paper 2008-31.

— (2009). "Handling Sparsity via the Horseshoe". In: *Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics*. Vol. 5, pp. 73–80.

— (2010). "The Horseshoe Estimator for Sparse Signals". In: *Biometrika* 97.2, pp. 465–480.

Chen, M. Keith, Venkat Lakshminarayanan, and Laurie R. Santos (2006). "How Basic are Behavioral Biases? Evidence From Capuchin Monkey Trading Behavior". In: *Journal of political economy* 114.3, pp. 517–537.

Chen, Qi (2024). "Behavioral Economics: Understanding the Psychological Factors Driving Consumer Decisions". In: *International Journal of Global Economics and Management* 4.1, pp. 99–1.

Cho, Kyunghyun, Bart van Merrienboer, Dzmitry Bahdanau, and Yoshua Bengio (2014). *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches*. arXiv.

Cooper, M. Lynne, Michael R. Frone, Marcia Russell, and Pamela Mudar (1995). "Drinking to Regulate Positive and Negative Emotions: a Motivational Model of Alcohol Use". In: *Journal of Personality and Social Psychology* 69 (5), pp. 990–1005.

Croxson, Karen and J. James Reade (2014). "Information and Efficiency: Goal Arrival in Soccer Betting". In: *The Economic Journal* 124.575, pp. 62–91.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam (1998). "Investor Psychology and Security Market Under- and Overreactions". In: *The Journal of Finance* 53.6, pp. 1839–1885.

Dare, William H. and S. Scott MacDonald (1996). "A Generalized Model for Testing the Home and Favorite Team Advantage in Point Spread Markets". In: *Journal of Financial Economics* 40.2, pp. 295–318.

Dickey, David A. and Wayne A. Fuller (1979). "Distribution of the Estimators for Autoregressive Time Series With a Unit Root". In: *Journal of the American statistical association* 74.366a, pp. 427–431.

Dillenberger, David (2010). "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior". In: *Econometrica* 78 (6), pp. 1973–2004.

Edmans, Alex, Diego Garcia, and Øyvind Norli (2007). "Sports Sentiment and Stock Returns". In: *The Journal of Finance* 62 (4), pp. 1967–1998.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica (2015). "Suspense and Surprise". In: *Journal of Political Economy* 123 (1), pp. 215–260.

Engle, Robert F. and Jose Gonzalo Rangel (2008). "The Spline-GARCH Model for Low-Frequency Volatility and its Global Macroeconomic Causes". In: *The Review of Financial Studies* 21.3, pp. 1187–1222.

Epskamp, Sacha, Mijke Rhemtulla, and Denny Borsboom (2017). "Generalized Network Psychometrics: Combining Network and Latent Variable Models". In: *Psychometrika* 82, pp. 904–927.

Eren, Ozkan and Naci Mocan (2018). "Emotional Judges and Unlucky Juveniles". In: *American Economic Journal: Applied Economics* 10 (3), pp. 171–205.

Falk, Armin and James Joseph Heckman (2009). "Lab Experiments are a Major Source of Knowledge in the Social Sciences". In: *Science* 326 (5952), pp. 535–8.

Fama, Eugene F. (1970). "Efficient Capital Markets". In: *The Journal of Finance* 25.2, pp. 383–417.

Fischer, Lukas, Michael Nagel, Augustin Kelava, and Tim Pawlowski (2023). "Celebration Beats Frustration: Emotional Cues and Alcohol Use During Soccer Matches". In: *Available at SSRN: https://ssrn.com/abstract=4569227 or http://dx.doi.org/10.2139/ssrn.4569227*.

— (2024). "Emotional Drinking: Surprise, Suspense, and Alcohol Use During Soccer Matches". In: *Available at SSRN: https://ssrn.com/abstract=4912703*.

Forrest, David and Robert Simmons (2008). "Sentiment in the Betting Market on Spanish Football". In: *Applied Economics* 40.1, pp. 119–126.

Gandar, John, Richard Zuber, Thomas O'brien, and Ben Russo (1988). "Testing Rationality in the Point Spread Betting Market". In: *The Journal of Finance* 43.4, pp. 995–1008.

Gandar, John M., William H. Dare, Craig R. Brown, and Richard A. Zuber (1998). "Informed Traders and Price Variations in the Betting Market for Professional Basketball Games". In: *The Journal of Finance* 53.1, pp. 385–401.

Gao, Xiaohui and Tse-Chun Lin (2015). "Do individual investors treat trading as a fun and exciting gambling activity? Evidence from repeated natural experiments". In: *The Review of Financial Studies* 28.7, pp. 2128–2166.

Garicano, Luis, Ignacio Palacios-Huerta, and Canice Prendergast (2005). "Favoritism Under Social Pressure". In: *Review of Economics and Statistics* 87 (2), pp. 208–216.

Gelman, Andrew and Donald B. Rubin (1992). "A Single Series From the Gibbs Sampler Provides a False Sense of Security". In: *Bayesian statistics* 4.1, pp. 625–631.

Genesove, David and Christopher Mayer (2001). "Loss Aversion and Seller Behavior: Evidence From the Housing Market". In: *The quarterly journal of economics* 116.4, pp. 1233–1260.

George, Edward I. and Robert E. McCulloch (1993). "Variable Selection via Gibbs Sampling". In: *Journal of the American Statistical Association* 88 (423), pp. 881–889.

Germain, Laurent, Nour Meddahi, and Eric Renault (1996). "The Speed of Learning in Financial Markets". In: *Manuscript. Toulouse: Toulouse Univ*.

Ghosh, Joyee, Yingbo Li, and Robin Mitra (2018). "On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression". In: *Bayesian Analysis* 13.2, pp. 359–383.

Glorot, Xavier and Yoshua Bengio (2010). "Understanding the Difficulty of Training Deep Feedforward Neural Networks". In: *AISTATS*. Ed. by Yee Whye Teh and D. Mike Titterington. Vol. 9. JMLR.org, pp. 249–256.

Goodell, John W., Satish Kumar, Purnima Rao, and Shubhangi Verma (2023). "Emotions and Stock Market Anomalies: A Systematic Review". In: *Journal of Behavioral and Experimental Finance* 37, p. 100722.

Gray, Philip K. and Stephen F. Gray (1997). "Testing Market Efficiency: Evidence From the NFL Sports Betting Market". In: *The Journal of Finance* 52.4, pp. 1725–1737.

Greeley, Janet and Tian Oei (1999). "Alcohol and Tension Reduction". In: *Psychological Theories of Drinking and Alcoholism*. Second. The Guilford Press, pp. 14–53.

Griswold, Max G., Nancy Fullman, Caitlin Hawley, Nicholas Arian, Stephanie R. M. Zimsen, Hayley D. Tymeson, et al. (2018). "Alcohol Use and Burden for 195 Countries and Territories,

1990-2016: A Systematic Analysis for the Global Burden of Disease Study 2016". In: *Lancet* 392 (10152), pp. 1015–1035.

Hansen, Lars Peter (1982). "Large Sample Properties of Generalized Method of Moments Estimators". In: *Econometrica: Journal of the econometric society* 50.4, pp. 1029–1054.

Hansen, Lars Peter, John Heaton, and Amir Yaron (1996). "Finite-Sample Properties of Some Alternative GMM Estimators". In: *Journal of Business & Economic Statistics* 14.3, pp. 262–280.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2015). *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. arXiv.

Hegarty, Tadgh and Karl Whelan (2024). *Comparing two Methods for Testing the Efficiency of Sports Betting Markets*. Tech. rep. UCD Centre for Economic Research Working Paper Series.

Hodrick, Robert J. (1987). "The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets". In: *Fundamentals of Pure and Applied Economics* 24.

Hoffman, Matthew D. and Andrew Gelman (2014). "The No-U-Turn Sampler: Adaptively Setting Path Lengthsin Hamiltonian Monte Carlo". In: *Journal of Machine Learning Research* 15 (47), pp. 1593–1623.

Isen, Alice M. (2008). "Handbook of Emotions". In: ed. by Michael Lewis, Jeannette M. Haviland-Jones, and Lisa Feldman Barrett. Third. New York: The Guilford Press. Chap. Some ways in which positive affect influences decision making and problem solving, pp. 548–573.

Ishwaran, Hemant and J. Sunil Rao (2005). "Spike and Slab Variable Selection: Frequentist and Bayesian Strategies". In: *The Annals of Statistics* 33 (2), pp. 730–773.

Johnstone, Iain M. and Bernard W. Silverman (2004). "Needles and Straw in Haystacks: Empirical Bayes Estimates of Possibly Sparse Sequences". In: *The Annals of Statistics* 32.4, pp. 1594–1649.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem". In: *Journal of political Economy* 98.6, pp. 1325–1348.

Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision Under Risk". In: *Econometrica* 47 (2), pp. 263–291.

Kelava, Augustin and Holger Brandt (2019). "A Nonlinear Dynamic Latent Class Structural Equation Model". In: *Structural Equation Modeling: A Multidisciplinary Journal* 26 (4), pp. 509–528.

Kessler, Judd B., Andrew McClellan, James Nesbit, and Andrew Schotter (2022). "Short-Term Fluctuations in Incidental Happiness and Economic Decision-Making: Experimental Evidence from a Sports Bar". In: *Experimental Economics* 25 (1), pp. 141–169.

Kingma, Diederik P. and Jimmy Ba (2014). *Adam: A Method for Stochastic Optimization*. arXiv.

Klick, Jonathan and John MacDonald (2021). "Sobering Up After the Seventh Inning: Alcohol and Crime Around the Ballpark". In: *Journal of Quantitative Criminology* 37 (611), pp. 813–34.

Klostermann, Keith C. and William Fals-Stewart (2006). "Intimate Partner Violence and Alcohol Use: Exploring the Role of Drinking in Partner Violence and its Implications for Intervention". In: *Aggression and Violent Behavior* 11 (6), pp. 587–597.

Kőszegi, Botond and Matthew Rabin (2006). "A Model of Reference-Dependent Preferences". In: *The Quarterly Journal of Economics* 121 (4), pp. 1133–1165.

Krieger, Kevin, Justin L. Davis, and James Strode (2021). "Patience is a Virtue: Exploiting Behavior Bias in Gambling Markets". In: *Journal of Economics and Finance* 45.4, pp. 735–750.

Laibson, David (2001). "A Cue-Theory of Consumption". In: *The Quarterly Journal of Economics* 116 (1), pp. 81–119.

Lang, Alan R., Christopher J. Patrick, and Werner G. K. Stritzke (1999). "Alcohol and Emotional Response: A Multidimensional-Multilevel Analysis". In: *Psychological Theories of Drinking and Alcoholism*. Second. The Guilford Press, pp. 328–371.

Leng, Chenlei, Minh-Ngoc Tran, and David Nott (2014). "Bayesian Adaptive Lasso". In: *Annals of the Institute of Statistical Mathematics* 66, pp. 221–244.

Lindo, Jason M., Peter Siminski, and Isaac D. Swensen (2018). "College Party Culture and Sexual Assault". In: *American Economic Journal: Applied Economics* 10 (1), pp. 236–265.

Liu, Xiao, Matthew Shum, and Kosuke Uetake (2020). "Passive and Active Attention to Baseball Telecasts: Implications for Content (Re-)Design". In: *New York University Stern School of Business Forthcoming*.

Loewenstein, George (2000). "Emotions in Economic Theory and Economic Behavior". In: *American Economic Review* 90 (2), pp. 426–432.

Loomes, Graham and Robert Sugden (1982). "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty". In: *The Economic Journal* 92 (368), pp. 805–824.

Marquardt, Donald W. and Ronald D. Snee (1975). "Ridge Regression in Practice". In: *The American Statistician* 29.1, pp. 3–20.

Marzilli Ericson, Keith M. and Andreas Fuster (2011). "Expectations as Endowments: Evidence on Reference-Dependent Preferences From Exchange and Valuation Experiments". In: *The Quarterly Journal of Economics* 126.4, pp. 1879–1907.

Meij, Leander van der, Fabian Klauke, Hannah L. Moore, Yannick S. Ludwig, Mercedes Almela, and Paul A. M. van Lange (2015). *Football Fan Aggression: The Importance of Low Basal Cortisol and a Fair Referee*.

Mitchell, T. J. and J. J. Beauchamp (1988). "Bayesian Variable Selection in Linear Regression". In: *Journal of the American Statistical Association* 83 (404), pp. 1023–1032.

Moskowitz, Tobias J. (2021). "Asset Pricing and Sports Betting". In: *The Journal of Finance* 76.6, pp. 3153–3209.

Munyo, Ignacio and Martín Rossi (2013). "Frustration, Euphoria, and Violent Crime". In: *Journal of Economic Behavior and Organization* 89 (C), pp. 136–142.

Nagel, Michael (2024). "Price Formation Dynamics and Learning in the Tennis Sports Betting Market". In: *Available at SSRN: http://dx.doi.org/10.2139/ssrn.5043714*.

Nagel, Michael, Lukas Fischer, Tim Pawlowski, and Augustin Kelava (2024). "An Alternative Prior for Estimation in High-Dimensional Settings". In: *Structural Equation Modeling: A Multidisciplinary Journal* 31 (6), pp. 939–951.

Neal, Dan J. and Kim Fromme (2007). "Hook 'em Horns and Heavy Drinking: Alcohol Use and Collegiate Sports". In: *Addictive Behaviors* 32 (11), pp. 2681–2693.

Nelson, Daniel B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach". In: *Econometrica: Journal of the econometric society* 59.2, pp. 347–370.

Palacios-Huerta, Ignacio (1999). "The Aversion to the Sequential Resolution of Uncertainty". In: *Journal of Risk and Uncertainty* 18 (3), pp. 249–269.

— (2023). "The Beautiful Dataset". In: *Journal of Economic Literature*. Forthcoming.

Park, Trevor and George Casella (2008). "The Bayesian Lasso". In: *Journal of the American Statistical Association* 103.482, pp. 681–686.

Patrick, Megan and John Schulenberg (2011). "How Trajectories of Reasons for Alcohol Use Relate to Trajectories of Binge Drinking: National Panel Data Spanning Late Adolescence to Early Adulthood". In: *Developmental Psychology* 47 (2), pp. 311–7.

Pawlowski, Tim, Dooruj Rambaccussing, Philip Ramirez, J. James Reade, and Giambattista Rossi (2024). "Exploring Entertainment Utility from Football Games". In: *Journal of Economic Behavior & Organization* 223, pp. 185–198.

Petrone, Sonia, Judith Rousseau, and Catia Scricciolo (2014). "Bayes and empirical Bayes: do they merge?" In: *Biometrika* 101.2, pp. 285–302.

Piironen, Juho and Aki Vehtari (2015). "Projection Predictive Variable Selection Using Stan+ R". In: *arXiv preprint arXiv:1508.02502*.

— (2017). "Sparsity Information and Regularizationin the Horseshoe and Other Shrinkage Priors". In: *Electronic Journal of Statistics* 11 (2), pp. 5018–5051.

Polson, Nicholas G. and James G. Scott (2010). "Shrink Globally, Act Locally: Sparse Bayesian Regularization and Prediction". In: *Bayesian statistics* 9.501-538, p. 105.

Pope, Devin G. and Maurice E. Schweitzer (2011). "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes". In: *American Economic Review* 101.1, pp. 129–157.

Ramirez, Philip, J James Reade, and Carl Singleton (2023). "Betting on a Buzz: Mispricing and Inefficiency in Online Sportsbooks". In: *International Journal of Forecasting* 39.3, pp. 1413–1423.

Rao, Vennapu Lakshmana and N. Vijay Mohan (2021). "Impact of Weather on Sports and Sport Injuries". In: *International Journal of Physical Education, Sports and Health* 8 (3), pp. 9–13.

Reddi, Sashank J., Satyen Kale, and Sanjiv Kumar (2019). *On the Convergence of Adam and Beyond*. arXiv.

Richardson, Travis, Georgios Nalbantis, and Tim Pawlowski (2023). "Emotional Cues and the Demand for Televised Sports: Evidence from the UEFA Champions League". In: *Journal of Sports Economics* 24 (8).

Room, Robin, Pia Mäkelä, Vivek Benegal, Siri Greenfield Thomas K .and Hettige, Nazarius M. Tumwesigye, and Richard Wilsnack (2013). "Times to Drink: Cross-Cultural Variations in Drinking in the Rhythm of the Week". In: *International Journal of Public Health* 57 (1), pp. 107–117.

Salvatier, John, Thomas V. Wiecki, and Christopher Fonnesbeck (2016). "Probabilistic Programming in Python Using PyMC3". In: *PeerJ Computer Science* 2, e55.

Schwarz, Gideon (1978). "Estimating the Dimension of a Model". In: *The annals of statistics* 2.6, pp. 461–464.

Scott, James G. and James O. Berger (2006). "An Exploration of Aspects of Bayesian Multiple Testing". In: *Journal of statistical planning and inference* 136.7, pp. 2144–2162.

Simonov, Andrey, Raluca M. Ursu, and Carolina Zheng (2023). "Suspense and Surprise in Media Product Design: Evidence from Twitch". In: *Journal of Marketing Research* 60 (1), pp. 1–24.

Thaler, Richard and William Ziemba (1988). "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries". In: *Journal of Economic perspectives* 2.2, pp. 161–174.

Tibshirani, Robert (1996). "Regression Shrinkage and Selection Via the Lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, pp. 267–288.

Tymula, Agnieszka A. and Paul W. Glimcher (2018). "The Nature of Emotion: Fundamental Questions". In: ed. by A. S. Fox, R. C. Lapate, A. J. Shackman, and R. J. Davidson. Second. New York: Oxford University Press. Chap. Emotions Through the Lens of Economic Theory, pp. 338–342.

Van Der Pas, Stéphanie L., Bas J. K. Kleijn, and Aad W. Van Der Vaart (2014). "The Horseshoe Estimator: Posterior Concentration Around Nearly Black Vectors". In: *Electronic Journal of Statistics* 8, pp. 2585–2618.

Vecer, Jan, Frantisek Kopriva, and Tomoyuki Ichiba (2009). "Estimating the Effect of the Red Card in Soccer: When to Commit an Offense in Exchange for Preventing a Goal Opportunity". In: *Journal of Quantitative Analysis in Sports* 5 (1), p. 8.

Ventura-Cots, Meritxell, Ariel E. Watts, Monica Cruz-Lemini, Neil D. Shah, Nambi Ndugga, Peter McCann, A. S. Barritt 4th, Anant Jain, Samhita Ravi, Carlos Fernandez-Carrillo, Juan G. Abraldes, Jose Altamirano, and Ramon Bataller (2019). "Colder Weather and Fewer Sunlight Hours Increase Alcohol Consumption and Alcoholic Cirrhosis Worldwide". In: *Hepatology* 69 (5), pp. 1916–1930.

Vives, Xavier (1995). "The Speed of Information Revelation in a Financial Market Mechanism". In: *Journal of Economic Theory* 67.1, pp. 178–204.

Walras, Léon (1889). *Éléments d'économie politique pure: ou, théorie de la richesse sociale*.

Watt, Helen M. G. (2004). "Development of Adolescents' Self-Perceptions, Values, and Task Perceptions According to Gender and Domain in 7th-Through 11th-Grade Australian Students". In: *Child Development* 75.5, pp. 1556–1574.

Williams, Donald R. and Josue E. Rodriguez (2022). "Why Overfitting is Not (Usually) a Problem in Partial Correlation Networks". In: *Psychological Methods* 27.5, pp. 822–840.

Wolfe, Jeannette, Ricardo Martinez, and Warren A. Scott (1998). "Baseball and Beer: An Analysis of Alcohol Consumption Patterns Among Male Spectators at Major-League Sporting Events". In: *Annals of Emergency Medicine* 31 (5), pp. 629–632.

Wolfers, Justin (2006). "Point Shaving: Corruption in NCAA Basketball". In: *American Economic Review* 96.2, pp. 279–283.

Wood, Stacy, Melayne Morgan McInnes, and David A. Norton (2011). "The Bad Thing about Good Games: The Relationship between Close Sporting Events and Game-Day Traffic Fatalities". In: *Journal of Consumer Research* 38 (4), pp. 611–621.

Zonda, Tamás, KÁroly Bozsonyi, Elõd Veres, David Lester, and Michael Frank (2009). "The Impact of Holidays on Suicide in Hungary". In: *OMEGA - Journal of Death and Dying* 58 (2), pp. 153–162.