SPORTS BETTING AS AN ALTERNATIVE INVESTMENT

ELLIOT CHRISTOPHERS

Bachelor Thesis

Stockholm School of Economics

2024

Sports Betting as an Alternative Investment

Abstract:

This paper provides a broad overview of sports betting, considering its markets and evaluating whether it can provide positive risk-adjusted returns as an alternative financial investment. Across a large dataset of closing odds from over 165,000 soccer games from top European leagues, we develop a trading strategy which generates promising returns. We then conduct a survey of sports-betting models in academic literature, obtaining a representative distribution of returns from betting across sports, combining this with detailed data for prices and events to simulate a fictional sports-betting asset. With a returns profile for this asset, we then include it in an optimised mean-variance portfolio, integrating it with conventional asset indices, finding that the sports-betting asset compares favourably in various investing scenarios, including a value-at-risk analysis. Finally, we discuss financial frictions specific to this market, concluding that sports betting does not provide sufficient liquidity for large wealth managers, but that shrewd investors with a nuanced understanding of the market can likely generate strong returns.

Keywords:

Sports betting, Investment returns, Liquidity, Value at risk, Portfolio optimisation

Author:

Elliot Christophers (25487)

Tutors:

Ran Xing, Assistant Professor, Swedish House of Finance

Examiner:

Adrien d'Avernas, Assistant Professor, Swedish House of Finance

Bachelor Thesis
Bachelor Program in Business and Economics
Stockholm School of Economics

1 Introduction

In conventional finance, capital is allocated for investment towards stocks and bonds, derivative contracts of these, foreign exchange, real estate, commodities, and the like. Such asset classes are well-researched, well-known and generally well-returning. Alternative finance is investment in less well-recognised assets, ranging from the mundane to the obscure. Financial research has been performed regarding the investment in wine (Chu, 2014), art (Worthington and Higgs, 2004), classic cars (Laurs and Renneboog, 2019), collectible toys (McAlister and Cornwell, 2012), stamps (Dimson and Spaenjers, 2011), luxury watches (Kylämarkula, 2023), sports memorabilia (Burton and Jacobsen, 1999), rare coins (Dickie et al., 1994), musical instruments (Graddy and Margolis, 2011), and more. Beyond having the potential to provide high risk-adjusted returns due to lower competition from high-end financial analysts; regardless of the alpha provided, alternative investments allow strong portfolio diversification, being relatively uncorrelated with the typical financial assets listed above. This paper seeks to evaluate the potential of sports betting as another such source of alternative investment. Although evidently more closely related to conventional financial markets than wine or art, with similar markets and price mechanisms; sports betting is hardly a mainstay in wealth-management contexts. This paper analyses whether it should possess a more prominent role in investment scenarios.

Sports betting has received extensive coverage in financial and economic research. It is not a particularly prestigious nor renowned corpus in broader financial circles, but to the initiated it is vast and diverse. Analyses range from evaluating Fama-style efficiency in betting markets (Fama, 1970; Dixon and Pope, 2004; Kuypers, 2000) to the design of various varieties of calibrated forecasting models (Maher, 1982; Dixon and Coles, 1997) to searching for evidence of behavioural biases (Moskowitz, 2021; Law and Peel, 2002) to the quest for arbitrage-style profit opportunities. It is this last category which this paper, in part, seeks to make a contribution towards.

Arbitrage, in finance, is the pursuit of risk-free positive returns through, in different markets, executing simultaneous transactions of financial contracts which pertain to the same underlying asset. In a sports-betting context, this would entail betting on a collectively exhaustive set of outcomes on an underlying event, with different bookmakers, in such proportions which guarantee a certain positive payoff irrespective of the event's realised outcome. Arbitrage in sports betting is easily identified and there is publicly-available information regarding these opportunities. The homepage of betmonitor.com, for instance, invariably displays numerous events where arbitrage is available, and highly lucrative. However, it is equally well-known that arbitrageurs are banned from betting-sites as soon as they are detected.

A second class of arbitrage exists in the sports-betting literature, namely, quasi-arbitrage, where profitable betting-opportunities are identified by analysing the distribution of prices offered on an event in the cross-section of bookmakers, but selecting only one outcome to back, rather than the complete set, as under regular arbitrage. This amounts to identifying undervalued assets. For all events covered by multiple bookmakers, there will almost inevitably be some discrepancy between the odds provided, as the bookmakers will have slightly different opinions regarding the probabilities of different outcomes, as well as having different pricing strategies. Quasi-arbitrage betting is the endeavour to exploit any eventual pricing inefficiencies resulting from such odds discrepancies, by comparing some aggregate price across the different bookmakers, to the longest, most favourable price amongst the cohort. If the aggregate bookmaker opinion represents the "true", or "fundamental" probability of the event in question occuring, then the implied probability of the longest odds will be smaller than this probability, suggesting

that this bookmaker has underestimated the possibility of the given event being realised. This presents an edge that can in theory be exploited.

There is thorough discussion of quasi-arbitrage — or simply, quarb — strategies in betting literature. Amongst several other instances, Paton and Williams (2005) coined the term and analysed its profitability in spread betting markets, Smith et al. (2009) investigated quarbs in fixed-odds horse racing markets, Cortis (2016) applied quarb methods to the 2012 soccer European Championships, and Kaunitz et al. (2017) — although they did not use the term quasi-arbitrage — compared mean to longest odds in prominent soccer leagues.

The purpose of this paper, on the one hand, is to leverage quasi-arbitrage methodologies to evaluate betting opportunities, identify profitable ones, and attain a positive cumulative growth rate. In a sense, this exercise is a proof of concept, a demonstration of the existence of positive returns in sports betting, in a characteristically general manner. For our purposes, the salient property of quarb strategies is that they are market-agnostic, needing only access to a set of odds provided by different bookmakers on an event to be applicable. There is no need to perform extensive modelling regarding the particular features of the market (as did the seminal sports-betting papers, for instance, by Maher, and Dixon and Coles, discussing the Poisson-distribution of goals in soccer). As such, quasi-arbitrage is fully general, applicable to all sports, all markets, all outcomes.

This analysis will be performed with respect to historical soccer data. But to understand the validity of sports betting as a source of investment returns, it is important to consider the broader sports landscape. Therefore, having demonstrated the existence (and simplicity) of positive returns, we conduct a survey of sports-betting models presented in other research papers. The objective here is to more fully comprehend the scope of such models, namely, their returns and the frequency of their opportunities. Given this practical information, we can approximate a sports-betting asset, generating a fictional distribution of monthly returns.

We then integrate the sports-betting investment into a more conventional financial portfolio to evaluate whether the alternative investment contributes positively towards an already optimised position. For the various asset classes under consideration, we iteratively remove each from the asset pool and calculate the Sharpe-ratio (Sharpe, 1966) of this reduced portfolio, comparing it to the Sharpe-ratio of the full portfolio. We also perform a Value at Risk analysis, optimising the portfolio with respect to a well-defined risk appetite. In these distinct investing scenarios, the sports-betting asset consistently contributes more value than the remaining assets.

Finally, we discuss market frictions inherent to the sports-betting market, and evaluate the scale at which sports-betting might be leveraged as a wealth-management strategy. The two primary frictions which impede sports betting as an investable asset are that a certain class of bookmakers ban profit seekers, and that all bookmakers set limits to the stakes that they accept from bettors in an enforced liquidity deficit. We discuss obvious tells which bookmakers use to identify problematic customers, while estimating a level of investable capital after which diminishing returns to scale commences, with sports betting thereafter increasingly unattractive to investors with extremely large assets under management.

2 Dataset

The primary dataset, accessed from football-data.co.uk, analysed in this paper consists of 22 seasons' worth of association football (soccer) games from 22 of the biggest leagues across Europe. From a total of 167,536 games spanning the 2001/2002 season through 2022/2023, the dataset contains the result of each game and match-outcome closing odds (home win, draw,

away win) from as many as thirteen individual bookmakers. These bookmakers are listed with abbreviations in the results below, with B365 corresponding to Bet365, BS to Blue Square, BW to Betway, GB to Gamebookers, IW to Interwetten, LB to Ladbrokes, SB to Sportingbet, SJ to Stan James, VC to Victor Chandler, and WH to William Hill. The dataset is ordered chronologically. We also collect data for odds and events for five other sports, as well as the indices used in the portfolio integration section. Being secondary, these data are described in the appendix, in subsections "Games and Odds by Sport", as well as "Assets", respectively.

3 Theoretical Background

In this paper, the analyses conducted in two separate research studies will be replicated and extended. These are "On determining probability forecasts from betting odds" by Štrumbelj (2014) and "Beating the bookies with their own numbers - and how the online sports betting market is rigged" by Kaunitz et al. (2017).

There are myriad mechanisms for removing bookmaker margins and extracting probabilistic information from bookmaker prices. Štrumbelj presents a framework for comparing these methods. In his original paper, he compared two established methods as well as one which he designed, evaluating which method extracted the most outcome-relevant information from the odds. Since our quasi-arbitrage ambitions require leveraging the information provided by the cross-section of bookmaker odds, we commence our analysis by replicating Štrumbelj in comparing several margin-standardisation algorithms. We extend the analysis by also evaluating several opinion-aggregation mechanisms, identifying an optimal method for extracting a true probability distribution from a set of bookmaker forecasts.

Kaunitz et al. conduct a quasi-arbitrage analysis, in a highly practical investment analysis, where the probabilistic forecast obtained from the mean odds is used to bet at the longest odds. We replicate and extend their model, hoping to find the same superiority of aggregate odds over longest odds. We also further parameterise the Kaunit et al. model, finding increased profits. There exists a large set of quasi-arbitrage trading models in the sports-betting literature (see introduction). We choose the Kaunitz et al. model, specifically, because it is the most simplistic, while nonetheless generating strong returns, which corroborates our argument regarding the ease with which favourable investment returns can be produced with sports betting.

There is more to sports betting than the win market outcome for soccer. When conducting the survey of betting strategies in research literature, our aim is to gain an understanding of what ranges of profitability exist for sports betting, in general. As such, we do not restrict ourselves to soccer, but also explore American football, baseball, basketball, cricket, tennis and ice hockey. The content and narrative of these analyses are not of direct interest to us, as we merely extract their conclusions regarding the profitability and availability of betting opportunities. For each paper, we extract a per-bet return and a metric representing the availability of such profitable bets.

Regarding alternative investment, the introduction already presents a set of studies discussing unconventional methods of wealth-management in numerous guises. These are of course interesting in themselves, but are particularly pertinent for our purposes when we aspire to integrate our sports-betting investment with a broader investment portfolio. We use mean-variance portfolio construction (Markowitz, 1952) to integrate our sports-betting asset with these alternative investments, as well as the more conventional indices: stocks, bonds, and so on.

4 Sports Betting

Sports betting is a service provided by bookmakers, allowing individuals to make bets regarding the outcome of sporting events, as well as various submarkets within these main events. In soccer, the main market is the win market, where punters can decide to back three outcomes: whether a fixture between two teams will be won by the home team, the away team, or whether it will end as a draw. The myriad subevents that are also offered by bookmakers include the outcome of each half, whether the total number of goals scored in the game will or will not exceed one of several thresholds, whether individual players will score one or several goals, and which team will have more corners. Essentially, all aspects of the underlying sport can be bet on, and indeed, are bet on.

Prices for sports-betting markets are known as odds, typically quoted in one of three different formats: decimal, fractional, American. Whenever we refer to odds o, these will be in decimal format. This notation entails that if a unit is staked at o and the bet is successful, the total sum returned from the bookmaker to the bettor will be o units, and the profit will be $o-1$, units. Naturally, if the bet fails, the entire unit is lost. As an example, the very first game in our dataset was contested between Charlton and Everton during the opening weekend of the 2001/02 English Premier League season on the 18th of August 2001, where the most competitive (longest) odds, that is, those providing the highest return, were 2.0, 3.25 and 3.75, for the home win, draw and away win outcomes, respectively.

If the probability of the event occurring is p, then there is a p chance of increasing capital, with a unit bet, by $o-1$ and a $1-p$ chance of losing the unit, meaning that the expected value of the bet is $EV = p(o-1)-(1-p) = po-1$. If the bet is fair, then $EV = 0 = po-1$ which gives that $p = \frac{1}{o}$: when odds are expressed in decimal form, the odds residual can be interpreted as a probability. However, this interpretation is not fully accurate. Considering the example game between Charlton and Everton above, the residuals are 0.5, 0.3077 and 0.2667 (to four decimal places), which sum to 1.0744, clearly greater than one and thus not consistent with probability axioms. Although not a true probability distribution across the event, there is still information to be gained from the odds, for instance, in this case, that home win (Charlton) is the most likely outcome. It is also possible to standardise the odds residuals so that they do sum to one and can then be interpreted as probabilities. Indeed, we often refer to the implied probabilities of odds before and after standardisation. This excess implied probability above 1 is known as the bookmaker's margin, whereby they set prices lower than the true probabilities dictate to increase their profits (higher odds residual means lower odds, meaning lower profits for bettors and more profits for the bookmaker). This is a transaction cost that prospective bettors must overcome.

Quasi-arbitrage relies solely upon implied probabilities to identify profitable bets. Consider the fictitious set of odds presented on a two-outcome market in Table 1. For each bookmaker, odds for each of the two outcomes are quoted. The second number is an implied probability after standardisation (removing the bookmaker margin). Note that we say "an" implied probability, not "the" implied probability: there are an infinite number of different methods available for removing the bookmaker margin (as will be explored shortly). Here, Basic Standardisation (see be-

	Outcome 1	Outcome 2
Bookmaker 1	1.95	1.95
	0.5	0.5
Bookmaker 2	1.95	1.85
	0.4868	0.5132
Bookmaker 3	2.05	1.8
	0.4675	0.5325

Table 1: Fictional set of odds and standardised implied probabilities for a two-outcome, three-bookmaker scenario.

low) was used.

Quasi-arbitrage considers an aggregate of the available opinions, and weighs these against the most exposed opinion, capitalising if this discrepancy is above a certain threshold. In this case, if we take the arithmetic mean of the implied probabilities for each outcome, and see these as the event's true probability distribution, then we get that $p_1 = 0.4848$ and $p_2 = 0.5152$. In terms of expected value, against the longest odds for each outcome, we get $EV_1 = 0.4848 \cdot 2.05 - 1 = -0.0062$ and $EV_2 = 0.5152 \cdot 1.95 - 1 = 0.0046$. EV_2 is greater than zero, meaning that there is a quarb opportunity betting with Bookmaker 1 on Outcome 2. This is the family of bets that will be explored below in the development of our betting model.

5 Odds Standardisation and Information Extraction

The purpose of this section is twofold: first, to evaluate which of several methods most effectively standardises bookmaker implied probabilities; and second, to determine how to most effectively aggregate the collective bookmaker opinions. Quasi-arbitrage leverages a wisdom-of-the-crowds effect where an aggregated bookmaker opinion is compared to the longest, must vulnerable opinion. Using a framework presented by Štrumbelj (2014), we optimise the manner in which the aggregated opinion is constructed, an essential component of any quarb strategy.

5.1 Ranked Probability Score and Odds-Standardisation Techniques

Using the notation of Štrumbelj, let $\boldsymbol{o} = (o_1, o_2, \ldots, o_m)$ for $m \geq 2$ outcomes be the decimal odds set by a bookmaker for a particular event. Then $\boldsymbol{\pi} = \frac{1}{\boldsymbol{o}} = (\pi_1, \pi_2, \ldots, \pi_m)$ are the implied probabilities before standardisation for the m outcomes, with the bookmaker's total take $\Pi = \sum_{j=1}^{m} \pi_j > 1$. Alternatively, $\boldsymbol{p} = f(\boldsymbol{\pi}) = (p_1, p_2, \ldots, p_m)$ are the standardised implied probabilities, derived via some standardisation function f such that $\mathbf{1}^{\boldsymbol{T}}\boldsymbol{p} = 1$.

We will in this section evaluate multiple such standardisation functions f, determining which extracts the most outcome-relevant information from the original odds. Using the Ranked Probability Score (RPS) (Epstein, 1969), these different methods will be evaluated across the dataset, where the RPS score is calculated

$$RPS = \frac{1}{m-1} \sum_{k=1}^{m-1} (F_k - O_k)^2 \tag{1}$$

where F_k is the cumulative probability forecast for all categories indexed j up to and including k, and O_k is the cumulative outcome for all categories up to and including k, where $O_j = 1$ if the jth category is realised and 0 otherwise. The RPS score is at most 1 and at least 0, with 0 being optimal performance.

We use RPS rather than the Brier score (Brier, 1950), because for a sport like soccer it is worse to be wrong going from a home win to an away win than from a draw to an away win, an ordinal dynamic which RPS captures and the Brier Score does not. For instance, consider the probabilistic forecast $\boldsymbol{p} = (0.79, 0.09, 0.12)$ for a three-outcome event. The forecast is heavily weighted towards the first outcome (1,0,0) but for distinguishing between the Brier Score and RPS (measured on the same scale), consider each of the other two outcomes. For the second outcome (0,1,0), the Brier score of 0.489 is higher than the RPS of 0.319, while the reverse is true for the third outcome (0,0,1), where the RPS of 0.699 is higher than the Brier score of 0.469. The RPS recognises that the forecast heavily favours the first outcome, and that the third outcome is further away from the first than is the second (in soccer terms, a draw is at least

one goal removed from a home win, while an away win is at least two goals removed). The Brier score fails to account for this ordinal phenomenon (merely assigning a lower score to the higher probability forecast), which is why RPS – which, like the Brier Score, is a strictly proper scoring rule, in that a forecaster optimises their expected score by providing their true forecast – is chosen for this analysis.

The appendix (subsection "Odds-Standardisation Techniques") provides a detailed explanation of each of the seven methods evaluated here. Three of these (Basic, Shin and Logit) were compared by Štrumbelj, with the remaining four being our additions. Of these seven mechanisms, three are common in the betting literature (Basic, Shin, Exponential); of the other four, one is a contribution of Štrumbelj (Logit), with the remaining three our own suggestions (Price Difference, Discrete Profit, Continuous Profit). Three of these standardisations require parameters to be fit (Logit, Discrete Profit, Continuous Profit), three require a self-contained numerical solution (Shin, Exponential, Price Difference), while the Basic is solved with a simple vector operation.

5.2 Standardisation Comparison

The dataset covers 22 leagues and thirteen bookmakers. In the vein of Štrumbelj we seek to evaluate the different methods on bookmaker-league pairs, each consisting of all the odds provided by a single bookmaker for a single league. Taking only those pairs with at least 1,000 games, we find 242 pairs (two of the thirteen bookmakers, SO and SY, have no such pairs, leaving 22 pairs for each of the remaining eleven bookmakers). The first third of data for each pair, separated chronologically, is allocated towards training the regression-based standardisation methods: the logit model and the profit models. On the final two thirds of each pair, a method is given a RPS for each game, across which we then take the arithmetic mean. We thus have 242 data points, each data point being seven RPS.

	mean	median
Basic	0.2044	0.2064
Shin	0.2042	0.2062
Logit	0.2051	0.2066
Exponential	0.2042	0.2062
Price Difference	0.2042	0.2063
Discrete Profit	0.2047	0.2066
Continuous Profit	0.2046	0.2069

Table 2: Mean and median Ranked Probability Score for each standardisation method.

	Basic	Shin	Logit	Exponential	Price Difference	Discrete Profit	Continuous Profit
Basic	1						
Shin	0.001	1					
Logit	0.001	0.001	1				
Exponential	0.001	0.3134	0.001	1			
Price Difference	0.001	0.9	0.001	0.1577	1		
Discrete Profit	0.001	0.001	0.0029	0.001	0.001	1	
Continuous Profit	0.9	0.001	0.001	0.001	0.001	0.001	1

Table 3: Nemenyi test p-values, where each p-value corresponds to a pair-wise comparison of two methods, with the null being statistically-similar performance.

For an initial overview of performance for the different methods, we take the mean and median for each method, as shown in Table 2. Performing the Friedman test, the non-parametric

analysis of variance test, with 6 degrees of freedom (having seven models), we observe a test statistic of 608.564 and a p-value of 3.31×10^{-128}, where the null is thus rejected, meaning that there exists at least one method with significantly different performance from another.

Following the successful Friedman test, Table 3 shows the Nemenyi test (Nemenyi, 1963; Demšar, 2006), a pairwise comparison, where the null hypothesis is that the two methods have performed the same. As can be seen, the Exponential, Shin and Price Difference methods are statistically similar, as are the Basic and Continuous-Profit techniques. To visualise the different strategies further, in Figure 1 we display the mean rank of each method (for each bookmaker-league pair, each method has an RPS: we rank these mean RPS for each pair, then take the mean of these ranks), with the critical distance CD (see Demšar) for pairwise comparison, where CD is calculated:

$$CD_{\alpha,K,N} = q_{\alpha,K}\sqrt{\frac{K(K+1)}{6N}} \tag{2}$$

where α is the confidence level (in this case, 0.01), N is the number of observations (242), K is the number of models, and $q_{\alpha,K}$ is the Studentised range statistic (Student, 1927) at the given confidence level and number of models.

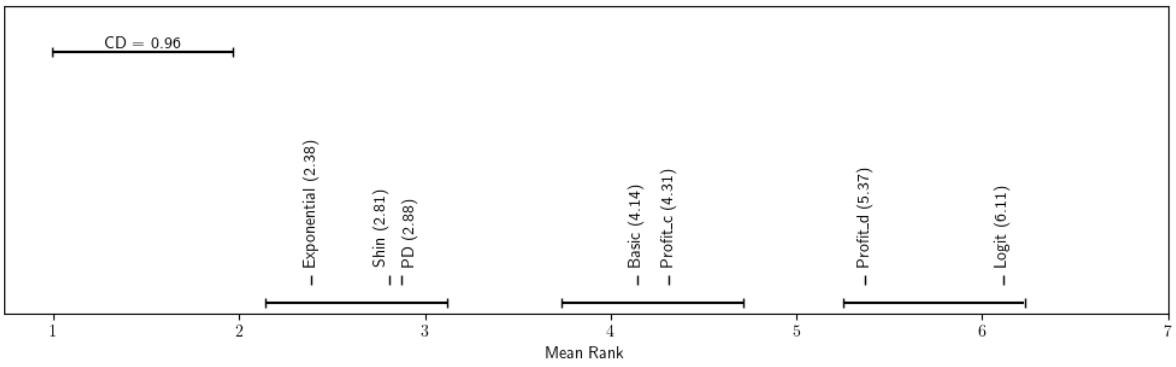

Figure 1: Mean ranks and critical distance. Methods connected by a single line have statistically performed the same. Values in brackets are that method's mean rank.

The relationships that were found from the Nemenyi test are not quite observed here: Exponential, Shin and Price Difference are familiarly in one cluster, with Basic and Continuous Profit in another; however, Discrete Profit and Logit perform the worst, and do so statistically equivalently, which was not the case according to the Nemenyi test. This difference arises due to the Nemenyi test being a pair-wise comparison, considering the distributions of ranks for each method, while the critical distance test is superficial, giving one aggregate-level test-statistic for comparison without evaluating the nuances of each method. Interestingly, the three regression-based models are inferior to the zero-shot methods (arguably not for Continuous Profit). Additionally, the three methods which require some manner of optimisation with each use (finding z for Shin, k for Exponential, and pd for Price Difference, see appendix, subsection "Odds-Standardisation Techniques") perform better than Basic (statistically), which requires merely one element-wise vector operation, $\boldsymbol{p} = \frac{\boldsymbol{\pi}}{\sum \boldsymbol{\pi}}$. Most importantly, we can say that any of these three methods are a prudent choice.

5.3 Bookmaker Comparison

Returning to replicating Štrumbelj's results, we now compare the forecasting abilities of different bookmakers. Across the entire dataset, we searched for the largest sample of games covered

by the same group of bookmakers (not all bookmakers set odds for all games in all leagues). One such collection had 35,290 games where ten of thirteen bookmakers set odds for each game. This was actually the second largest independent sample of games upon which multiple bookmakers uniformly set odds, the largest consisting of 35,877 games with six bookmakers. Table 4 shows the mean RPS for each method, with each bookmaker.

	B365	BS	BW	GB	IW	LB	SB	SJ	VC	WH	Mean
Basic	0.2048	0.2053	0.2053	0.2052	0.2058	0.2056	0.2052	0.2052	0.2047	0.2052	0.2052
Shin	0.2046	0.205	0.2051	0.2049	0.2054	0.2053	0.205	0.2051	0.2046	0.2049	0.205
Logit	0.2049	0.2053	0.2054	0.2051	0.2056	0.2056	0.2052	0.2054	0.2052	0.2055	0.2053
Exponential	0.2046	0.205	0.205	0.2049	0.2053	0.2053	0.205	0.205	0.2046	0.2049	0.205
Price Difference	0.2046	0.205	0.2051	0.2049	0.2053	0.2053	0.205	0.2051	0.2046	0.205	0.205
Discrete Profit	0.2047	0.205	0.2051	0.2049	0.2053	0.2054	0.205	0.2051	0.2049	0.205	0.205
Continuous Profit	0.2046	0.205	0.2051	0.2049	0.2053	0.2053	0.205	0.2051	0.2046	0.2049	0.205
Mean	0.2424	0.2428	0.2428	0.2426	0.2431	0.243	0.2428	0.2428	0.2425	0.2428	

Table 4: Ranked Probability Score for each probability-standardisation method, with each bookmaker, across the 35,290 games shared by all ten bookmakers. Arithmetic mean by bookmaker and method.

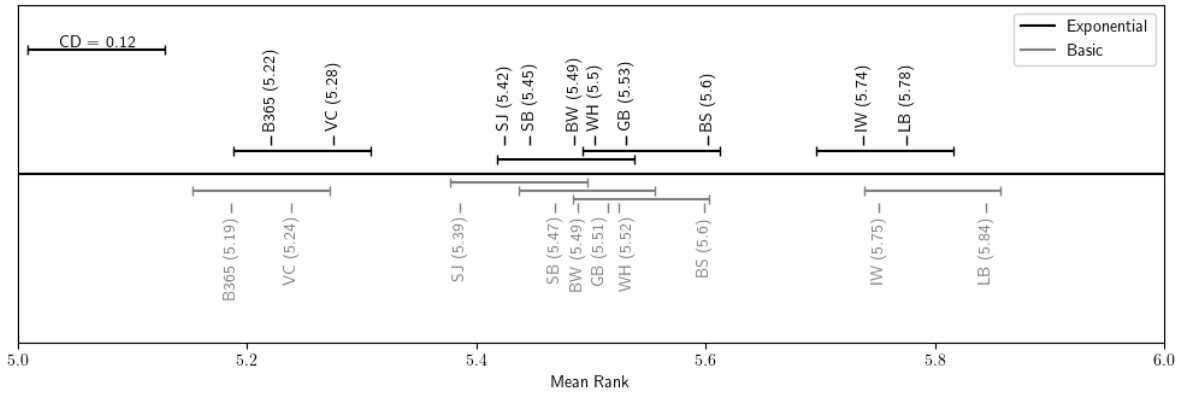

Figure 2: Mean rank for each bookmaker across a shared 35,290-game sample, with Exponential standardisation (above) and Basic standardisation (below).

Following Štrumbelj, we focus on two methods, Basic and Exponential (Exponential rather than Shin, due to the (marginal) out-performance of the former). The Friedman test is highly significant for both methods (p-value 4.92×10^{-228} and 3.30×10^{-305} for Exponential and Basic, respectively), meaning that, applying each method to the odds from each individual bookmaker, there is at least one bookmaker with significantly different forecasting performance than another. Figure 2 displays the mean ranks for each bookmaker, with each method, with critical distance. There is a clear trend, with all bookmakers having the same order of mean ranks for both methods, other than GB and WH, who swap places in sixth and seventh. B365 and VC are clearly much stronger forecasters than the rest, while IW and LB are significantly worse than their peers. Štrumbelj showed that bookmaker forecasting performance was strongly correlated with the amount of margin they apply to their odds. As such, Figure 3 plots, in the upper panel, RPS rank versus margin rank with Kendall's Tau (Kendall, 1938) the ranked correlation coefficient, while the lower panel shows mean RPS versus mean margin, accompanied by Pearson's correlation coefficient. The lower the margin a bookmaker sets, the more improved will be

their forecasting. This makes sense: the RPS measures the informational content of probability forecasts which have been deciphered by removing margin from unstandardised implied probabilities; the more margin a bookmaker applies, the weaker will be the signal between their true probabilities and their implied probabilities.

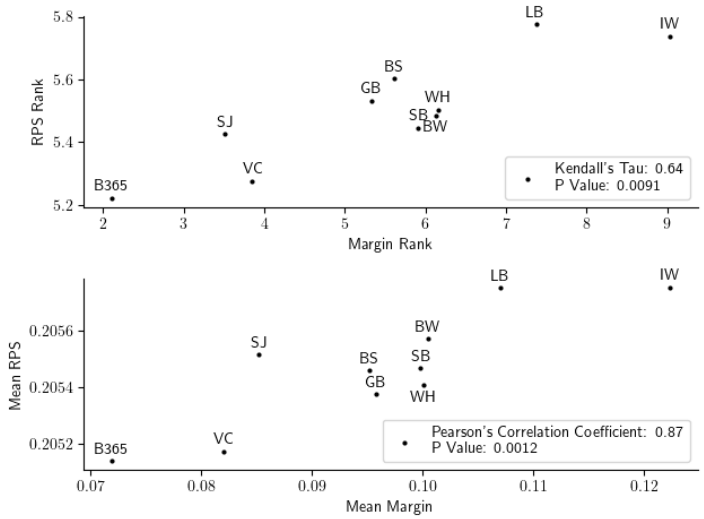

Figure 3: Correlation between forecasting performance (in terms of RPS) and margin, with respect to rank (upper) and mean (lower). There is a strong correlation in both cases. With Exponential standardisation.

We seek a method of aggregating bookmaker opinions into an optimal forecast. The most basic such would be the mean forecast: standardise implied probabilities for all bookmakers providing odds on an event, then take the mean of these. This is the most conventional method, used both by Smith et al. (2009) and (to an extent) Kaunitz et al. Inspired by Figure 3, we propose two margin-based methods. First, the weighted-mean opinion, where instead of giving all bookmaker standardised probabilities equal weight, we weigh their forecasts according to the inverse of their margin (lower margin means superior RPS, so we take the inverse to assign higher weight to these bookmakers).

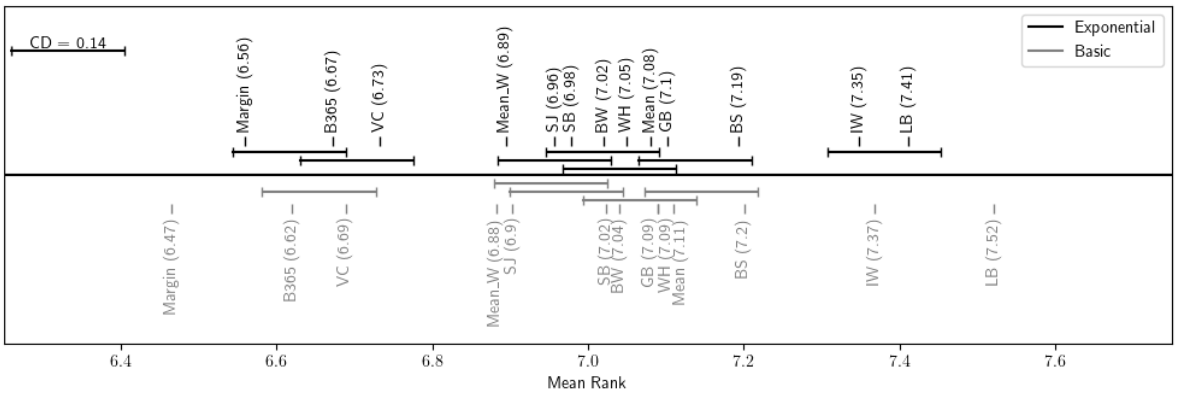

Figure 4: Bookmaker and aggregation mean ranks: Mean, Mean_W, Margin. The two mean aggregations do no better than the bulk of bookmakers. The lowest-margin aggregation, however, is the best with both methods, although not significantly so versus B365 with Exponential standardisation.

Second, for each bet, simply take the bookmaker with the lowest margin for that match and standardise their probabilities. Figure 4 shows the mean ranks from Figure 2 for each bookmaker, adding these three aggregations (Mean, Mean_W and Margin, respectively) as well. Clearly, the margin aggregation performs the best. By selecting, for a given event, the bookmaker with the lowest margin and standardising their probabilities, this gives the most efficient

possible probabilistic forecast for the event outcome. This will be our aggregation moving forward: Exponential standardisation coupled with Margin aggregation. Note that this combination did not outperform Exponential standardisation with B365 statistically, however, its mean rank was superior, while it is also applicable for matches where B365 does not offer odds.

6 Naive Quasi-Arbitrage

In an instructive analysis, Kaunitz et al. demonstrate the great potential of Quasi-arbitrage by generating very strong returns from a rather rudimentary strategy. They do not avail themselves of any probability standardisation method, instead, simply subtract a constant and use the remainder as their forecast: $\boldsymbol{p} = \boldsymbol{p_c} - \alpha$, where c stands for consensus and $\boldsymbol{p_c} = \frac{1}{\overline{\boldsymbol{o}}}$, where they use a mean aggregation of the odds for the outcomes, $\overline{\boldsymbol{o}}$. This betting strategy is executed whenever the expected value is positive, calculated as $(p_{cj} - \alpha)max(\boldsymbol{o_j}) - 1 > 0$, where p_{cj} is the consensus probability for outcome j, α is some optimised constant and $max(\boldsymbol{o_j})$ selects the longest odds from the cross-section of odds $\boldsymbol{o_j}$ for outcome j. This rearranges to the condition

$$max(\boldsymbol{o_j}) > \frac{1}{p_{cj} - \alpha} \tag{3}$$

Note that if multiple outcomes fulfill this condition for a single match, then only one bet is placed on that outcome with the highest expected value. Figure 5 shows the optimisation rationale for α.

As did Kaunitz et al., we apply linear staking, meaning that each bet is worth a unit, which is lost in full if the bet loses, while a profit of $o - 1$ is made when the bet wins with odds o. r is the mean per-bet return, which increases in α, which is reasonable: as α increases, this increases the minimum value with which $max(\boldsymbol{o_j})$ exceeds the threshold set in equation (3). As such, we become increasingly selective, as can be seen in the second panel, where the number of bets n made in the training set decreases in α. Total returns will thus equal $n \cdot r$, since stakes are linear. We choose the value for α which maximises $n \cdot r$ in the training set, since this metric is the optimal balance between getting higher returns on each bet (r) but also ensuring that sufficiently many bets are placed (n). The optimal value, here, is 0.6225. The out-of-sample performance of this model when applied to the test dataset is shown in Figure 6.

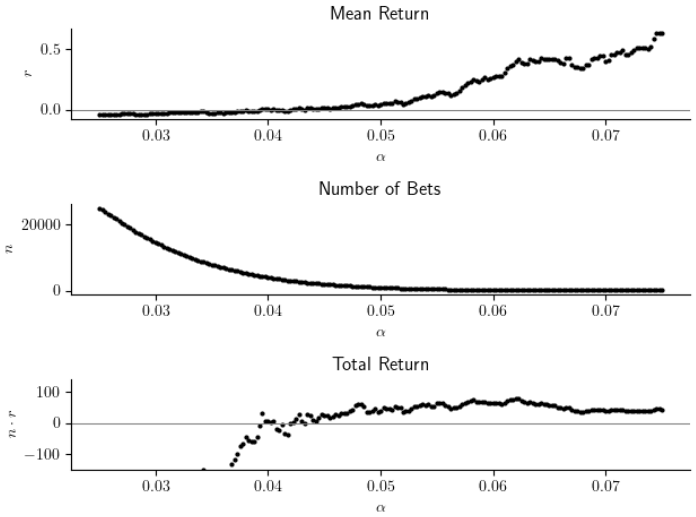

Figure 5: Selecting α. r is the per-bet linear return (from unit staking), while n is the number of bets. Choose the α which maximises the product $n \cdot r$

The black line shows how profit accumulated at the ith bet, π_i. The grey lines are 2,000 simulated runs with the same number of bets as the true iteration. For each of these runs, 129 games were randomly selected from the test dataset; for each game, an outcome was randomly

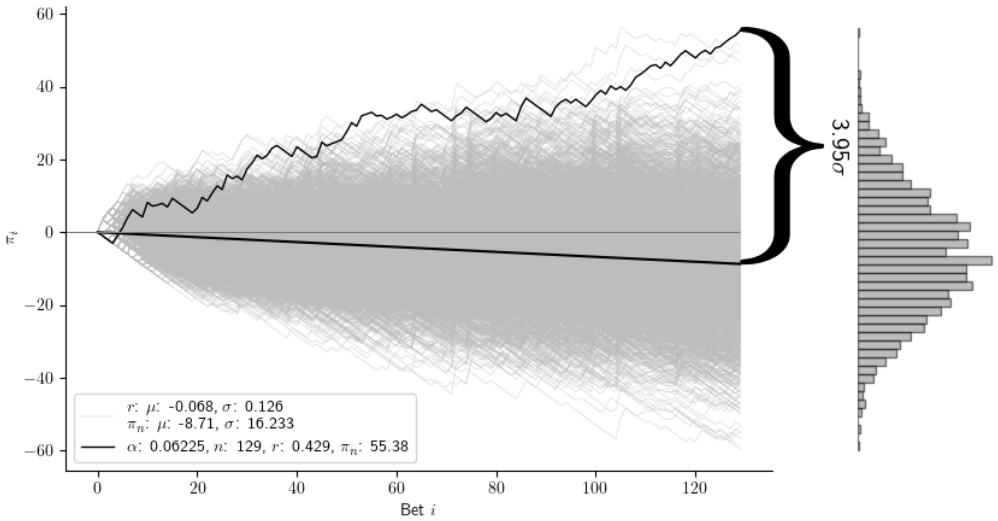

Figure 6: Out-of-sample profit development for the Kaunitz et al. model. Compared to a random-betting simulation of 2,000 runs. With model parameter α, r is the per-bet return, with π_n the total profit after n unit bets. μ and σ are mean and standard-deviation for the simulation.

selected, with probabilities equal to the base rate at which the different outcomes occur in the training dataset. The perpendicular histogram shows the distribution of final accumulated profit for these simulated runs, allowing us to calculate the number of standard deviations above randomness the true run performed, in this case, 3.95.

Merely 129 games are selected for betting, a bet selection rate of 0.12%. However each bet had an arithmetic return of 42.9%, on average, giving a total profit of 55.38 units. Kaunitz et al. chose $\alpha = 0.05$, giving a bet selection of 11.77% (56,435/479,400) and a per-bet return, arithmetically, of 3.5%. With these two values, Kaunitz et al. would have bet on 0.1177·111,021≈13,067 games, giving a total profit of 0.035·13,067=457.35, clearly much more than we achieve here. We suspect that this difference in performance is due to quality of data, where they "analysed 479,440 games from 818 leagues and divisions across the world during the period 2005-2015", while we analyse 166,532 games from 22 leagues during the period 2001-2023. Clearly, they accessed much lower quality leagues, potentially with less efficient pricing. Therefore, it is encouraging to note that their methodology remains extremely successful (although not to the same extent) on a more qualitatively sound dataset. These quasi-arbitrage profits are encouraging, especially given the lack of sophistication in this one-parameter, one-calculation model.

We elaborate with the following amendment:

$$max(\boldsymbol{o_j}) > \frac{1}{p_{cj} - \alpha_j} \tag{4}$$

assigning an α_j parameter to each outcome j, as it seems reasonable that probabilities are not extracted equally for each outcome. These parameters are optimised using the same framework as for α in equation (3). Returns to this model, as with Figure 6, are shown in Figure 7. Customising the α_j parameters leads to significant improvement, where total profits almost double compared to the original one-parameter model. Most encouragingly, the bet selection increases drastically, from 0.12% to 0.96%, while per-pet returns remain high at 10.4% linearly. Total profits increase by a factor just under two. One final point here is that there seems to exist an

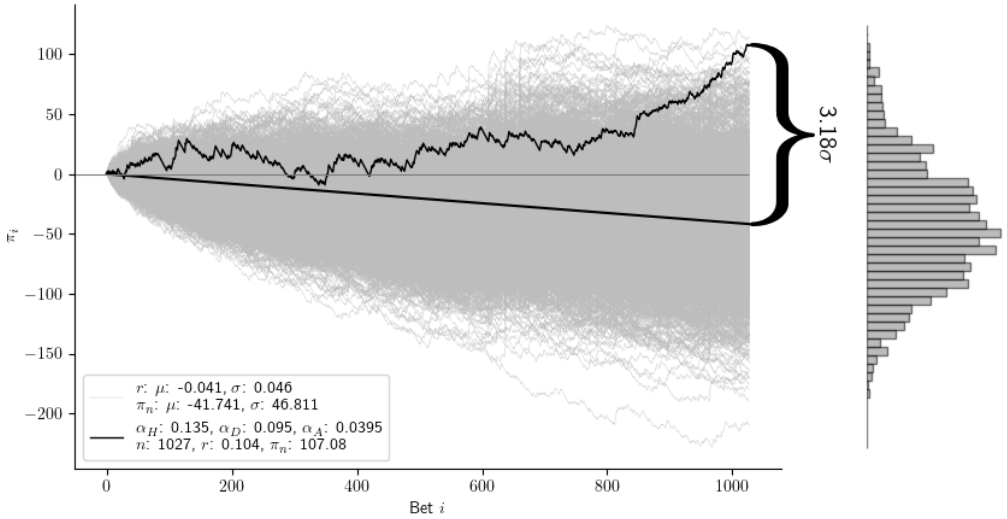

Figure 7: Returns to Kaunitz et al. model, with an α parameter for each outcome. r is the per-bet return, with π_n the total profit after n unit bets. μ and σ are mean and standard-deviation for the 2,000-run random-bet simulation.

away bias in the odds, as the overwhelming majority of bets for the multi-α model are with away: 5, 0 and 1023 bets at home, draw, and away, respectively, for the three-parameter model. This is fascinating, as per the favourite-longshot bias – where returns decrease as odds increase – and soccer outcome base rates – where away win is the least common outcome, thus, on average, possessing the longest odds – the mean returns per outcome are the worst for away: across the entire dataset, against mean odds, the mean return for home, draw and away, respecitvely, is -0.0711, -0.0966 , -0.1171; while for longest odds they are -0.0302, -0.0539, -0.0542. It would seem that although these bets to away odds are indeed worse on average, there also exist lucrative opportunities that can be readily identified with basic quarb methods. Indeed, this section has demonstrated the ease with which profits can be generated using rudimentary quasi-arbitrage. In the next section, we show how a more sophisticated model can generate even larger returns.

7 Quasi-Arbitrage Betting Model

For a match with m outcomes, with each of n bookmakers indexed i providing odds $\boldsymbol{o_i} = (o_{ij} \mid j = 1, \ldots, m)$ for each outcome j, our probability forecast for this match is

$$\boldsymbol{p} = \left(\frac{1}{o_{ij}^k} \;\middle|\; j = 1, \ldots, m \right) \tag{5}$$

where k is determined such that $\mathbf{1}^{\boldsymbol{T}}\boldsymbol{p} = 1$, and i is determined according to

$$\arg\min_i \mathbf{1}^{\boldsymbol{T}} \left(\frac{1}{\boldsymbol{o_i}} \right) \tag{6}$$

We also extract

$$\boldsymbol{o_l} = (max(\boldsymbol{o_j}) \mid j = 1, \ldots, m) \tag{7}$$

the set of longest odds provided, where $\boldsymbol{o_j}$ is the cross section of odds offered on the jth outcome. The set of expected values for the n outcomes is then $\boldsymbol{ev} = \boldsymbol{p} \odot \boldsymbol{o_l} - \mathbf{1}$, where $\odot$ signifies element-wise multiplication, while the set of price differences is

$$\boldsymbol{pd} = \left(ln\left(\frac{1}{1-p_j}\right) - ln\left(\frac{o_{lj}}{o_{lj}-1}\right) \;\middle|\; p_j \in \boldsymbol{p}, o_{lj} \in \boldsymbol{o_l}, j = 1, \ldots, m \right) \tag{8}$$

Here price difference corresponds to the distance measure in the Price Difference standardisation technique discussed above; see equation (35) in the appendix subsection "Odds-Standardisation Techniques" for greater detail. We consider betting on outcome j at odds o_{lj} (the longest odds on outcome j) if pd_j (the price difference for the jth outcome) is $pd_j > 0$, where j is the outcome with the largest pd. If this condition is fulfilled, we would stake a proportion k of our current capital, according to

$$k = \frac{(p_j o_{lj} - 1)^3}{(o_{lj} - 1)((p_j o_{lj} - 1)^2 + o_{lj}^2 \sigma_j^2)} \tag{9}$$

where p_j is the probability from $\boldsymbol{p}$ for outcome j. This expression is per Baker and McHale (2013), and is a derivation of the conventional Kelly Criterion (Kelly, 1956), adjusting the stake downwards due to uncertainty in the probability estimate p_j, given by σ_j^2. The Kelly Criterion, in general, is to stake a fraction

$$k^* = \frac{po - 1}{o - 1} \tag{10}$$

of current capital on an investment opportunity with probability p and odds o, maximising the geometric growth rate of capital. Numerous sources (MacLean et al., 1992; Maclean et al., 2005; MacLean et al., 2011; Davis and Lleo, 2011) recognise the benefits of Fractional Kelly (betting less than the full Kelly-Criterion fraction k^*), primarily to reduce returns volatility in the short to medium-term. Baker and McHale's method accomplishes this by reducing the stake as the uncertainty in the estimation of p_j increases. Here, we calculate

$$\sigma_j^2 = c\frac{\sigma_{\boldsymbol{o_j}}}{\mu_{\boldsymbol{o_j}}} \tag{11}$$

where c is a constant, and $\sigma_{\boldsymbol{o_j}}$ and $\mu_{\boldsymbol{o_j}}$ are the standard deviation and mean, respectively, of $\boldsymbol{o_j}$. For an outcome j, the uncertainty is thus the coefficient of variation, scaled by c, which is optimised such that, in the training set, for all outcomes where $ev_j > 0$ (this can be multiple per match), the average fractional Kelly ratio is 0.5, where this ratio is calculated as

$$\frac{k}{k^*} = \frac{(p_j o_{lj} - 1)^2}{(p_j o_{lj} - 1)^2 + c\frac{\sigma_{\boldsymbol{o_j}}}{\mu_{\boldsymbol{o_j}}} o_{lj}^2} \tag{12}$$

That is, we set c such that, for all bets that would have been candidate bets pending further filtering, on average we would bet half of what the full Kelly Criterion would suggest. We find that $c = 0.00147$, while the mean coefficient of variation among these bets was 0.048. On average, then, Baker-McHale's $\sigma_j^2 = 0.00007056$ where, although perhaps not particularly relevant in a vacuum, the low order of magnitude demonstrates the extreme sensitivity of these Kelly fractions to uncertainty in the probability estimate. As was the case in the Štrumbelj and Kaunitz et al. analyses above, the training set is the first third of the entire dataset, with the second two thirds reserved for backtesting. Typically, larger portions of data are allocated towards training, however, we feel that the emphasis in this case should lie in the performance of the trading model out-of-sample, with the generous allocation of data to these tests enabled by the large quantity of data in the dataset.

Inspired by Kaunitz et al., we filter these candidate bets according to

$$pd_j - \alpha > 0 \tag{13}$$

here α is optimised in the training data in the same manner as was done for the first model in the Kaunitz et al. replication, as seen in Figure 8.

We see that returns are positive at all levels of α, even 0, meaning that the simple $pd_j > 0$ rule is sufficient for reasonable profits. However, the objective $(1+r)^n$ does improve, peaking at $\alpha = 0.4575$, with a cumulative return of 360.28, starting from 1. Applying this value of α to the test dataset, capital grows to 3,959,583.13 over 3,738 bets, with a per-bet (geometric) return of $r = (3{,}959{,}583.13)^{\frac{1}{3{,}738}} - 1 = 0.407\%$ and bet selection of $s = \frac{3{,}738}{111{,}021} = 3.367\%$. We perform the same random-betting simulation as was done for Kaunitz et al., creating 2,000 fictional runs of 3,378 bets taken randomly from the test dataset (with replacement), the outcome for each bet chosen according to the training-set base rate. As seen in equations (9) and (10), Kelly fractions are not constant in odds o. Therefore, for the random simulations, we select Kelly fractions from the distribution of fractions from the true run (see Figure 9), where the chosen fraction corresponds to the odds from the true run that is closest to the odds at which the random bet is placed. For instance, if a random run places a bet at odds 3.45, we find the odds from the true run closest to 3.45 and select the Kelly fraction that was bet at these odds in the true run. The distribution of final capital for the random simulation is shown in Figure 10.

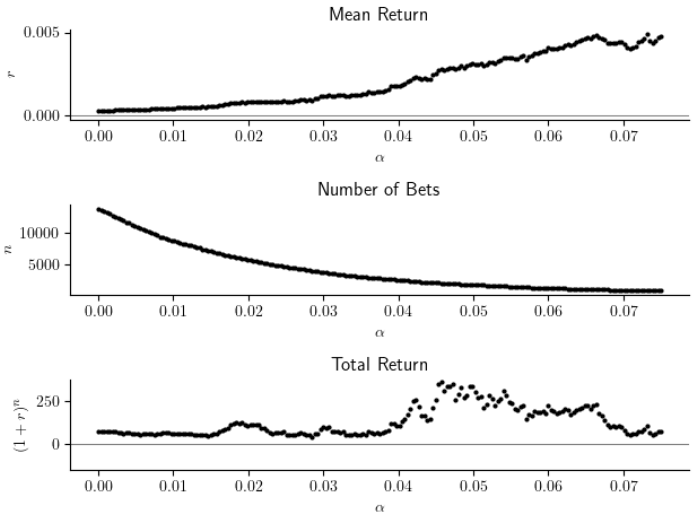

Figure 8: Selecting α. With geometric staking, the objective is the geometric cumulative return $(1+r)^n$, where r is the per-bet geometric return.

The random strategy does extremely poorly, going bankrupt without failure. The least ruinous run had final capital of 8.25×10^{-5}, which corresponds to a per-bet return of -0.251%, while the worst run loses 1.442% per bet. Not disastrous values (-1.442% as a worst case is better than one might anticipate), however, obviously enough to completely eviscerate any investment capital, given the geometric Kelly staking. This, then, is strong proof of the significance of the true-model's returns, which instead rose to 3,959,583.13 over the same period. In terms of standard deviations above the mean, we can look instead at per-bet returns, which have mean (and approximately median) of -0.00842 and standard deviation -0.00181, making the per-bet return of 0.0041 for the true run 6.92 standard deviations above the mean.

Continuing the rationale from the Kaunitz et al. replication, we develop the following thresholds for a bet to be accepted

$$pd_j - (\alpha + \beta p_j) > 0 \tag{14}$$

$$pd_j - \alpha_j > 0 \tag{15}$$

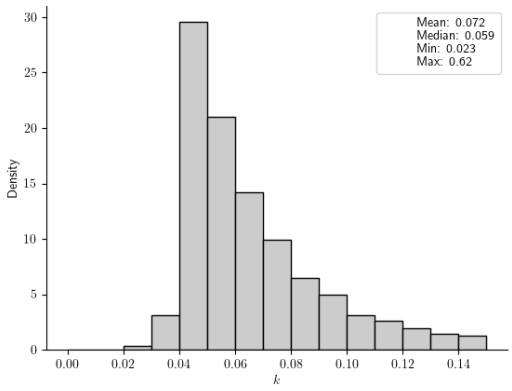

Figure 9: The distribution of Baker-McHale Kelly fractions k with which the one-parameter model's bets were placed.

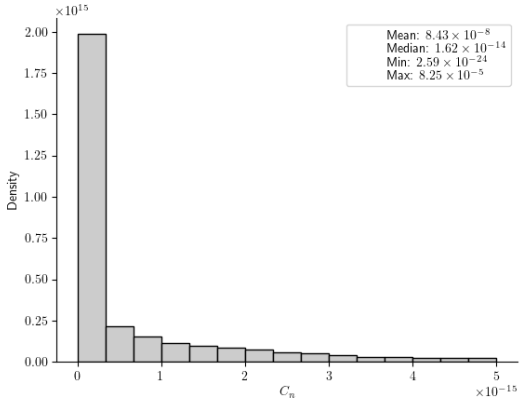

Figure 10: The distribution of final capital (starting at 1) for 2,000 random runs, with 3,378 bets.

$$pd_j - (\alpha_j + \beta_j p_j) > 0 \tag{16}$$

with two, three and six parameters, respectively. The β parameters alter the pd threshold based on probabilities; in theory, this should be redundant, as the price-difference mechanism (see equations (8) and (35)) should account for such probability-level effects. We optimise for these parameters by maximising $(1 + r)^n$ in the training data. The results for the four models are shown in Figure 11.

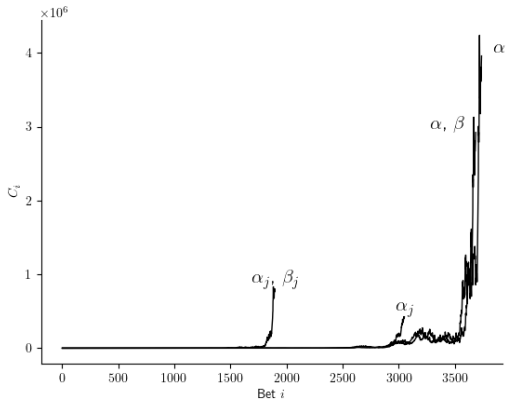

Figure 11: Out-of-sample performance for four models. The plot labels indicate which parameters are included in the model. j indicates that this parameter has been trained for each outcome.

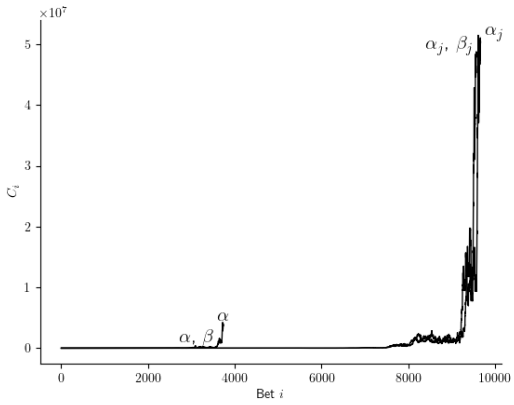

Figure 12: Out-of-sample performance for four models following bootstrapped training, having maximised the median return across ten 75% samples, with replacement.

All four models do very well. Surprisingly, with the exception of the two worst models, performance decreases in number of parameters, as the one-parameter model has the highest final capital, followed by the two-parameter, six-parameter and, finally, three-parameter models. However, the per-bet returns increase in parameters (again, other than for the two inferior models). This suggests that the training process has failed to optimise for total returns $(1 + r)^n$,

perhaps overfitting the parameters to the training data, giving increasingly high per-bet returns, but selecting fewer matches as a result. As such, we perform a bootstrapped optimisation, where we create ten samples of 75% of the data in the training set (with replacement), and maximise the median total returns of these ten samples. The rationale is that by providing incomplete data and optimising a lower percentile, the parameters will not overfit to the particular bets available in the training set. Having fit the same four models on the training set, Figure 12 shows the results when applied to the test set.

Usefully, the single-parameter, α model has the same value following both optimisations, so its capital path is the same in both Figures 11 and 12. This gives some reference to the improvement in the outcome-customised models, where now both the three and six-parameter models give returns around 50,000,000 (the three-parameter model just above, the six-parameter just below). Across both sets of models, the β parameters provide very little value, with the final results basically equal between the three and six-parameter models, as well as between the one and two-parameter models, with the α-only models better (with the exception mentioned above). Being, also, more computationally practical in general to optimise half as many parameters, we therefore select α_j, the three-parameter model, as our quasi-arbitrage optimal model. Table 5 presents some key metrics for the model across the test data.

n	s	C_n	r	$\overline{o}$	Acc	σ		H	D	A
9,668	8.71%	50,416,091.95	0.184%	2.712	0.489	7.60	α_j	0.05	0.05	0.01
							n_j	2333	26	7309
							$\overline{o}_j$	1.60	3.49	3.06
							Acc_j	0.683	0.385	0.428

Table 5: Optimal Quasi-Arbitrage Model: key out-of-sample metrics for the three-parameter, multi-α model, namely, the number of bets, bet selection rate, final capital, per-bet geometric return, average odds, accuracy, and standard deviations above the mean random per-bet return; and then for each outcome j, the optimal α value, number of bets placed at that outcome, average odds for that outcome and accuracy for that outcome.

As was done for the Kaunitz et al. models and then the non-bootstrap-trained, single-α model, we simulate 2,000 random runs of the 9,668 bets, taken from the test dataset (with replacement), drawing Kelly fractions from those seen in the main run as described above. Across these runs, the mean final capital was 0.00924 (from an initial 1), while the median final capital was 0.0000137. Two runs gave positive returns, with the maximum final capital found in the simulation being 5.09. In terms of per-bet returns, Figure 13 displays this distribution for the random simulation. The mean (and median, approximately) is -0.00115, while the standard deviation is 0.00039. Comparing to the distribution of returns for the non-bootstrap-trained, single-α model above, returns are higher to the random simulation here, despite each run consisting of more bets (9,668 rather than 3,378). This is because the mean Kelly fraction from the true run bets here is lower: 0.0357 rather than 0.0725. For the true run, the per-bet returns were 0.00184, 7.60 standard deviations above the mean random per-bet return, better than the noted performance of the non-bootstrap-trained, single-α model, which was 6.92 standard deviations above its corresponding random simulation. Clearly statistically robust, the edge is sizeable: from Table 5, the average odds is 2.712, while the model selects winning bets 48.9% of the time. Although it is a naive metric in this instance, this gives $ev = 0.3262$; it is naive because the mean odds, obviously, but also accuracy are aggregates: more relevant is the nuanced interaction between individual probabilities and odds. Also of note from Table 5 is the number

of bets placed at each outcome. Although not quite as extreme as for the Kaunitz et al. model above, the away bias identified there remains, as we place more than three times as many bets at away as we do home and draw combined. That draw only sees 26 bets across the entire 111,021-game test dataset is surprising. Especially considering that $\alpha_H = \alpha_D = 0.05$. Clearly, seeing price differences above this magnitude is more common for home odds than draw odds.

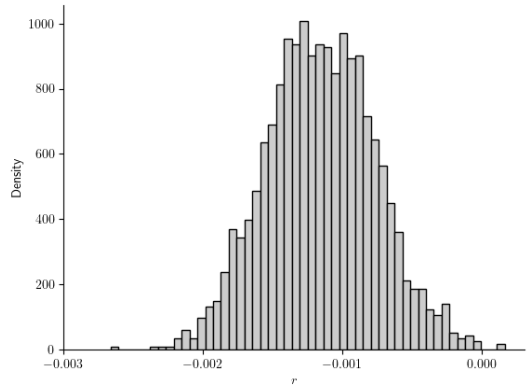

Figure 13: Distribution of per-bet returns for 2,000 random runs of length 9,668.

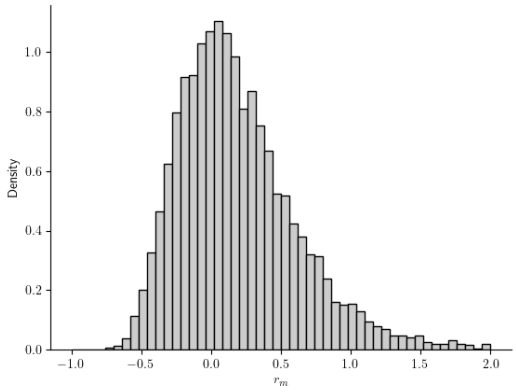

Figure 14: Distribution of r_m, monthly returns, from 10,000 simulated 69-bet runs.

The model is successful. But the returns were gained over more than ten years' worth of betting. Consider therefore the following heuristic argument, meant only as a loose approximation to a more understandable unit. There are 22 seasons in the dataset, of which the test set is the latter two thirds. Make this fourteen seasons, each ten months long (see Table 21). We thus have $10 \cdot 14 = 140$ months, giving $9{,}668/140 = 69.06$ monthly bets. If the per-bet return is made across such an average month, the accumulated return would be $(1+0.00184)^{69.06}-1 = 0.1354$, just above 13.5% per month. We simulate 10,000 69-game samples from the 9,668 bets selected by the model (with replacement), with the results shown in Figure 14. With positive skew, the mean, median and standard deviation monthly return is 0.211, 0.124, 0.497, respectively. Only 36.66% of months gave negative cumulative returns. Over a more familiar period, the model performs remarkably well, for what the approximation here is worth.

One final aspect worth inspecting is whether the overall returns are supported by a small class of extremely profitable bets, while the majority of the bets are actually unprofitable, but at a lower scale, and hence unnoticeable. Figure 15 plots the per-bet returns to all bets with k equal to or above a threshold k_t, while Figure 16 corresponds to bets below the threshold. Considering Figure 15 first, the more selective one gets, choosing only bets with $k \geq k_t$, returns improve. The upper panel closely resembles the upmost panels in Figures 5 and 8. The lower panel here similarly resembles the middle panels in those figures. Acknowledging that these bets have already been filtered by pd, k thus also seems to provide a valid quasi-arbitrage edge. Considering Figure 16, this edge might be complementary: the upper panel shows negative returns across the bets with $k < 0.0045$, or so (amid the jumbled zig-zag where the line crosses the r axis). If one were to apply post-hoc filtering, one would certainly remove all such bets. In terms of number of bets, this corresponds to roughly 290. Although no progress is made during the subsequent 3000 or so bets (when the line crosses r and remains positive), this is an average including those first 290 bets, which if removed, would be positive. Thus, that including k as a complementary post-hoc filter would only have removed roughly 290 bets, merely 0.26% of the test sample, further strengthens pd, demonstrating its efficacy in selecting strong bets.

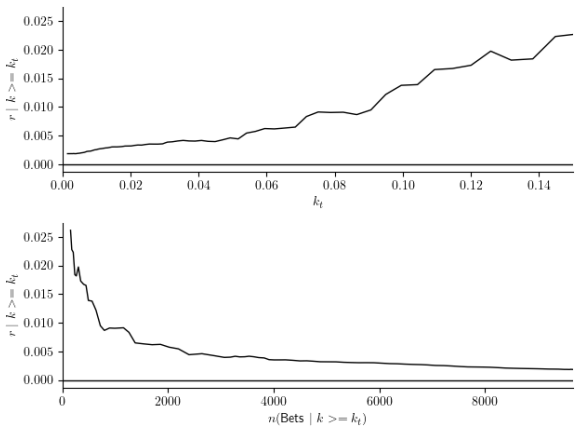

Figure 15: Upper panel: returns per bet where $k \geq k_t$. Lower panel: returns per bet to the n bets where $k \geq k_t$. That is, the n bets with the highest k.

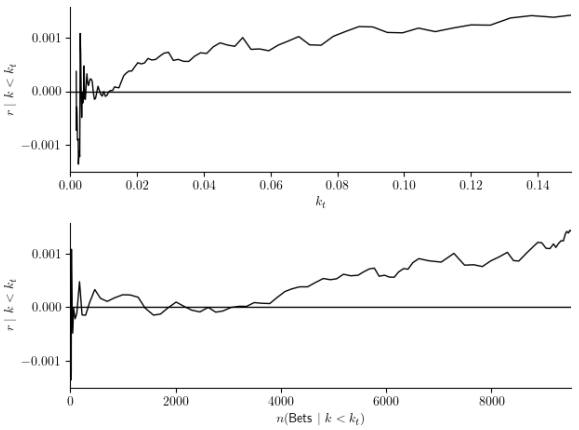

Figure 16: Upper panel: returns per bet where $k < k_t$. Lower panel: returns per bet to the n bets where $k < k_t$. That is, the n bets with the lowest k.

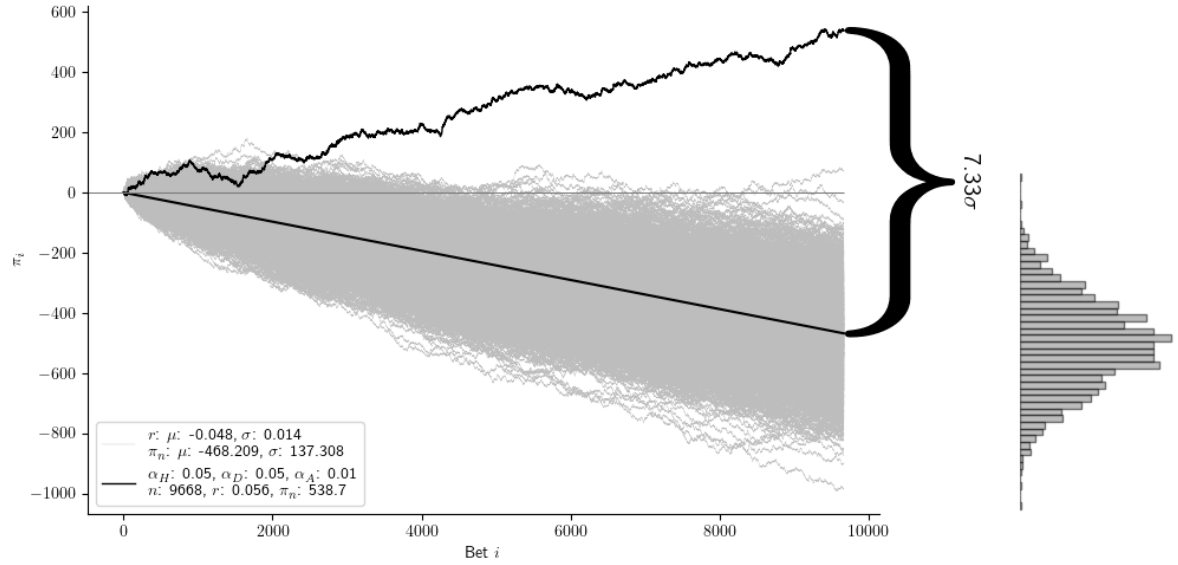

Figure 17: Profit for optimal quasi-arbitrage model with linear staking. With model parameters α_j, r is the per-bet return, with π_n the total profit after n bets. μ and σ are mean and standard-deviation for the random-bet simulation.

By all accounts, these results are wonderful. However, this is not an entirely realistic scenario. Kelly betting is optimal, and guarantees maximal wealth growth in the long-term. However, when capital grows to extremely large values, betting any non-trivial proportion will give enormous stakes. For instance, the selected model here ended with final capital at 50,416,091.95 units. If a unit were, say, \$100, then a 1% Kelly fraction would at this point give a bet of \$50,416,091.95. While not necessarily an abnormal amount by itself, bookmakers invariably set upper limits to stake sizes (the details vary, as discussed below). Therefore, Kelly staking is not realistic in practice, beyond this upper stake limit. The above results show an incredibly lucrative proof of concept in an idealised situation. Naturally, linear staking is an option. Therefore, Figure 17 shows the linear returns, which are very strong, growing to 538.7 units,

more than four times better than the best performance of either of the Kaunitz et al. models, which ended with at most 107.08, over the same dataset. Compared to the familiar 2,000-run simulation, it is 7.33 standard deviations above the mean, making it extremely robust.

Due to staking linearly, we can apply a second test of significance, conceived by Tryfos et al. (1984). Developed originally for American Football spread betting models, where, over n matches, the mean return is calculated as

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = Acc - 1.1(1 - Acc) \tag{17}$$

where Acc is the model's accuracy – the proportion of bets won – and X_i, is the return for bet i, which in our case would be

$$X_i = \boldsymbol{I} o_i - 1 \tag{18}$$

where o_i are the odds for bet i, and $\boldsymbol{I}$ is the indicator function

$$\boldsymbol{I} = \begin{cases} 1 & \text{if the event ends with the bet outcome } j \\ 0 & \text{if the event ends with outcome } k \neq j \end{cases} \tag{19}$$

while for the spread betting model, it would be 1 if the bet wins and -1.1 if the bet loses (stake 1.1 units for the chance to win 1 unit), which allows them to make the jump from the summation of X_i to $\overline{X} = Acc - 1.1(1 - Acc)$. The equivalent result for our model would be that $\overline{X} = Acc \cdot \overline{o} - 1$. However, this is not the case: we have already seen above that this equals 0.3276, when the true mean (linear) per-bet return is 0.056 (Figure 17), calculated using the first equality in the above Tryfos et al. equation. Given $\overline{X}$, we can calculate

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \tag{20}$$

the estimate for the population variance σ^2, corresponding to $\overline{X}$, which estimates the population mean μ. If the number of bets placed n is large, then, by the central limit theorem, we have that

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim N(0, 1) \tag{21}$$

allowing us to test the null hypothesis of unprofitability ($H_0 : \mu = 0$) against the alternative hypotehsis of profitability ($H_1 : \mu \neq 0$), in a two-sided test. We find that $S^2 = 1.74$, giving a Z-score of $\frac{0.0557 - 0}{\sqrt{1.74/9{,}668}} = 4.157$, and a p-value of $2(1 - \phi(4.157)) = 0.000032$.

Even with the more practically realistic and actionable linear staking, the quasi-arbitrage model performs excellently. This concludes the development of our quarb model, a proof of concept: all a prospective bettor requires to bet intelligently and gain positive returns is a set of odds provided by a moderate number of bookmakers. They need no domain-specific knowledge, require no complicated modelling or calculations. Merely the possession of a set of prices is sufficient. Ideally, they would train their model for parameters which specifically suit the sport or market in question, but in the absence of this, setting sufficiently high pd thresholds is a relatively low-risk strategy: we know that exponentially-standardised odds provided by the lowest-margin bookmaker are more informative than the longest odds in high-pd regions, and although what "high-pd" signifies may vary between sports and markets, one can set the threshold as low as risk-tolerance allows, knowing that any threshold above 0 (as seen in Figure 8, in this case) is

sufficient, in expectation. Also with respect to risk-tolerance, the bettor can adjust their value for c (or perhaps they prefer a simpler, raw fractional-Kelly strategy of, say, 0.25 or 0.5 uniformly) to moderate the stake sizes and amounts bet, also controlling risk. All this to say that quasi-arbitrage is a tractable system for making positive effort-adjusted returns.

The following sections focus on developing and integrating a realistic sports-betting asset with conventional (and less conventional) investing methods and assets. There is more to sports-betting than the win market for a subset of European soccer leagues. What is a reasonable estimate and proxy for how many bets a sports-betting investor might make in a set period of time, and how do the returns that they are likely to see compare to what they could achieve investing their capital elsewhere?

8 The Sports-Betting Asset

In this section we aim to approximate what a sports-betting asset might look like, if one were to bet on different sports, markets and leagues, while using a broad variety of methods. There are three components to this analysis. First, what is, in general, a plausible range of returns to sports-betting models? In the previous section we presented a quasi-arbitrage model which bet on 8.71% of matches in 22 top European soccer leagues, on the win market, and generated 0.184% per-bet returns geometrically and 5.6% linearly. But this is only one model in one domain, there are numerous others that can be leveraged for profit. The first step to creating the sports-betting asset is then to survey betting models in academic literature to gain a comprehensive understanding of what ranges of returns are plausible, on a per game basis, across multiple sports.

The second facet of this analysis is gathering data on the frequency with which the sports events are actually contested: what good is knowing a model's per-bet returns, or a range of models' per-bet returns, when we do not know how frequently these can be harvested? Thus, for the various sports covered in the academic literature for which we find several profitable betting models, we gather data on the most recently contested season of that sport (with mentioned exceptions). These figures are aggregated on a monthly basis: when we eventually compare the sports-betting asset to stocks, bonds and the like, it must be done on common ground; therefore, we collect data for how frequently games are played for different sports over different months.

The third part of the analysis is using the monthly-games data, as well as the range of possible sports-betting models, to simulate how these models would perform, on a monthly basis, to generate a distribution of returns which can be integrated with a stock-bond portfolio. This simulation is, in certain respects, general and approximate, while in others, granular and detailed. Instead of plucking exact values found in research papers, we create a distribution of betting models, from which we randomly assign per-bet returns and bet-selection rates to different sports. However, we also honour the real-world distribution of games across the calendar year, so that we incorporate the consequences of June being high-season for baseball and tennis but summer break for soccer and American football, and December being the opposite.

This practice will never be exact. Short of actively constructing betting models for all of these sports, or perhaps applying those that we will shortly outline from existing research papers, and documenting the results on a large variety of different datasets, a complete comprehension of the potential of sports betting will never be attained. This is our most earnest attempt, which is at any rate closer to the truth than an arbitrary, or even informed, guess.

8.1 Surveying Betting Models

For each paper presented, we shortly describe in the appendix (subsection "Surveying Betting Models") our method for extracting the two values of interest: the linear per-bet returns r and the bet-selection rate s. As with this entire endeavour, these deductions are approximate at best; however, there is a systematic approach, giving consistent treatment of the study data. When external data was needed to complement the article numbers, these are described in full detail. We limit ourselves to the past half decade, selecting (with a few exceptions) only papers published in this span. We group the papers by sport, wherein they are not sorted in any particular order.

Figure 18 shows the distribution of models, with per-bet returns versus bet selection. Ignoring the accuracy-based models with $s = 100\%$, there is an evident inverse relationship, where very high returns are possible only at very low bet-selections. This is seen, to a lesser extent, in Figures 5 and 8 above, where, when the α parameters increase, r increases while n decreases, as well as in Figure 15, with respect to k_t instead of α.

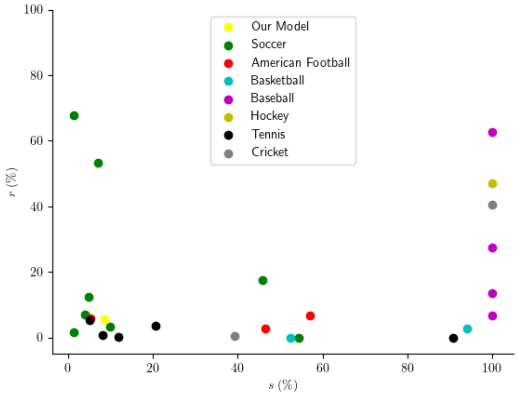

Figure 18: Distribution of sports-betting models in a sample of academic literature, with per-bet linear returns r against bet-selection rate s, both in percent. The points with $s = 100\%$ are accuracy-based models, where their returns are estimated using small samples of odds.

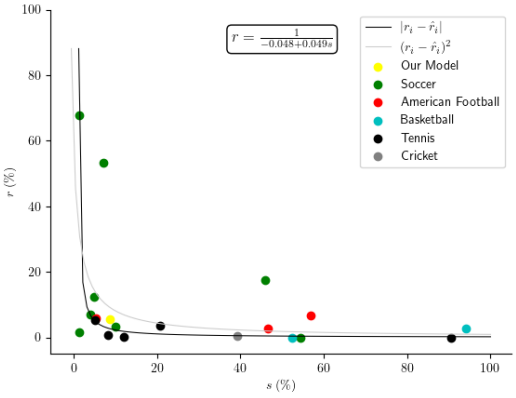

Figure 19: When removing the accuracy-based models, a strong fit between s and r becomes evident, which is captured by $r = \frac{1}{-0.048+0.049s}$, with r and s both entered in terms of percent. The model was estimated by minimising the sum of absolute differences (rather than squared differences).

To capture this inverse relation, we fit the regression

$$r = \frac{1}{\alpha + \beta s} \tag{22}$$

When expressing r and s as percent (that is, enter the data-point (5.6%,8.71%) as (5.6, 8.71)), we find that $r = \frac{1}{-0.048+0.49s}$ is the best fit, as seen in Figure 19. We perform this regression by minimising the sum of absolute differences between the r values and their values $\hat{r}$

$$\arg\min_{\alpha,\beta} \sum_{i=1}^{n} |r_i - \hat{r}_i| \tag{23}$$

where $\hat{r}_i = \frac{1}{\alpha+\beta s_i}$. This is done because, rather than using the more conventional regression technique of minimising the sum of squared differences, as can be seen in Figure 19, the fit is

better to the main cluster of points (low-s,low-r region), the line does not leave the accepted domain for s over the accepted range of r, and primarily because it is a more pessimistic estimation, having lower r to s across almost the entire domain.

When we simulate returns for our sports-betting asset, we will draw betting models from this line and assign them randomly to these five sports, applying any (r, s) point on this line to odds and games from any of the sports. This two-dimensional distribution of sports-models is representative of the interplay between s and r for real-world betting models, while being more general than plucking the exact (r, s) pairs found in the papers. It is also reasonable to assume that the specific returns found for the different sports are likely available to the other sports as well: if a betting model can achieve (r, s) for one sport, we can imagine that these attributes can be achieved for the other sports as well.

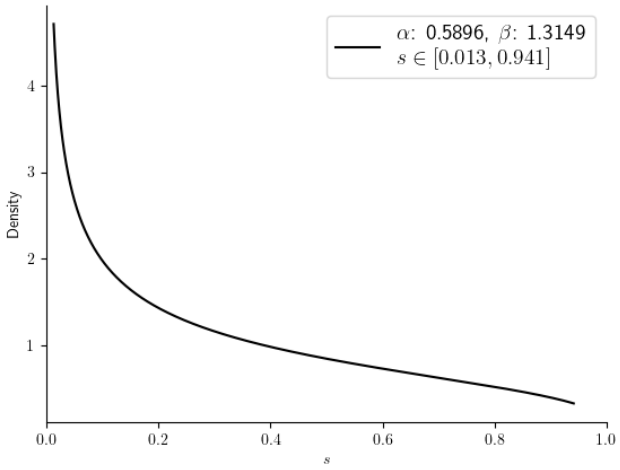

Figure 20: Beta distribution for bet-selection rate s, fit and then truncated to $s \in [0.013, 0.941]$.

Considering Figure 19, there are no distinct clusters of sports, no characteristic regions where betting models for a given sport tend to be located. We confirm this by considering mean silhouette scores (Rousseeuw, 1987) for each cluster (our model added to the soccer group). For a single data point, the silhouette score compares the distance to other data points within its cluster to points outside its cluster. The mean silhouette score for a cluster is then a measure of how well assigned the clusters are, with a score of 1 being optimal and -1 the opposite. Cricket has a score of 1, which is expected given that there is only one cricket model amongst the $s \neq 100\%$ group. The American football, soccer and tennis models are all unrelated, their silhouette scores being negative, with -0.302, -0.097 and -0.045, respectively. Even for basketball, which has two points in the upper s region and is therefore the most cohesive cluster of the four multi-model sports, only has a slightly positive silhouette score of 0.126, suggesting that its members are by no means identical. Additionally, given that both basketball models are in the high bet-selection area, if we were to assign uncharacteristic models to it, these would be from the low-s region, giving fewer bets and thus, for the vast majority of the domain where the line is essentially flat, lower returns. This pessimism principle also applies to cricket, which has only one model, in the high-s domain as well.

The sampling of models will be performed by drawing an s value and then calculating the corresponding r, using the fitted line. We fit a beta distribution to the vector of s values, as shown in Figure 20, finding the parameterisation $\alpha = 0.5896, \beta = 1.3149$. Wanting to stay within the observed domain of models found in the survey of existing literature, the domain for s is up to 0.941. Similarly, we do not want to exceed the maximum per-bet return found in the sample, namely 67.9%, which gives a lower bound to s of 0.01294, which we round to 0.013. Based on this Beta distribution of s and the inverse relation between r and s, we can calculate

the expected value of r as

$$E(r) = \frac{\int_{0.013}^{0.941} \frac{Beta(s;0.5896,1.3149)}{-0.048+0.049\cdot 100s} ds}{100 \int_{0.013}^{0.941} Beta(s; 0.5896, 1.3149) ds} = 0.0295 \tag{24}$$

where $Beta$ is the probability density function of the beta distribution, and the denominator integral is an adjustment due to truncating the s domain (typically beta-distributions are integrated from 0 to 1). $r = 2.95\%$ is lower than the arithmetic mean of 9.8% among the $s \neq 100\%$ models, which, again, is due to the chosen fit of r as a function of s being pessimistic, to not exaggerate the possible returns to a sports-betting asset.

Finally, baseball has four models with $s = 100\%$, with r equal to 6.74%, 13.62%, 27.32% and 62.7%, respectively. For baseball models in our simulation, keeping $s = 100\%$, we draw $r \sim U(0.0674, 0.1362)$, where U is the uniform distribution, and its bounds are r for the two lowest-returning models. We restrict ourselves to this lower-r range to stay consistent to the spirit of caution and under-estimation, not wanting to exaggerate the profitability of a sports-betting asset.

8.2 Games and Odds by Sport

We now have a distribution of sports models – a reasonable manner in which we can generate (r, s) pairs – to simulate real-world betting across multiple sports. These models are (other than for baseball) sport-agnostic, with cricket having the same array of models as does soccer, and so on. What does differ is their distribution of odds and the distribution, across a calendar year, of competitive games on which to bet. Baseball and tennis are summer sports, while soccer and American football are predominantly played during the winter. In this section we aim to capture these nuances, so that a representative distribution of sports-betting returns can be generated on a per-month basis, meaning that the distribution of results will reflect the betting high-season in April and the low-season in December.

For each sport we gather data, over the most recently-contested season, from all "major" leagues for that sport, major in quotation marks, as this definition is somewhat arbitrary. For some sports, tennis for instance, there is a clear pro tour, and it is evident how many competitions are "major", how many games are played in each, and when these games are played. However, for cricket, international competition is extremely important, while the players also compete in franchise leagues across the world. Deciding whether a given short-format tournament should be included in the data or not is hardly self-evident. A distinction has in all cases been made, with the results in the appendix below (subsection "Games and Odds by Sport").

We also collect odds for each sport. This is to understand what the general distribution of odds is for each particular sport. For instance, baseball and tennis are both games with two outcomes, but in the datasets that we collect for each of the sports, respectively, the largest underdog in a baseball game was priced at $o = 4.3$, while for tennis this was 35.0. Clearly then, assuming that odds are identically distributed across sports is not realistic. Therefore, we collect odds for each sport, so that we can match the detailed calendar of games for each sport that we gather with an equally detailed set of representative prices. As with the betting-model survey described above, the exact information is in the appendix. Note that ice hockey was omitted from this analysis, due to only having one betting model in the survey.

8.3 Creating the Sports-Betting Asset

With the above information, we can now simulate how a real-world sports-betting asset's returns might be distributed on a monthly basis. We have a range of models that are general and applicable to the sports in question. For each sport we have a representative set of odds, characteristic especially in the range of possible odds, as well as a calendar of games per month. We can simulate how many bets an investor would make for each sport in a certain month, given the betting model that they have for that sport, as well as the returns that they experience as a result. This investor bets on each sport, meaning that we can aggregate returns across sports to obtain a sports-betting asset monthly return, not a sport-specific monthly return.

The simulation is performed as follows. We simulate 1,000 years. For each year, for each sport, draw a bet selection rate s from the beta distribution described above, and then calculate the corresponding per-bet return r. This model belongs to the sport for the entire year. For each month in the year, calculate the number of games that will be bet on, which is s multiplied by the total number of games (across all competitions) for that sport and month, rounded down. Randomly select this many games from the odds data for that sport. For each game, randomly select one of the odds, o. For each game, we can calculate the outcome probability, given o and the expected return r, using $p = (1 + r)/o$ (since $po - 1 = r$). Draw a random number $q \sim U(0, 1)$. If $q \leq p$, this game's returns are $o - 1$, otherwise they are -1. Do this for each of the games in the month for the sport and sum the return. Do this for each sport in the month. Sum the total returns across the sports and the total number of bets across the sports, and divide the latter from the former, giving the monthly return r_m. Do this for each of the months in the year, then select new models for each of the models for the next year and repeat 1,000 times. We thus have one r_m value for each month. With twelve months and 1,000 years, this gives 12,000 values for r_m. Their distribution in one such random simulation is shown in Figure 21.

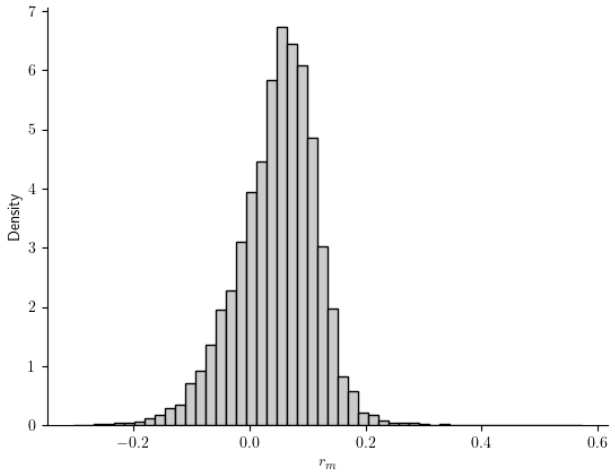

Figure 21: Distribution of monthly returns r_m to the sports-betting asset. Mean is 0.047, volatility is 0.0689.

As expected, given that the sports-betting models which underlie the simulation are all profit-making, the returns are positive on average, with a mean of 0.047 and median of 0.0539. Indeed, only in 22.45% of months are losses made. The worst month saw losses of 30.5%, while the best month returned 57.43%. The volatility of the asset is a modest 0.0689, giving a Sharpe ratio (with $r_f = 0.0196$, as explained below) of 0.398. These are flattering numbers, making the sports-betting asset rather attractive. In the next section we integrate it with a conventional stock-bond portfolio to identify just how attractive.

9 Portfolio Integration

In a mean-variance portfolio (Markowitz, 1952), weights are allocated to different assets in consideration of their mean return vector $\boldsymbol{\mu}$, their volatility vector $\boldsymbol{\sigma}$ and their correlation matrix $\boldsymbol{C}$. In this section we will construct such a portfolio by combining indices for different asset classes, including the sports-betting asset developed in the previous section. The purpose of this exercise is to evaluate how the sports-betting asset compares to these established forms of investment. We proceed as follows. First, we present the different indices and asset classes under consideration, including their monthly time series. Then, we consider their correlation structure with the sports-betting asset, arguing, briefly, that they are uncorrelated. Finally, we create an optimal portfolio, or rather, portfolios: for each asset, first removing it from the asset pool, and subsequently observing the change in Sharpe ratio when it is reentered.

9.1 Assets

For each asset class a major index was chosen. In the case of the more obscure, alternative investments, the choice of index was forced, selection due to availability. For the more conventional assets, an effort was made to select the most renowned or relevant index: the S&P500 for stocks, and so on. For each, we collect monthly data over ten years from March 2014 through to March 2024, or if not available over the entire period, for as much of it as possible. For all assets, we take the first available index value in each month. See the appendix for the exact details (subsection "Assets"). The asset classes considered are stocks, bonds, real estate, commodities, cryptocurrency, art, wine, classic cars, rare coins, luxury watches and collectibles.

Figure 22 plots the various index trajectories over the period of time where data was available, since 2014. Art and stocks perform particularly well, with cryptocurrency substantially the most volatile. Interestingly, many of the asset classes perform poorly, the bond index particularly surprising. Table 6 shows the correlation matrix between the eleven indices, measured over the longest common period between August 2021 and February 2024. Also shown are the monthly mean and volatility estimates, μ_m and σ_m, for each index, across this same span. During this period, the mean risk-free return r_f equalled 3.0%.

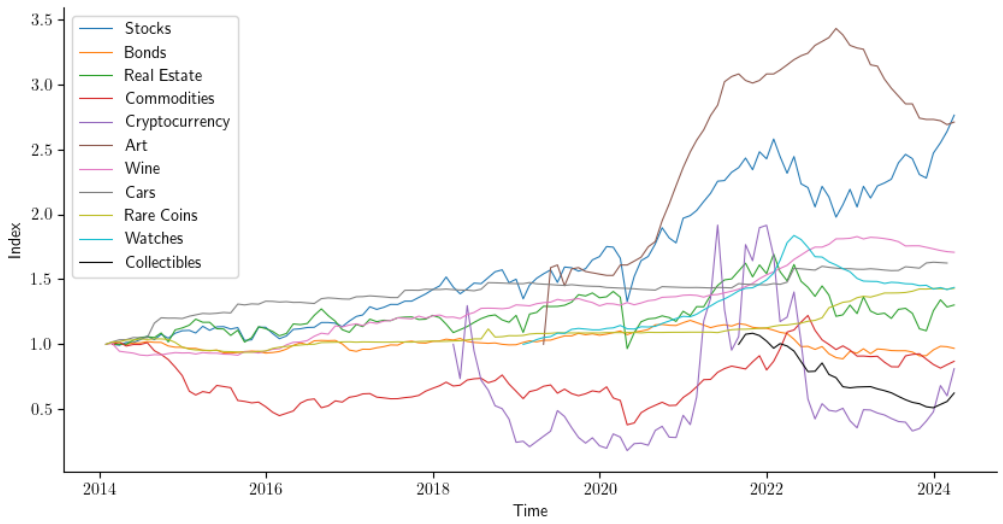

Figure 22: For all asset classes, their index trajectories over the period of available data.

	Stocks	Bonds	Real Estate	Commodities	Cryptocurrency	Art	Wine	Cars	Rare Coins	Watches	Collectibles
Stocks	1										
Bonds	0.657	1									
Real Estate	0.8957	0.6035	1								
Commodities	-0.0376	-0.1787	-0.0705	1							
Cryptocurrency	0.4795	0.2772	0.5089	-0.2144	1						
Art	-0.1924	-0.1315	-0.03	-0.0069	-0.124	1					
Wine	-0.4317	-0.5281	-0.302	0.1368	-0.2419	0.592	1				
Cars	0.2741	-0.1011	0.3443	0.0942	0.2287	0.1851	0.1236	1			
Rare Coins	-0.1745	-0.315	-0.1528	-0.1547	-0.17	0.2282	0.1673	-0.1224	1		
Watches	0.0296	-0.0574	0.0552	0.5184	0.1705	0.1909	0.272	0.3051	-0.2819	1	
Collectibles	0.1742	0.3402	0.1449	-0.1447	0.491	-0.0991	-0.1599	-0.0379	-0.389	0.1989	1
μ_m	0.005	-0.0052	-0.0043	0.0026	0.0053	-0.0044	0.0063	0.0035	0.0083	0.0006	-0.0181
σ_m	0.051	0.0278	0.0657	0.0592	0.225	0.0161	0.0122	0.0146	0.012	0.0262	0.0447

Table 6: Correlation matrix for the eleven indices, with monthly mean μ_m and volatility σ_m.

9.2 Sports-Betting Asset Correlation

For the sports-betting asset, we have μ_m and σ_m. However, we do not know how sports-betting returns are correlated, if at all, with the selected asset indices. Edmans et al. (2007) find that national stock markets decline following the nation's sporting team's elimination from a major international tournament, robust to soccer, cricket, rugby and basketball. Baddour (2010) shows that, during the course of the Fifa world cup (soccer), the S&P 500 has historically returned 2% lower than average. Levy (2015) shows that championship wins for New York sports teams generates abnormally large returns to the major US stock markets. However, these are all major events, which occur exceedingly seldom, not at all comparable to, for instance, the 2,430-game MLB regular season. The fundamental, causal force supporting the results of these papers, is a uniform, collective, national (or even global) sentiment, whereby an entire population is systematically affected by some emotional event: when a nation is eliminated from world-championship competition, the entire population is negatively affected, hurting national investment sentiments; when the Fifa world cup is held, the entire world pauses, with global stock returns similarly delayed; when a New York (note, major US financial centre) team accomplishes a once-every-few-years feat, the US stock markets return highly. These are all seminal events, and thus impact financial returns. However, the sporting events under consideration here are a normal feature of the everyday routine in all nations, never being large enough or significant enough to capture an entire population's attention and spirit. But even if, say, one game per month might do so, the mean of the beta distribution in Figure 20 is $\frac{0.5896}{0.5896+1.3149} = 0.310$, meaning there is a 31% chance (on average) of it being selected for betting (if it is non-baseball); and if it is selected, it is one of 862 bets placed in a given month (mean number of bets underlying r_m in Figure 21, rounded), thus having a mean return of $0.047/862 = 0.00545\%$, where 0.047 is the mean monthly return in Figure 21. So even if this once-a-month event is correlated with the stock market, its individual return is so inconsequential to the overall sports-betting asset, that the aggregate correlation is most certainly trivial.

9.3 Portfolio Construction and Analysis

We construct a mean-variance portfolio, given a vector of mean returns $\boldsymbol{\mu}$, a vector of volatilities $\boldsymbol{\sigma}$ and a correlation matrix $\boldsymbol{C}$, where we find a vector of portfolio weights $\boldsymbol{w}$ which determines what proportion of our investment capital is allocated to each asset. The portfolio's expected return is

$$E(R_p) = \boldsymbol{w}^T \boldsymbol{\mu} \tag{25}$$

From the volatility vector and correlation matrix, we calculate the covariance matrix

$$\boldsymbol{\Sigma} = diag(\boldsymbol{\sigma})\ \boldsymbol{C}\ diag(\boldsymbol{\sigma}) \tag{26}$$

where $diag(\boldsymbol{\sigma})$ denotes a diagonal matrix with $\boldsymbol{\sigma}$ on the diagonal. From this, we can calculate the portfolio volatility

$$\sigma_p = \sqrt{\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}} \tag{27}$$

The portfolio Sharpe ratio (Sharpe, 1966) is thus

$$S_p = \frac{E(R_p) - r_f}{\sigma_p} \tag{28}$$

with which we optimise for $\boldsymbol{w}$:

$$\begin{aligned} &\arg\max_{\boldsymbol{w}} \quad S_p \\ &\text{Subject to} \\ &\sum_{i=1}^{n} w_i = 1 \\ &w_i \in [w_{li}, \infty^+] \ \forall\ i,\ w_{li} \in \{\infty^-, 0\} \end{aligned} \tag{29}$$

where w_{li} , as specified, is the lower bound for w_i, either ∞^- if short selling is allowed for the asset class, or 0, if not. In our case, short selling is allowed for the stock, bond, real estate, commodities and cryptocurrency indices. For all assets, we calculate what we call the Sharpe-ratio contribution $\frac{S_p}{S_i}$, where S_i is the Sharpe ratio of the optimal portfolio in the absence of asset i in the asset pool. The larger is this ratio, the more does asset i contribute to the overall portfolio's Sharpe ratio, S_p. Table 7 shows, for each asset, this ratio and the asset's optimal portfolio weight, w_i.

	Stocks	Bonds	Real Estate	Commodities	Cryptocurrency	Art	Wine	Cars	Rare Coins	Watches	Collectibles	Sports Betting
w_i	9,852.524	-17,834.525	-3,881.325	-3,481.172	441.88	2.274×10^{-13}	13.293	1.137×10^{-13}	2,201.846	1.137×10^{-13}	1.137×10^{-13}	12,688.479
$\frac{S_p}{S_i}$	1.211	1.304	1.178	1.011	1.023	1.016	1.009	1.007	1.021	1.005	1.048	4.066

Table 7: Optimal portfolio weights and Sharpe-ratio contribution, for all assets.

The sports-betting asset has the highest Sharpe-ratio contribution, by a wide margin. Other than bonds, it has the largest weight in terms of magnitude, while having the largest positive weight. Of the conventional assets, the portfolio is long in only stocks and cryptocurrency (if the latter is indeed considered conventional). Bonds, real estate and commodities are sold short. In fact, bonds have the second highest Sharpe-ratio contribution due to how short the portfolio is in it. Of the alternative investments, art, cars, watches and collectibles are given essentially zero weight. The inclusion of collectibles is therefore rather wasteful, as it currently limits the period over which the mean, volatility and correlation parameters are measured, providing data only since 2021. Therefore, we remove it from the asset pool, and recalculate these arrays, as well as reperforming this analysis. The results are shown in Table 8.

With the new parameter estimates, including the risk-free rate now at $r_f = 1.96\%$, the sports-betting asset retains the largest Sharpe-ratio contribution, but not quite as overwhelmingly. Now, too, the stock asset has a larger positive weight (bonds still hold the largest absolute weight). Over this period, only the classic cars index has no weight assigned. Despite having

zero weight, its Sharpe-ratio contribution is slightly above one, suggesting that its mere presence of availability improves the portfolio, despite it not receiving any investment capital. Presumably this is an optimisation inefficiency, with its correlations with the other assets causing convergence to an alternative local maximum via a slightly different path.

	Stocks	Bonds	Real Estate	Commodities	Cryptocurrency	Art	Wine	Cars	Rare Coins	Watches	Sports Betting
Stocks	1										
Bonds	0.5346	1									
Real Estate	0.8564	0.5279	1								
Commodities	0.5075	0.0332	0.405	1							
Cryptocurrency	0.4173	0.2277	0.4343	0.1228	1						
Art	0.0066	0.0249	0.0345	0.0467	0.0653	1					
Wine	-0.3525	-0.379	-0.2447	-0.0202	-0.1987	0.0835	1				
Cars	0.1707	-0.1026	0.2278	0.0341	0.1075	-0.0126	0.1633	1			
Rare Coins	-0.1701	-0.3347	-0.1459	-0.1213	-0.1988	-0.0745	0.22	-0.0444	1		
Watches	0.1017	0.007	0.1955	0.3792	0.228	0.1345	0.2136	0.2612	-0.304	1	
Sports Betting	0	0	0	0	0	0	0	0	0	0	1
μ_i	0.0109	-0.0006	0.0023	0.0069	0.048	0.0198	0.0046	0.0019	0.0048	0.0054	0.047
σ_i	0.0564	0.0226	0.0671	0.0763	0.2858	0.0821	0.0111	0.0115	0.0096	0.0213	0.0689
w_i	4,457.464	-7,099.519	-1,749.027	-516.516	101.013	1,344.76	143.962	0	201.387	88.394	3,029.083
$\frac{S_p}{S_i}$	1.124	1.423	1.049	1.031	1.028	1.068	1.006	1.005	1.006	1.016	2.681

Table 8: After removing the collectibles index, the correlation matrix for the then remaining indices plus the sports-betting asset, as well as monthly mean and volatility for asset i, μ_i and σ_i. New optimal portfolio weights and Sharpe-ratio contribution are calculated and displayed. Across this new time period, $r_f = 1.96\%$.

As was the case for the previous portfolio, which included the collectibles index, with parameters measured over a truncated time period, the overall portfolio has enormously inflated values, with an expected return of $E(R_p) = 22{,}121.05\%$ and a volatility of $\sigma_p = 28{,}398.13\%$. This is unsurprising given that, although the weights do sum to 1, their absolute values sum to 18,731.12. An investor with this exact portfolio would be making transactions worth almost 19,000 times their own capital. As is typically done in the literature, we assume that portfolio returns are normally distributed (which is realistically not the case, but the point being made is valid nonetheless), which gives a monthly probability of bankruptcy of $\phi\left(\frac{-1-221.2105}{283.9813}\right) = \phi(-0.787) = 0.2156$! Certainly not something that the everyday investor has a risk-appetite for. Therefore, we limit the maximum absolute weight to 1: from equation (29), we now get $w_i \in [w_{li}, 1] \ \forall \ i$, where $w_{li} \in \{-1, 0\}$, depending on short-selling allowance. With these conditions, we obtain the results in Table 9.

	Stocks	Bonds	Real Estate	Commodities	Cryptocurrency	Art	Wine	Cars	Rare Coins	Watches	Sports Betting
w_i	1	-1	-0.73092	-0.1208	0.0811	0.3188	0	0	0.2593	0.1926	1
$\frac{S_p}{S_i}$	1.135	1.147	1.096	1.004	1.031	1.053	1	1	1.001	1.001	2.621

Table 9: Optimal portfolio weights and Sharpe-ratio contribution for all assets, sans Collectibles. Maximum absolute weight is 1.

Sports betting remains the most valuable asset in terms of its Sharpe-ratio contribution. Now both wine and cars are ignored, and their contributions are correctly 1: most likely the inconsistency above with the cars contribution being greater than one arose from the considerable weight magnitudes. This portfolio has a more modest returns profile, with $E(R_p) = 6.85\%$ and $\sigma_p = 8.29\%$, but is considerably more stable (giving a probability of bankruptcy in a single month of 2.55×10^{-38}), with the sum of absolute weights now 4.70. Additionally, if we do not allow short selling at all, only cryptocurrency and the sports-betting asset are given weight, with 5.7% and 94.7%, respectively.

We consider one final scenario. Above, we have calculated value at risk (Jorion, 2007; Kuester et al., 2006) as

$$\phi\left(\frac{\beta - E(R_p)}{\sigma_p}\right) = \alpha \tag{30}$$

to approximate the probability α of a portfolio with mean $E(R_p)$ and volatility σ_p experiencing a monthly return of β or worse. Above, we have optimised portfolios both without constraints on the absolute maximum weight for an asset, as well as with a maximum absolute weight of 1. Here, we instead optimise for a maximum absolute weight which renders the subsequent portfolio consistent with a desirable level of value at risk, per equation (30). We choose our confidence level α and value at risk β, and find the value of w_b, giving $w_i \in [w_{li}, w_b] \ \forall \ i$, where $w_{li} \in \{-w_b, 0\}$, depending on short-selling allowance, such that the equality in equation (30) holds, with $E(R_p)$ and σ_p calculated through equations (25) and (27), with the weights optimised in (29).

This process works well for α-β pairs where the resultant w_b is a reasonably large value. However, when it is extremely small, say, $w_b = 0.01$, then it is impossible to satisfy the condition that $\sum_{i=1}^{n} w_i = 1$, since, for our eleven assets, $\sum_{i=1}^{11} w_i$ is at most $\sum_{i=1}^{11} w_b = 11 \cdot 0.01 < 1$. Even if w_b is only somewhat small, at say $w_b = 0.1$, while the weight-summation condition can be satisfied, the weights have much less freedom to be individually optimised: $0.1 \cdot 11 = 1.1$, meaning that there is essentially $1.1 - 1 = 0.1$ worth of weight-space for the separate w_i to be optimised across, which will give forced solutions rather than optimised solutions. Indeed, typically for these scenarios where w_b is in this range where the weight-summation condition can be satisfied but not with many degrees of freedom, one or two weights will be some negative number and the rest will equal w_b. In these low-w_b cases, we therefore attempt to create a slightly different portfolio-optimisation mechanism. We allow $\sum_{i=1}^{n} w_i \in [w_b/2, 1]$ and invest the remainder $1 - \sum_{i=1}^{n} w_i$ in the risk-free asset. We then get $E(R_p) = \boldsymbol{w^T \mu} + r_f(1 - \sum_{i=1}^{n} w_i)$ and $S_p = \frac{\boldsymbol{w^T \mu} - r_f \sum_{i=1}^{n} w_i}{\sigma_p}$, with σ_p as in equation (27), being unaffected by the zero-volatility risk-free asset. We select $w_b = 0.444$ as our threshold for switching from one portfolio-optimisation method to the other, as this is the point where, using the conventional optimisation, as w_b increases above 0.444, the number of unique absolute w_i plateaus. In the example given where one or two w_i are negative and the remaining equal w_b, the number of unique absolute w_i is small, meaning the optimisation has been forced. Above $w_b = 0.444$ weights are unique to the utmost extent.

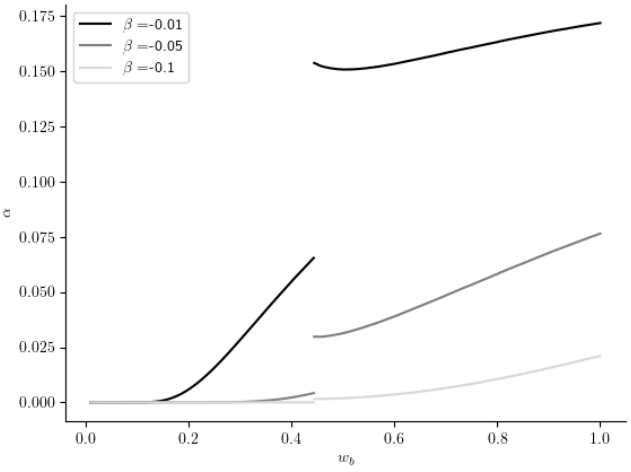

Figure 23: The probability α of a monthly return below β as a function of w_b. Our naive alternative portfolio-optimisation is inadequate in the domain $w_b < 0.444$.

The hope is that this abrupt threshold does not cause a discontinuity in α as a function of w_b, however, as seen in Figure 23, there is a large jump in the estimated value at risk probability. Therefore, we discard our naive model and restrict ourselves to α-β pairs that require $w_b > 0.5$. We increase the threshold to 0.5 since the $\beta = -0.01$ line is decreasing on the domain $w_b \in$

$[0.444, 0.5]$, which should not be the case: a lower w_b gives a less extended portfolio, meaning loss probabilities should increase monotonically in w_b. Of course, a lower absolute β would give this minimum at a larger w_b, so we therefore restrict our analysis to $\beta \leq -0.01$, choosing the combinations of $\alpha, \beta \in \{0.05, 0.1\}$ (positive for α, negative for β), as shown in Table 10. It should be noted that the value-at-risk metric is only relevant for the main portfolio, not for the sub-portfolios generated when iteratively removing each asset i; these sub-portfolios, rather, are calculated with the w_b found for the main portfolio.

α	β	w_b		Stocks	Bonds	Real Estate	Commodities	Cryptocurrency	Art	Wine	Cars	Rare Coins	Watches	Sports Betting
0.05	-0.05	0.7159	w_i	0.7159	-0.7159	-0.5614	-0.0718	0.0717	0.2702	0	0	0.4196	0.1558	0.7159
			$\frac{S_p}{S_i}$	1.155	1.115	1.116	1.003	1.045	1.07	1	1	1.005	1.002	2.844
	-0.1	1.481	w_i	1.481	-1.481	-1.037	-0.1965	0.1024	0.4207	0	0	0.0017	0.2279	1.481
			$\frac{S_p}{S_i}$	1.119	1.209	1.082	1.006	1.026	1.044	1	1	1	1.001	2.409
0.1	-0.05	1.3208	w_i	1.321	-1.321	-0.9342	-0.172	0.095	0.387	0	0	0.0857	0.2177	1.321
			$\frac{S_p}{S_i}$	1.123	1.19	1.085	1.005	1.026	1.046	1	1	1	1.001	2.475
	-0.1	2.938	w_i	2.938	-2.938	-1.955	-0.4283	0.1645	0.6558	0	0	0	0	2.563
			$\frac{S_p}{S_i}$	1.101	1.295	1.071	1.011	1.02	1.035	1	1	1	1	1.985

Table 10: Portfolio optimisation under value at risk, maximising the Sharpe ratio, subject to the absolute weight bound w_b giving a probability of α of experiencing a monthly return of β or worse. Weights for all asset classes w_i and their Sharpe-ratio contribution $\frac{S_p}{S_i}$.

For each of the four value-at-risk analyses, the sports-betting asset retains the highest Sharpe-ratio contribution. Interestingly, in the scenario representing a 10% chance of losing at least 10% in a given month, the sports-betting asset does not receive the full w_b allocation, while stocks and bonds receive w_b and $-w_b$ in each case, respectively. This suggests that, while the sports-betting asset already has the largest Sharpe-ratio contribution, it is also more important than these other two assets when controlling for downside risk: to create a portfolio with a relatively high probability of a relatively large loss, we want less of the sports-betting asset, not more of it. The rare coins asset, strangely, has a positive weight in the two five and ten percent scenarios, yet has a Sharpe-ratio contribution of one, which is impossible. This, however, is a rounding trick, as its contribution is larger than, say, that of cars (for example, 1.0001875 compared to 1.0000013 for $\alpha = 0.1$, $\beta = -0.05$).

All things considered, the sports-betting asset is the most attractive of the group, consistently contributing the most to the portfolio, in various different investment situations. Having demonstrated first the relative ease with which a domain-agnostic, highly lucrative betting model can be constructed; we then surveyed existing betting literature, searching for models that have been applied and proven in practice, from which we created a distribution of betting models; to match these models, we gathered data for odds and games for different sports, with which we simulated returns to the betting models with maximal realism; finally, we argued that the betting asset is uncorrelated with the broader financial market, and have now seen that, when integrated with conventional financial assets, it prevails as the most valuable. With this information, then, sports betting seems to be an underused source of financial returns, one that should be acknowledged and embraced immediately, by individual investors and wealth managers alike. It is not quite so simple, however, as there exists a set of market frictions which prevent investors from fully leveraging these seemingly tractable markets to their financial benefit. The following section discusses these frictions, and evaluates the extent to which they can be overcome.

10 Market Frictions

The two predominant sports-betting market frictions have already been mentioned, albeit briefly, above. Firstly, it is well-known that sports-betting arbitrage (full arbitrage, not merely quasi) is readily available, with, for instance, betmonitor.com having a list of present-time opportunities to bet simultaneously on off-setting outcomes (say, home win and away win in baseball) with different bookmakers. However, it is also common knowledge that if these opportunities are abused, the bookmakers will restrict such naive bettors – decrease their maximum stake size – or even ban them outright: if you know that a bookmaker is exposed to full arbitrage, you know for certain that they know this as well. Therefore, a degree of subtlety is required.

Secondly, we estimated our quasi-arbitrage model above by optimising for geometric staking with the Baker-McHale Kelly criterion, where stakes are a fraction of the current capital, rather than linear staking, where a certain number of units are staked on each bet, uniformly. The reason why we then instead proceeded with a linear staking system, extracting per-bet linear returns from the surveyed sports-betting models, is that bookmakers set an upper limit to the bets that they accept from bettors, meaning that only at a relatively low capital level is geometric staking an option. This is not necessarily an issue in general, as the fundamental principle behind the sports-betting asset is that an enormous number of perhaps only marginally profitable bets are made, with a slow but steady accumulation of profits as returns converge to the positive expectation asymptotically. However, if a hedge fund or large investment bank with significant assets under management were to consider sports betting as a potential complementary investment, this decreasing return to scale would likely prevent them from making the initial investment of time needed to generate the appropriate models and staking systems, as their return as a percentage of total capital would be minuscule. And this requirement of placing an extraordinary number of bets is a third financial friction, compared to conventional financial assets where excessive trading incurs large transaction costs. Trade detection and automation is possible, but it raises the technical barriers to entry above merely having some rather elementary mathematical understanding of sports-betting processes.

The art of sports betting, indeed, lies not in the mathematical estimation of profitable models, but rather in giving due respect to these frictional considerations: firstly, ensuring that one does not arouse the bookmaker's attention by abusing excessively profitable opportunities, while, secondly, simultaneously making sufficiently many bets to guarantee a substantial enough return to validate this strenuous process, because, thirdly, bettors cannot make arbitrarily large bets, due to the stake limits bookmakers apply. We here estimate these factors in an attempt to evaluate (given that a sports-betting asset is indeed financially viable in a vacuum, as seen above) the scale at which profits can be made: whether high assets-under-management firms should invest in sports betting, or whether it should instead be the realm of moderately-wealthy, reasonably-technical individuals.

10.1 Suspicious Betting Behaviour

The principal friction for prospective betting investment is bookmakers restricting or banning winning customers. Kaunitz et al. relate how they had stake sizes restricted by at least four bookmakers (their Figure 4), ostensibly due to their successful betting. However, given their returns profile, we suspect that they were actually restricted due to their conspicuous betting behaviour. From their Table 1, they made a total of 265 actual bets with real funds, winning 125 of these, generating a profit of \$957.5 from 50-dollar bets (note that their return value of 8.5% is incorrect, rather being $\frac{957.5}{50 \cdot 265} = 7.23\%$, a calculation consistent with the other rows in

the table which correspond to backtested betting). Their sample size of 265 is small, especially considering how these bets were spread out across at least eight bookmakers (Figure 4 and Supplementary Figure 1, both theirs). Sports-betting is a high-variance process in the short-to-medium term: consider our Figures 6 and 7, where in the short-term random-betting is profitable in a large number of simulations; and our Figure 17, where even after almost 10,000 random bets, a few runs remain positive. A bookmaker is not going to ban all punters who are profitable over the short-term, needing evidence that these profits are significant: otherwise, they would be banning what are fundamentally losing customers, who are simply experiencing quite likely streaks of luck.

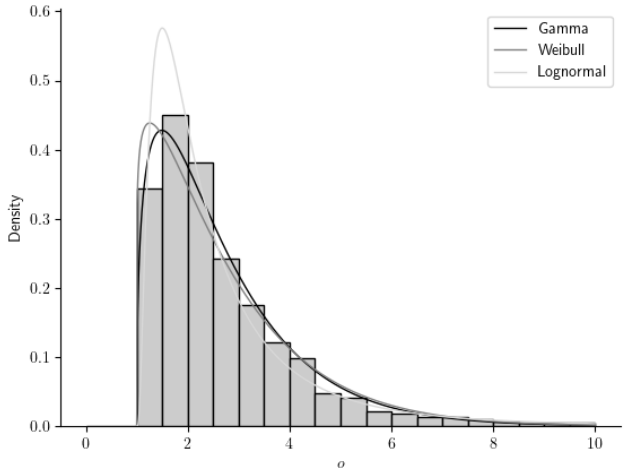

Figure 24: Gamma, Weibull and Lognormal distributions fit to the odds at which bets were placed during the true run of the quasi-arbitrage model (equation 15). Distributions are shifted, starting at 1.

Therefore, we test whether the Kaunitz et al. profits are significant or not, using the Tryfos test outlined above (equations 17 through 21). However, although we have $\overline{X} = 0.0723$, to calculate the estimate of the population variance S^2 (equation 20), we need the set of odds at which their bets were placed, which Kaunitz et al. do not disclose. We therefore assume that these odds are drawn from the same distribution as were the odds at which our optimal quasi-arbitrage model bet. This distribution is shown in Figure 24, fit with Gamma, Weibull and Lognormal models, where these distributions are shifted to start at 1, rather than 0, accounting for odds in decimal form being at least 1. Although all three models reject the Kolmogorov-Smirnov test null hypothesis (Massey Jr, 1951), suggesting that the data is not drawn from any of these distributions, the visual fit is adequate for our estimation, especially considering how there is no guarantee that the Kaunitz et al. odds are identically distributed to ours. We therefore select the Lognormal distribution (which had the least poor fit), and use this to generate a sample of odds to approximate the Kaunitz et al. odds.

From equation (18), we see that odds o_i only influence X_i for winning bets. Therefore, we generate a vector of 125 odds (the number of winning bets), such that be achieve $\overline{X} = 0.0723$ with these winning bets and the $265 - 125 = 140$ losing bets, while minimising the test statistic of the Kolmogorov-Smirnov test applied to these generated odds and the fit Lognormal distribution: that is, we maximise the fit of the generated odds to our odds, subject to them resulting in the desired profit. To perform this optimisation, we provide an initial-guess vector, which we randomly draw from the Lognormal distribution, ensuring all values are in the range found in the odds from the main run. Because of this random guess, we perform the optimisation 100 times, as it converges to different solutions depending on the initial guess. After generating this vector of odds, we calculate the Tryfos et al Z-score. The mean is 0.698, with a standard deviation of 0.084, minimum 0.412 and maximum 0.858. A Z-score of 0.858, corresponds to a one-sided p-value of 0.195, meaning there is an 19.5% chance that the Kaunitz et al. profits are random, even in the most extreme scenario: certainly not significant by any accounts. And these

calculations assume that all bets were made with only one bookmaker. In reality we know that at least eight bookmakers were used, where the betting with each would then have much smaller sample sizes, and therefore the probability of bookmakers restricting based on profitable betting would be substantially lower.

Instead, Kaunitz et al. were most certainly restricted due to their furtive profit-seeking. If they are able to observe the prices offered by various bookmakers, then the bookmakers definitely do so as well. They are therefore fully aware that Kaunitz et al. only place bets with them if they offer the longest odds on the market. This transparent behaviour is among a list of obvious tells which bookmakers utilise in a rule-based manner to easily identify customers that are bad for business. Both Wong (2001) and Buchdahl (2022) provide such lists, with advice including to make bets as soon as an account is created; to not merely abuse bonuses and never bet otherwise; to never hedge offsetting bets both with the same bookmaker; to use the bookmaker's corresponding online casino and to make accumulator bets occasionally; and to not make bets with bizzare stake sizes (likely to occur with Kelly staking, betting an obscure percentage of some arbitrary capital level). Predominant, however, among their advice is to, one, as noted, not bet with a bookmaker only if they offer the longest odds, and, two, to not bet on lower-level leagues and events. Bookmakers desire business from the everyday sports fan, who watches from home and places a small wager on the event to make it more interesting for themselves: no fan is sufficiently dedicated to engage with the Belarusian Cup, the Macedonian First League as well as the Oman Sultan Cup, which are three competitions that Kaunitz et al. bet on, making their restriction entirely predictable. To clarify, a punter might be interested in any one of these leagues, but for one person to be engaged with all three is most unlikely.

Given this principle of not betting with bookmakers only if they offer the most favourable price in the market, the returns to our quasi-arbitrage model become unrealistic. We thus redid our analysis, however, instead betting not with the longest odds, but with the nth longest odds, as illustrated in Table 11. We use the model specified in equation (15), with the same α_j parameters as fit above. We backtest across the entire dataset, not only the test data, knowing that, since the parameters were fit to the longest odds, the profits calculated here are not overfit to training data. Evidently, returns are still positive when betting with the second, third, fourth and fifth longest bookmakers. Interestingly, returns increase in n, other than for the fifth longest bookmakers, while returns are significant at the 5% level for $n = 3$ (two-tailed) and 4 (one-tailed), but not for $n = 2$, where returns are just short of significance at the one-tailed 10% level. Certainly, the number of profitable bets decreases substantially, but the main point is that returns remain profitable, which is encouraging and suggests that profits can be made without attracting bookmaker attention immediately. This is corroborated by findings among the survey of betting models outlined above and described in the appendix (subsection "Surveying Betting Models"). Alesina (2021, 2022) tests his betting strategies at mean odds, generating profitable returns. Ramirez et al. (2023) similarly tested their strategy at both longest odds and Bet365 odds, receiving $r = 0.0553\%$ and $s = 12.02\%$ when using only this one bookmaker. Additionally, the accuracy-based models for Home/Away-style markets outperform the entire range of spreads set for these markets, meaning profitable returns are in theory possible at mean odds and even below. There is therefore substantial evidence suggesting that there are no

nth Bookmaker	r	Bets	Z-Score
2	2.46%	3,820	1.269
3	5.79%	1,650	2.089
4	7.76%	679	1.794
5	3.49%	255	0.492

Table 11: Linear returns r when betting using the optimal quasi-aribtrage model at the nth best odds. Returns from entire dataset.

constraints imposed on the prospective sports-betting investor to bet exclusively at the longest odds, allowing them to avoid notice and restriction for an extended period of activity.

An additional attribute to the nth bookmaker analysis is that it addresses a potential issue with our odds dataset, where prices are gathered by a third-party scraper, rather than having the prices sent directly from the bookmakers. This creates a host of potential issues, where these prices might be those most recently offered by the bookmakers, without necessarily being those offered at the time of collection; there might be ambiguity regarding when the market closes, whereby it is difficult to identify exactly which odds are the closing odds; bookmakers might retroactively adjust odds, accounting for a previous error or similar; or the odds might simply be misquoted. That our model is profitable even at shorter odds thus demonstrates that these occasional data issues are unlikely to influence returns significantly.

10.2 Staking Limits

We now discuss the second main market friction, that stake limits on bets are imposed by all bookmakers. The ceilings are set by default, varying between bookmakers and markets, and are familiarly lowered for restricted customers. The website maxpayout.co.uk tracks these figures for UK bookmakers across sports and leagues. For English football, Bet365 has the highest maximum payout of £2,000,000, shared with William Hill, while 888 Sport and Unibet have a maximum payout of £250,000, which is the lowest. For horse racing, the highest maximum payout is £1,000,000, offered by seven of the thirteen covered bookmakers, with 888 Sport and Unibet again the most miserly, paying out, at most, £100,000. These limits are of course still rather generous, especially given that sports betting is a large sample-size process, where a fortune is made not over a single bet but over a prolonged period of betting. However, these are only the upper limits to the largest and most significant competitions (where pricing is therefore the most accurate and the edges for bettors the most scarce). Considering Bet365, their website bet365.com provides detailed information regarding maximum winnings for different competitions and even markets. Quoted in the Swedish currency SEK, the maximum payout for the English Premier League is 20,000,000 SEK, which with current exchange rate equates roughly to £1,470,000. We know that this figure is actually £2,000,000, from above, and can therefore assume that the 500,000 SEK maximum winnings for "All other [soccer] Competitions not listed", equivalent to £37,000, is more likely to be a maximum payout of £50,000 in the UK, 2.5% of the overall soccer maximum payout. For sub-markets, the maximum winnings for "Player to be Booked/Sent Off" and "Player to Start Match" are 250,000 SEK, or £25,000 by our approximation. Again, still large numbers, and we would never actually bet on any of the "other Competitions not listed" since this is an easy tell for the bookmakers; the point is that rate limits vary between bookmakers and markets, and that they can shrink rapidly the more obscure the sport and market becomes.

Using the max-winnings data from Bet365, we can estimate a level of investment capital beyond which the sports-betting asset becomes decreasingly useful, where diminishing returns to scale commences. For American football, the maximum winnings are £500,000 and £100,000, respectively, for the NFL and CFL. For baseball, the MLB maximum winnings are £500,000 while for other leagues this number is £50,000. For basketball, bettors can win at most £500,000 on the NBA and £25,000 at most for other leagues. For cricket, the maximum is £250,000, while for tennis it is £500,000. There is a broad distribution, as noted, across soccer leagues. In our sports-betting asset simulation (see Figure 21), returns were aggregated by sport; therefore, we calculate average maximum payoffs for each sport. Considering Table 16 in the appendix, 92.8% of American football games (with the competitions we consider)

are NFL games, with 7.8% being CFL games, which gives an average maximum winnings of $92.8\% \cdot$ 500,000 $+ 7.8\% \cdot$ 100,000 $\approx$ £471,000 for American football bets, assuming bets are uniformly distributed across competitions (which is probably not the case in practice, as pricing is more accurate for superior competitions and therefore edges rarer, but for the sake of this approximate estimation it suffices). Using the same method, the league-averaged maximum winnings by sport are roughly £247,000 for baseball, £150,000 for basketball, £954,000 for soccer, and then £250,000 for cricket and £500,000 for tennis where maximum payouts are competition-agnostic.

In general, these maximum payouts apply to individual bets, but also to cumulative winnings across a 24-hour period, whereby if several bets win and the combined winnings exceed the given payout ceiling, only this maximum payout is received. Therefore, we consider the average number of bets placed per sport, per day. In the sports-betting asset simulation, the mean monthly bets per sport (rounded) were 10 for American football, 465 for baseball, 80 for basketball, 35 for cricket, 215 for soccer and 55 for tennis. With 30-day months, rounding to nearest non-zero integer, this gives 1, 16, 3, 1, 7 and 2 bets per day, respectively. Given that there is a maximum cumulative daily payout and we are making multiple bets, we want to decrease our stakes so that we are unable to exceed this ceiling. For instance, if we make the daily-average two tennis bets (on two, independent games), both at $o = 2$ with stakes of £300,000, and assume that these bets are individually fair with $p = 0.5$, our expected value for this pair of bets would be negative, as there is a 25% chance of losing both for a return of -£600,000, a 50% chance of winning one and losing one for a return of 0, and a 25% chance of winning both, giving a fair return of £600,000 but an actual return of the maximum winnings £500,000, and thus an expected value of -£25,000. We therefore get the league-averaged maximum per-bet winnings by sport of £471,000 for American football, £50,000 for basketball, £15,438 for baseball, £250,000 for cricket, £136,286 for soccer and £250,000 for tennis, each found by dividing the league-averaged maximum winnings for each sport by that sport's mean bets per day. To calculate the maximum stake per bet for a sport, we can then divide by $\overline{o} - 1$, where $\overline{o}$ is the mean odds for that sport, namely 2.370 for American football, 2.463 for basketball, 2.061 for baseball, 1.939 for cricket, 2.712 for soccer (table 5) and 4.096 for tennis, giving average stakes of £343,795, £34,176, £14,550, £266,240, £79,606 and £80,749, respectively. The bet-weighted mean stake is then $\frac{1 \cdot 343{,}795 + 16 \cdot 34{,}176 + 3 \cdot 14{,}550 + 1 \cdot 266{,}240 + 7 \cdot 79{,}606 + 2 \cdot 80{,}749}{1+16+3+1+7+2} =$ £63,975. This is an estimate of the average stake for a linear-staking system, which is important as the linear-returns sports-betting asset simulation assumes that the stake-size stays constant.

When betting, one should only stake a fraction of the current investment capital on a given bet, due to the large per-bet volatility of returns. This is irrespective of whether linear or geometric staking is in use. For the quasi-arbitrage model discussed above, the mean Kelly fraction was 0.0357, which, if we assume that this is the fraction used on average for all sports in the sports-betting asset, gives sports-betting capital of $63{,}975/0.0357 =$ £1,792,000. This is the level of sports-betting capital where our optimal Kelly bet converges with the maximum allowed bet size: below this point, we will bet less than the £63,975 maximum bet, beyond it our maximum bet becomes an increasingly smaller fraction of our capital, eventually becoming trivial, in a rapid decreasing-returns-to-scale effect. For a wealth-manager considering the sports-betting asset for investment, if they adopt the value-at-risk portfolio optimised for a 5% chance of a loss of 5% or more in a given month, they will allocate 71.59% of their assets under management to this asset (see Table 10), meaning the total capital threshold becomes $1{,}792{,}000/0.7159 \approx$ £2,500,000. This is certainly not an insubstantial number, far greater than the wealth of most individuals. However, wealth managers and hedge funds have huge investment capital: the 251st through 500th largest wealth managers have assets under management of

$25.5 billion on average (Thinking Ahead Institue, 2023), while the top 100 hedge funds all manage at least $1.5 billion (Pensions and Investments, 2023). These firms are therefore unlikely to invest at this much lower scale – where sports betting essentially becomes an increasingly illiquid market – especially considering the logistical issues of betting with bookmakers that restrict winners. This analysis of an investment-capital ceiling, conducted with Bet365, assumes that this bookmaker will not ban profitable customers, which is of course unreasonable. The point has been to demonstrate how an upper limit exists even in an idealistic scenario, rather than to identify some concrete, physical boundary. There are of course ways to create new accounts once banned, but when betting at the enormous scale discussed here, the bookmakers will scrutinise such activity more closely and re-ban more swiftly. Conversely, for individuals impervious to decreasing returns to scale, having wealth significantly below £2,500,000, sports-betting investment is more feasible. The ever-present issue of restrictions and bans remains, of course, but when betting at a smaller scale, bookmakers will be less vigilant, and if care is taken to bet stealthily, the frequency of these obstacles will decrease. The main point here is that, while the scale of investment will deter large asset managers, it will not be an issue for retail-investor-type bettors – whether they can then navigate bookmaker surveillance is a different question.

10.3 Sharp Bookmakers, Syndicates, and Technical Barriers to Entry

The above discussion has centred on Bet365 and its peers: William Hill, 888 Sport, Unibet and the like. These are what Buchdahl (2022) and the lay betting community in general refer to as soft bookmakers (Livori, 2014; Pinnacle, 2024), as opposed to sharp bookmakers. Soft bookmakers perform limited modelling and event-pricing of their own, instead more or less copying the odds of the autonomously price-setting sharp bookmakers. Another policy difference between the two categories of bookmaker is that, as discussed, the soft bookmakers restrict and ban winning customers, while sharp bookmakers, conversely, do not. They instead embrace winning customers and, importantly, track their activity, so that these on-average winning bets can guide the price-discovery process: the sharp bookmakers can host winning customers if in return their prices are the most accurate, letting them, therefore, extract greater profits from non-winning customers. This is the crux: if a bettor can find a successful betting strategy with a sharp bookmaker, they can exploit it without bookmaker intervention; however, the sharp bookmaker's prices are more calibrated than the soft bookmaker's, making systematic edges more difficult to identify. Quasi-arbitrage, for instance, will never be possible with sharp bookmakers; much more likely, they will be the bookmaker whose implied probabilities are taken as the true forecast (equation 6). As mentioned above, models exist which have made profits with soft bookmakers without betting at the longest odds, which suggests that sharp bookmakers, too, might be beatable: indeed, by their own admission, sharp bookmakers host and therefore acknowledge winning bettors. Strategies, for instance, in the same innovative spirit as the inspiring "Betting on a buzz: Mispricing and inefficiency in online sportsbooks" (Ramirez et al., 2023) – where the authors forecast tennis matches using data for Wikipedia-page visits in the lead-up to the game, for each of the two competitors – are likely to succeed, utilising sources of information which the sharp bookmakers and their winning bettors do not.

Regarding successful betting, there exists evidence of organised groups of professional bettors, known as betting syndicates, which combine their betting experience and resources to collectively beat bookmakers. Although the word syndicate bears some negative connotations, the betting syndicates discussed here operate legally. One insightful account is provided by prolific sports-betting scholar William Ziemba, in "Parimutuel Betting Markets: Racetracks and Lotteries Revisited" (Ziemba, 2023), where he discusses his experiences working with, and

consulting for, various horse-racing syndicates in Asia. He relates interacting with Bill Benter, a professional gambler who made a vast fortune running syndicates, with cumulative profits approaching $1 billion across a thirty-year endeavour. Ziemba describes how the business environment for syndicates has become harsher, with increased competition (not necessarily by other syndicates, simply for betting returns in general), and with bookmakers improving their forecasting. However, profitability is still possible, as "the syndicates continue [to] trade in many markets today ... such as Japan and Korea as well as in the US, Canada and Europe". Finally, he touches on the third financial friction mentioned briefly above, the technical barriers to entry, where data-gathering, model-construction and trade-automation are all laborious and intellectually-demanding pursuits. He notes that his "personal experience consulting extensively for one other syndicate is that the setup cost for the research and computer implementation is a major time and financial undertaking", making professional sports betting, or for our purposes, sports betting as a wealth-management avenue, less feasible for the majority of individuals. Certainly these skills are possible to develop, but the opportunity cost of this time-demanding pursuit will likely be formidable and ultimately obstructive for most.

To conclude this section, we have considered the two main financial frictions in the sports-betting market, namely, the restriction of winners and ostentatious profit-seekers, as well as the upper limit set by bookmakers on the winnings that they pay out, which introduces in turn a limit to the maximum stake per bet, and therefore creates decreasing returns to scale when investors with large assets under management seek to diversify with the sports-betting asset. We have also provided discussions of sharp versus soft bookmakers, betting syndicates and the technical barriers to entry which exist in the market. In previous sections, the sports-betting asset has performed admirably in comparison with conventional financial investments; however, this section has provided some realism, forewarning that returns to sports betting, all else being equal, are not the same as returns to stocks or bonds. Developing a robust model which can beat sharp bookmakers, or exercising great caution and prudence while picking off soft bookmakers, are both viable avenues for medium-wealth individuals. However, both endeavours are exacting in their respective ways, and it is unclear how the ratio of financial returns to required effort compares between investment in sports betting contra more typical markets.

11 Conclusion

In this paper, we have outlined in some detail the most essential aspects of the sports-betting market. We rigorously developed a quasi-arbitrage trading model which produced favourable returns in an out-of-sample backtest. This was a proof of concept, demonstrating the relative ease with which sports-betting profits can be generated. We then conducted a survey of betting models in the academic literature, creating a distribution of models across sports, characterised by their per-bet returns and the availability of these bets. Continuing this analysis, we collected detailed data for odds and games for different sports, and together with the array of models simulated a high-veracity sports-betting asset, an approximation of what investment returns might look like for a broad-scope betting strategy. Equipped with these returns, we construct a mean-variance portfolio, integrating the sports-betting asset with stocks, bonds and other conventional assets, as well as with art, wine and other popular alternative investments. Optimising and evaluating this diversified portfolio in various scenarios, including a value-at-risk analysis, the sports-betting asset uniformly contributed the most value to the portfolio, as measured by the increase in the portfolio's Sharpe ratio when an asset is removed and then returned to the asset pool. Finally, we consider financial frictions in the sports-betting market, discussing the restric-

tion of naive and transparent arbitrageurs, as well as the upper limits that bookmakers place on bettors' staking. This latter friction imposes a ceiling to the level of investment capital allocated towards sports betting, beyond which a diminishing returns to scale effect commences, a friction which essentially eliminates sports betting from being a relevant market for high assets-under-management investment bodies. For technically-shrewd, moderately-wealthy investors, on the other hand, we conclude that sports betting may have financial upside. This would, however, require the investor outsmarting either the welcoming-yet-efficient sharp bookmakers, or the profligate-yet-prohibitive soft bookmakers: a tractable but hardly elementary problem.

Lacking the resources to perform many of the necessary analyses in practice, we have resorted to simulation techniques and Fermi-style approximations. The returns profile of the sports-betting asset depends entirely on the models that we happened to survey, the number of sports chosen for analysis and the competitions included for each sport. Our decisions were hardly arbitrary, choosing six of the most popular and gambled-upon sports in the world, paired with their largest leagues, while the presented models were those found in an exhaustive search across the past half decade of betting literature. However, it remains an approximation nonetheless. A preferable methodology would be to survey the returns of professional bettors and betting syndicates; however, survivorship bias aside, data for the successful practitioners is understandably proprietary and inaccessible. In a similar vein, we estimate £2,500,000 as an inflection point where bets become increasingly trivial fractions of investable capital, reducing the relative attractiveness of the low-liquidity sports-betting market to wealth managers. Ignoring the methodology with which this figure was attained (any number of calculations could just as well apply), the estimate is equally dependent on the assumptions, inputs and results of the aforementioned simulation as is the portfolio integration, and would therefore similarly require data from the field to ascertain the exact value with more confidence. Naturally, the general existence of this threshold is guaranteed by bookmaker staking limits, with the overarching argument of sports betting being unattractive to large wealth managers therefore robust.

In conclusion, theoretical and simulatory approximations are only so good, particularly in our especially practice-embedded context. To truly evaluate sports betting as an alternative investment, further research would be directed towards creating a real-world sports-betting asset, one invested in beyond backtests and paper trades, with actual bookmakers. Only then can returns be judged faithfully and compared with other assets on common ground. In the interim – for what our results are worth – we conclude that sports betting can indeed be a profitable asset for an investor of the correct disposition.

12 References

Alesina, Davide. "An Effective Approach for Exploiting the Inefficiencies of the Italian Football Betting Market". *The Journal of Gambling Business and Economics*, vol. 14, no. 1, 2021.

———. "Exploiting the main biases in American soccer betting markets." *The Journal of Gambling Business and Economics*, vol. 15, no. 1, 2022, pp. 45–54.

Arscott, Robert. "Market efficiency and censoring bias in college football gambling". *Journal of Sports Economics*, vol. 24, no. 5, 2023, pp. 664–89.

ATP Tour. "Final 2024 ATP Calendar", 2024, www.atptour.com/-/media/files/final-2024-atp-calendar.pdf. Accessed 1 Apr. 2024.

aussportsbetting.com. "Historical NFL Results and Odds Data", www.aussportsbetting.com/data/historical-nfl-results-and-odds-data/. Accessed 31 Mar. 2024.

———. "Historical Twenty20 Big Bash Results and Odds Data", www.aussportsbetting.com/data/historical-twenty20-big-bash-results-and-odds-data/. Accessed 31 Mar. 2024.

Baddour, Ralph. "Is There a Correlation Between World Cups and S&P 500 Performance?" *Available at SSRN 1624822*, 2010.

Baker, Rose D, and Ian G McHale. "Optimal betting under parameter uncertainty: Improving the Kelly criterion". *Decision Analysis*, vol. 10, no. 3, 2013, pp. 189–99.

bet365.com. "Maximum Winnings", help.bet365.com/en/stand-alone-pages/maximum-winnings. Accessed 5 May 2024.

betmonitor.com. "Comparing Odds the Easy Way", www.betmonitor.com/. Accessed 1 Feb. 2024.

Bhattacharjee, Dibyojyoti, and Priyanka Talukdar. "Predicting outcome of matches using pressure index: evidence from Twenty20 cricket". *Communications in Statistics-Simulation and Computation*, vol. 49, no. 11, 2020, pp. 3028–40.

Brier, Glenn W. "Verification of forecasts expressed in terms of probability". *Monthly weather review*, vol. 78, no. 1, 1950, pp. 1–3.

Buchdahl, Joseph. *Monte Carlo or Bust*. Oldcastle Books, 2022.

Burton, Benjamin J, and Joyce P Jacobsen. "Measuring returns on investments in collectibles". *Journal of Economic Perspectives*, vol. 13, no. 4, 1999, pp. 193–212.

Canadian Football League. "CFL Schedule 2024 Season", 2024, www.cfl.ca/schedule/2024/. Accessed 31 Mar. 2024.

checkbestodds.com. "ATP Australian Open - Historical Betting Odds", checkbestodds.com/tennis-odds/historical-odds-atp-australian-open. Accessed 31 Mar. 2024.

Chen, Chi-Wen. "Construction of the Winner Predictive Model in Major League Baseball Games: Use of the Artificial Neural Networks". *Sport. Exerc. Res*, vol. 16, 2014, pp. 167–81.

chrono24.com. "ChronoPulse - Watch Index", www.chrono24.com/chronopulse.htm. Accessed 21 Mar. 2024.

Chu, P. K. K. "Study on the diversification ability of fine wine investment". *Journal of Investing*, vol. 23, 2014, pp. 123–39.

Constantinou, Anthony C. "Investigating the efficiency of the Asian handicap football betting market with ratings and Bayesian networks". *Journal of Sports Analytics*, vol. 8, no. 3, 2022, pp. 171–93.

Cortis, Dominic. "Betting Markets: Defining odds restrictions, exploring market inefficiencies and measuring bookmaker solvency". 2016. U of Leicester, **p**hdthesis.

Cui, Andrew Y. "Forecasting Outcomes of Major League Baseball Games Using Machine Learning". *University of Pennsylvania: Philadelphia, PA, USA*, 2020.

cultx.com. "Cult Wines Global Index", www.cultx.com/indices. Accessed 21 Mar. 2024.

Da Costa, Igor Barbosa, et al. "Forecasting football results and exploiting betting markets: The case of "both teams to score"". *International Journal of Forecasting*, vol. 38, no. 3, 2022, pp. 895–909.

Davis, Mark, and Sébastien Lleo. "Fractional Kelly strategies for benchmarked asset management". *The Kelly Capital Growth Investment Criterion: Theory and Practice*, World Scientific, 2011, pp. 385–407.

Demšar, Janez. "Statistical comparisons of classifiers over multiple data sets". *The Journal of Machine learning research*, vol. 7, 2006, pp. 1–30.

Dickie, Mark, et al. "Price determination for a collectible good: The case of rare US coins". *Southern Economic Journal*, 1994, pp. 40–51.

Dimson, Elroy, and Christophe Spaenjers. "Ex post: The investment performance of collectible stamps". *Journal of Financial Economics*, vol. 100, no. 2, 2011, pp. 443–58.

Dixon, Mark J, and Stuart G Coles. "Modelling association football scores and inefficiencies in the football betting market". *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 46, no. 2, 1997, pp. 265–80.

Dixon, Mark J, and Peter F Pope. "The value of statistical forecasts in the UK association football betting market". *International journal of forecasting*, vol. 20, no. 4, 2004, pp. 697–711.

Edmans, Alex, et al. "Sports sentiment and stock returns". *The Journal of finance*, vol. 62, no. 4, 2007, pp. 1967–98.

Epstein, Edward S. "A scoring system for probability forecasts of ranked categories". *Journal of Applied Meteorology (1962-1982)*, vol. 8, no. 6, 1969, pp. 985–87.

Fama, Eugene F. "Efficient capital markets". *Journal of finance*, vol. 25, no. 2, 1970, pp. 383–417.

football-data.co.uk. "Data Files: All Countries", www.football-data.co.uk/downloadm.php. Accessed 15 Jan. 2024.

fred.stlouisfed.org. "Market Yield on U.S. Treasury Securities at 1-Month Constant Maturity, Quoted on an Investment Basis", fred.stlouisfed.org/series/GS1M. Accessed 1 Apr. 2024.

Graddy, Kathryn, and Philip E Margolis. "Fiddling with value: violins as an investment?" *Economic Inquiry*, vol. 49, no. 4, 2011, pp. 1083–97.

Gu, Wei, et al. "A game-predicting expert system using big data and machine learning". *Expert Systems with Applications*, vol. 130, 2019, pp. 293–305.

Hassan, Andrés Ramírez, and Mateo Graciano Londoño. "Profiting from the English premier league: predictive elicitation, the Kelly criterion, and black swans". *International Journal of Sport Finance*, vol. 12, no. 4, 2017, pp. 386–95.

Huang, Mei-Ling, and Yun-Zhi Li. "Use of machine learning and deep learning to predict the outcomes of major league baseball matches". *Applied Sciences*, vol. 11, no. 10, 2021, p. 4499.

icc-cricket.com. "Fixtures and Results", 2023, www.icc-cricket.com/fixtures-results. Accessed 30 Mar. 2024.

Jorion, Philippe. *Value at risk: the new benchmark for managing financial risk*. McGraw-Hill, 2007.

Jullien, Bruno, and Bernard Salanié. "Measuring the incidence of insider trading: A comment on Shin". *The Economic Journal*, vol. 104, no. 427, 1994, pp. 1418–19.

k500.com. "The K500 Index", k500.com/the-index. Accessed 21 Mar. 2024.

Kaunitz, Lisandro, et al. "Beating the bookies with their own numbers-and how the online sports betting market is rigged". *arXiv preprint arXiv:1710.02824*, 2017.

Kelly, John L. "A new interpretation of information rate". *the bell system technical journal*, vol. 35, no. 4, 1956, pp. 917–26.

Kendall, Maurice G. “A new measure of rank correlation”. *Biometrika*, vol. 30, nos. 1/2, 1938, pp. 81–93.

Kuester, Keith, et al. “Value-at-risk prediction: A comparison of alternative strategies”. *Journal of Financial Econometrics*, vol. 4, no. 1, 2006, pp. 53–89.

Kuypers, Tim. “Information and efficiency: an empirical study of a fixed odds betting market”. *Applied Economics*, vol. 32, no. 11, 2000, pp. 1353–63.

Kylämarkula, Martin. “Luxury watches as an alternative investment–A qualitative study on how watch specialists see the luxury watch market in Finland”. 2023.

Laurs, Dries, and Luc Renneboog. “My kingdom for a horse (or a classic car)”. *Journal of International Financial Markets, Institutions and Money*, vol. 58, 2019, pp. 184–207.

Law, David, and David A Peel. “Insider trading, herding behaviour and market plungers in the British horse–race betting market”. *Economica*, vol. 69, no. 274, 2002, pp. 327–38.

Levy, Nir. “The Effect of New York City Sports Outcomes on the Stock Market”. *Undergraduate Economic Review*, vol. 12, no. 1, 2015, p. 8.

Li, Shu-Fen, et al. “Exploring and selecting features to predict the next outcomes of MLB games”. *Entropy*, vol. 24, no. 2, 2022, p. 288.

Livori, Nikolai. “Sportsbook - Soft or Sharp”, 2014, www.linkedin.com/pulse/20140627134122-44323252-sportsbook-soft-or-sharp/. Accessed 5 May 2024.

Lyócsa, Štefan, and Tomáš Vỳrost. “To bet or not to bet: a reality check for tennis betting market efficiency”. *Applied Economics*, vol. 50, no. 20, 2018, pp. 2251–72.

MacLean, Leonard C, et al. “Growth versus security in dynamic investment analysis”. *Management Science*, vol. 38, no. 11, 1992, pp. 1562–85.

MacLean, Leonard C, et al. “Medium term simulations of the full Kelly and fractional Kelly investment strategies”. *The Kelly Capital Growth Investment Criterion: Theory and Practice*, World Scientific, 2011, pp. 543–61.

Maclean, Leonard C, et al. “Time to wealth goals in capital accumulation”. *Quantitative Finance*, vol. 5, no. 4, 2005, pp. 343–55.

Maher, Michael J. “Modelling association football scores”. *Statistica Neerlandica*, vol. 36, no. 3, 1982, pp. 109–18.

Markowitz, H.M. “Portfolio Selection”. *The Journal of Finance*, vol. 7, no. 4, 1952, pp. 77–91.

Massey Jr, Frank J. “The Kolmogorov-Smirnov test for goodness of fit”. *Journal of the American statistical Association*, vol. 46, no. 253, 1951, pp. 68–78.

Mattera, Raffaele. “Forecasting binary outcomes in soccer”. *Annals of Operations Research*, vol. 325, no. 1, 2023, pp. 115–34.

maxpayout.co.uk. “Maximum Payout by Sport”, www.maxpayout.co.uk/. Accessed 5 May 2024.

McAlister, Anna R, and T Bettina Cornwell. “Collectible toys as marketing tools: understanding preschool children’s responses to foods paired with premiums”. *Journal of Public Policy & Marketing*, vol. 31, no. 2, 2012, pp. 195–205.

Moskowitz, Tobias J. “Asset pricing and sports betting”. *The Journal of Finance*, vol. 76, no. 6, 2021, pp. 3153–209.

myartbroker.com. “MAB100 The Print Index”, www.myartbroker.com/mab-100. Accessed 28 Mar. 2024.

Nemenyi, Peter Bjorn. *Distribution-free multiple comparisons.* Princeton U, 1963.

Norton, Hugh, et al. “Yes, one-day international cricket ‘in-play’trading strategies can be profitable!” *Journal of Banking & Finance*, vol. 61, 2015, S164–S176.

oddschecker.com. “English Premier League Betting Odds”, www.oddschecker.com/football/english/premier-league. Accessed 29 Mar. 2024.

———. “Win Market Indian Premier League”, www.oddschecker.com/cricket/ipl. Accessed 30 Mar. 2024.
———. “Win Market MLB”, www.oddschecker.com/baseball/mlb. Accessed 29 Mar. 2024.
oddsportal.com. “NHL Fixtures and Betting Odds”, www.oddsportal.com/hockey/usa/nhl/. Accessed 30 Mar. 2024.
Ottaviani, Marco, and Peter Norman Sørensen. “The favorite-longshot bias: An overview of the main explanations”. *Handbook of Sports and Lottery markets*, 2008, pp. 83–101.
———. *The Timing of Bets and Favorite-longshot Bias*. Citeseer, 2004.
Paton, David, and Leighton Vaughan Williams. “Forecasting outcomes in spread betting markets: Can bettors use ‘quarbs’ to beat the book?” *Journal of Forecasting*, vol. 24, no. 2, 2005, pp. 139–54.
Pensions and Investments. “The Largest Hedge Fund Managers 2023”, 2023, www.pionline.com/largest-hedge-funds/2023/full-list. Accessed 5 May 2024.
Pinnacle. “Winners Welcome at Pinnacle”, 2024, www.pinnacle.com/betting-resources/en/educational/winners-welcome-at-pinnacle. Accessed 5 May 2024.
Ramirez, Philip, et al. “Betting on a buzz: Mispricing and inefficiency in online sportsbooks”. *International Journal of Forecasting*, vol. 39, no. 3, 2023, pp. 1413–23.
rarecoins101.com. “Collectors Rare Coin Index”, www.rarecoins101.com/rare-coin-index.html. Accessed 21 Mar. 2024.
Rousseeuw, Peter J. “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis”. *Journal of computational and applied mathematics*, vol. 20, 1987, pp. 53–65.
S&P Dow Jones Indices. “Dow Jones U.S. Real Estate Index”, 2024, www.spglobal.com/spdji/en/indices/equity/dow-jones-us-real-estate-index/#overview. Accessed 1 Apr. 2024.
———. “S&P 500®”, 2024, www.spglobal.com/spdji/en/indices/equity/sp-500/#overview. Accessed 1 Apr. 2024.
———. “S&P Cryptocurrency LargeCap Ex-MegaCap Index”, 2024, www.spglobal.com/spdji/en/indices/digital-assets/sp-cryptocurrency-largecap-ex-megacap-index/#overview. Accessed 1 Apr. 2024.
———. “S&P Global Developed Aggregate Ex-Collateralized Bond Index”, 2024, www.spglobal.com/spdji/en/indices/fixed-income/sp-global-developed-aggregate-ex-collateralized-bond-index/#overview. Accessed 1 Apr. 2024.
———. “S&P GSCI”, 2024, www.spglobal.com/spdji/en/indices/commodities/sp-gsci/#overview. Accessed 1 Apr. 2024.
Shank, Corey A. “NFL betting biases, profitable strategies, and the wisdom of the crowd”. *Shank, C.(2019). NFL Betting Biases, Profitable Strategies, and the Wisdom of the Crowd. International Journal of Sport Finance*, vol. 14, no. 1, 2018, pp. 3–12.
———. “NFL betting market efficiency, divisional rivals, and profitable strategies”. *Studies in Economics and Finance*, vol. 36, no. 3, 2019, pp. 567–80.
Sharpe, William F. “Mutual fund performance”. *The Journal of business*, vol. 39, no. 1, 1966, pp. 119–38.
Shin, Hyun Song. “Measuring the incidence of insider trading in a market for state-contingent claims”. *The Economic Journal*, vol. 103, no. 420, 1993, pp. 1141–53.
———. “Optimal betting odds against insider traders”. *The Economic Journal*, vol. 101, no. 408, 1991, pp. 1179–85.
———. “Prices of state contingent claims with insider traders, and the favourite-longshot bias”. *The Economic Journal*, vol. 102, no. 411, 1992, pp. 426–35.

Smith, Michael A, et al. “Do bookmakers possess superior skills to bettors in predicting outcomes?” *Journal of Economic Behavior & Organization*, vol. 71, no. 2, 2009, pp. 539–49.

Song, Kai, et al. “Making real-time predictions for NBA basketball games by combining the historical data and bookmaker’s betting line”. *Physica A: Statistical Mechanics and its Applications*, vol. 547, 2020, p. 124411.

sports-statistics.com. “MLB Historical Odds and Scores Datasets”, sports-statistics.com/sports-data/mlb-historical-odds-scores-datasets/. Accessed 30 Mar. 2024.

sportsbookreviewsonline.com. “NBA Scores and Odds Archives”, www.sportsbookreviewsonline.com/scoresoddsarchives/nba/nbaoddsarchives.htm. Accessed 30 Mar. 2024.

Štrumbelj, Erik. “On determining probability forecasts from betting odds”. *International journal of forecasting*, vol. 30, no. 4, 2014, pp. 934–43.

Stübinger, Johannes, et al. “Machine learning in football betting: Prediction of match results based on player characteristics”. *Applied Sciences*, vol. 10, no. 1, 2019, p. 46.

Student. “Errors of routine analysis”. *Biometrika*, 1927, pp. 151–64.

Thinking Ahead Institue. “The World’s Largest 500 Asset Managers”, 2023, www.thinkingaheadinstitute.org/content/uploads/2023/10/PI-500-2023-1.pdf. Accessed 5 May 2024.

Tryfos, Peter, et al. “The profitability of wagering on NFL games”. *Management Science*, vol. 30, no. 1, 1984, pp. 123–32.

Walsh, Conor, and Alok Joshi. “Machine learning for sports betting: should forecasting models be optimised for accuracy or calibration?” *arXiv preprint arXiv:2303.06021*, 2023.

Wheatcroft, Edward. “A profitable model for predicting the over/under market in football”. *International Journal of Forecasting*, vol. 36, no. 3, 2020, pp. 916–32.

Wikipedia. “2023 Cuban National Series”, 2023, en.wikipedia.org/wiki/2023_Cuban_National_Series. Accessed 1 Apr. 2024.

Williams, Leighton Vaughan. “Can Bettors Win?” *World Economics*, vol. 2, no. 1, 2001, pp. 31–48.

Wong, Stanford. *Sharp sports betting*. Pi Yee P, 2001.

Worthington, Andrew C, and Helen Higgs. “Art as an investment: Risk, return and portfolio diversification in major painting markets”. *Accounting & Finance*, vol. 44, no. 2, 2004, pp. 257–71.

Yahoo Finance. “Total Collectables (^CLCTBLS.REGA)”, finance.yahoo.com/quote/%5C%5ECLCTBLS.REGA/chart. Accessed 1 Apr. 2024.

Ziemba, William T. “Pari-mutuel betting markets: racetracks and lotteries revisited”. *Annual Review of Financial Economics*, vol. 15, 2023, pp. 641–62.

13 Appendix

13.1 Odds-Standardisation Techniques

13.1.1 Basic Standardisation

One of three methods evaluated by Štrumbelj, and the simplest form of standardisation, we have that $p_j = \frac{\pi_j}{\Pi}$, where margin is proportionally extracted from each implied probability.

13.1.2 Shin Standardisation

Our second model was determined to be optimal by Štrumbelj, and is perhaps the most famous in the sports betting literature. Proposed in a series of papers by Hyun Song Shin (Shin, 1991; Shin, 1992; Shin, 1993), this standardisation assumes that there exists a proportion z of insider traders, who have perfect knowledge regarding the outcome of the sporting event in question. It is with respect to these insiders that bookmakers set their margins, based on their true beliefs $\boldsymbol{p}$ according to

$$\pi_j = \sqrt{zp_j + (1-z)p_j^2} \sum_{i=j}^{n} \sqrt{zp_i + (1-z)p_i^2} \tag{31}$$

Given $\boldsymbol{\pi}$, we must extract $\boldsymbol{p}$. Jullien and Salanié (1994) showed this is quite tractable, with

$$p_j = \frac{\sqrt{z^2 + 4(1-z)\frac{\pi_j^2}{\Pi}} - z}{2(1-z)} \tag{32}$$

where one can optimise for the value of z which satisfies the probability condition $\mathbf{1}^{\boldsymbol{T}}\boldsymbol{p} = 1$.

13.1.3 Logit Regression Standardisation

The third and final model evaluated by Štrumbelj, for our three-outcome case (home (H), draw (D), away (A)), an ordered logistic regression was proposed, where

$$\begin{aligned} p_H &= P(\mu_2 < Y) \\ p_D &= P(\mu_1 < Y < \mu_2) \\ p_A &= P(Y < \mu_1) \end{aligned} \tag{33}$$

with the underlying continues process modelled as

$$Y = \beta_0 + \beta_1 \pi_H + \beta_2 \pi_D + \varepsilon \tag{34}$$

where π_A has been removed due to potential multicollinearity issues (since $\pi_H + \pi_D + \pi_A$ is roughly constant). The probability distribution P used to compare Y with the μ parameters is Φ, the cumulative distribution function of the standard normal distribution.

This model thus contains five parameters: $\mu_1, \mu_2, \beta_0, \beta_1, \beta_2$. The first third of the relevant dataset is set aside to optimise these parameters (for this and an additional regression model described below). Štrumbelj also trained this model in a rolling fashion, reoptimising the parameters after each game in the test sample, by adding this game to the training data. Due to computational limitations, this was not done here. Especially considering the non-significant difference between the static and rolling models shown in Štrumbelj.

13.1.4 Exponential Standardisation

Another commonly applied margin standardisation technique, $p_j = \pi_j^k$, where k is optimised such that the p_j sum to 1.

13.1.5 Price Difference Standardisation

Proposed by Law and Peel (2002) the price difference pd between two probabilities p_1 and p_2 is calculated as

$$pd = ln\left(\frac{1}{1-p_1}\right) - ln\left(\frac{1}{1-p_2}\right) \tag{35}$$

Utilised to quantify how large a price movement is, the rationale is that a move from 0.01 to 0.02 is much more significant than a move from 0.51 to 0.52, even though they have the same arithmetic difference. However, pd has not previously been used for standardisation purposes. We want to convert o_j to p_j, where the price difference will be $pd = ln\left(\frac{1}{1-\frac{1}{o_j}}\right) - ln\left(\frac{1}{1-p_j}\right)$ (since $\frac{1}{o_j} > p_j$, we enter p_1 as $\frac{1}{o_j}$). Rearranging for p_j gives

$$p_j = 1 - \frac{1}{e^{ln\left(\frac{1}{1-\frac{1}{o_j}}\right) - pd}} \tag{36}$$

where, similar to the Shin and Exponential models above, we optimise for the value of pd which sets the sum of the p_j to one. The thought process behind this method is that the margins have been set such that the odds for each outcome are equally far from the true probability, in the price difference sense.

13.1.6 Profit Standardisation

This is a method we propose, which is regression-based similar to that of Štrumbelj. As analysed extensively in the literature (for discussion, see, for instance, Law and Peel (2002), Ottaviani and Sørensen (2004); Ottaviani and Sørensen (2008), Williams (2001)), the favourite-longshot bias in sports betting is the finding that returns are worse at longer odds (lower-probability events), known as longshots, compared to the returns at shorter odds (higher-probability events), the favourites. This standardisation method seeks to make standardised implied probabilities compatible with this bias, calculating the probabilities which maximise the probability of seeing the expected returns, as per the favourite-longshot bias.

Across the relevant training data, the returns to all odds are binned. The returns to o_j is $\boldsymbol{I}o_j - 1$, where $\boldsymbol{I}$ is the indicator function set to 1 if the bet wins and 0 otherwise (see equation (19)). These returns are binned by π_j, the unstandardised implied probabilities, rather than o_j. Then, for each bin, we calculate a normal distributional for these returns, using the arithmetic mean return, and standard error of the mean (SEM) return (which gives a much more reasonable scale of variation than the standard deviation). For an input set of odds $\boldsymbol{o}$, for which we want to find the standardised implied probabilities $\boldsymbol{p}$, we minimise the objective function

$$\begin{aligned} &\arg\min_{\boldsymbol{p}} - \sum_{j=1}^{m} \ln(\varphi(p_j o_j - 1 \mid \mu_j, \sigma_j)) \\ &\text{s.t.} \quad \sum_{j=1}^{m} p_j = 1 \end{aligned} \tag{37}$$

across the m outcomes, where μ_j and σ_j are the mean and SEM for the binned returns which π_j belongs to, φ is the normal distribution probability density function, and $p_j o_j - 1$ is the expected value of a unit bet at odds o_j with probability p_j, which follows the φ distribution with parameters μ_j and σ_j. We are essentially maximising the fit between the probabilities, the odds, and the expected returns.

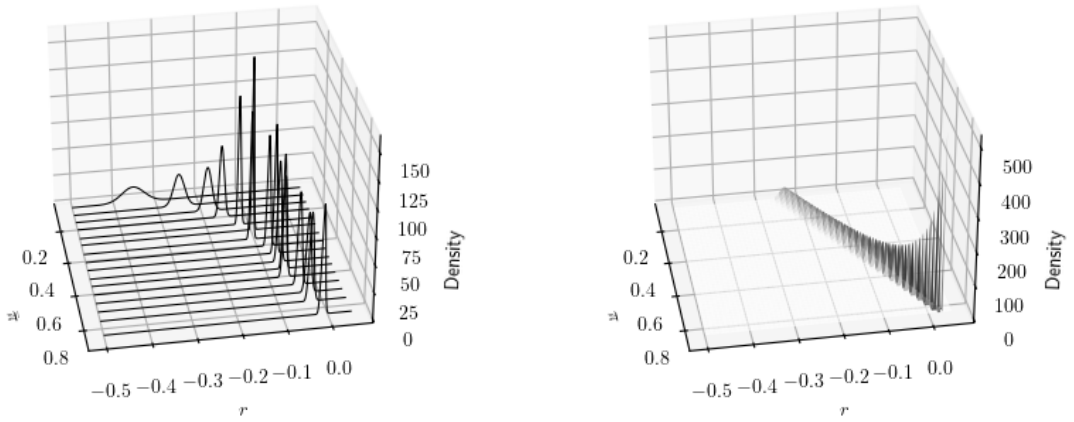

Figure 25: Discrete (left) and continuous (right) distributions of returns as a function of unstandardised implied probabilities.

This is the discrete version of the profit standardisation technique. It is discrete in the sense that the mean and SEM for odds o_j are taken from the corresponding bin. This means that odds at the upper threshold of the bin have the same returns distribution as those at the bottom of the bin, which is not realistic. Therefore, we create two linear regressions:

$$\begin{aligned} \mu &= \alpha_1 + \beta_1 \overline{\pi} + \varepsilon \\ \sigma &= \alpha_2 + \beta_2 \overline{\pi} + \varepsilon \end{aligned} \tag{38}$$

where each data point corresponds to a bin, the independent variable being the mean unstandardised implied probability $\overline{\pi}$ and the dependent variable the corresponding mean return and SEM, for the two regressions, respectively. This gives the continuous profit standardisation model:

$$\begin{aligned} &\arg\min_{\boldsymbol{p}} -\sum_{j=1}^{m} \ln(\varphi(p_j o_j - 1 \mid \alpha_1 + \beta_1 \frac{1}{o_j}, \alpha_2 + \beta_2 \frac{1}{o_j})) \\ &\text{s.t.} \quad \sum_{j=1}^{m} p_j = 1 \end{aligned} \tag{39}$$

The distributions of returns in the discrete and continuous senses are shown in Figure 25. Note that a quadratic regression would have been better for μ in the continuous model. However, it was deliberately kept linear, preferring a simpler, less parameterised model.

13.2 Surveying Betting Models

13.2.1 Soccer

13.2.1.1 Stübinger et al. (2019)

In their paper "Machine Learning in Football Betting: Prediction of Match Results Based on Player Characteristics", Stubinger and colleagues present a modelling-based strategy, using historical data on a team and player level to create machine learning models which predict event outcomes. Across twelve seasons' worth of games from the top two leagues in each of England, Spain, Germany, Italy and France, their model generates a $r = 1.58\%$ linear return. Their total profit is 9.47 units, meaning that they placed $9.47/0.0158 = 599$ bets. In a single season, the ten listed leagues have a total of 3,800 games combined (380 and 460 in England, 380 and 420 in Spain, 340 and 340 in Germany, 380 and 380 in Italy, and 340 and 380 in France), which over twelve seasons gives $12 \cdot 3{,}800 = 45{,}600$ games, making this model's bet selection $s = 599/45{,}600 = 1.314\%$.

13.2.1.2 Mattera (2023)

"Forecasting binary outcomes in soccer" presents machine learning methods for making profitable bets in the Red Card Yes/No, Over/Under and Both Teams To Score (BTTS) markets. We take data for BTTS, where the machine learning model had 92.44% accuracy on 353 bets over thirteen English Premier League seasons (380 games apiece), giving $s = 353/(13 \cdot 380) = 7.15\%$. To convert accuracy into a linear return, I gathered BTTS odds for ten upcoming Premier League games, as shown in Table 12.

Date	Home	Away	Yes	No
30-3-24	Newcastle	West Ham	1.52	2.63
30-3-24	Nottingham Forest	Crystal Palace	1.83	2
30-3-24	Bournemouth	Everton	1.62	2.38
30-3-24	Sheffield United	Fulham	1.67	2.2
30-3-24	Tottenham	Luton	1.65	2.38
30-3-24	Chelsea	Burnley	1.87	2
30-3-24	Aston Villa	Wolverhampton	1.63	2.38
30-3-24	Brentford	Man Utd	1.5	2.75
31-3-24	Liverpool	Brighton	1.6	2.45
31-3-24	Man City	Arsenal	1.68	2.3

Table 12: Odds for the subsequent ten Premier League games at the time of collection, gathered the 29th of March 2024 from online odds-comparison site oddschecker.com. Both Teams To Score market, longest odds available.

We take the average odds across each outcome, 1.657 for Yes and 2.347 for No. With an accuracy of 92.44%, this would (on average) give per-bet returns of 0.5317 for Yes and 1.17 for No. We take the lower value in 53.2%, meaning that this model has parameters $(r, s) = (0.532, 0.0715)$

13.2.1.3 Alesina (2021)

"An Effective Approach for Exploiting the Inefficiencies of the Italian Football Betting Market" presents a model leveraging the Favourite-Longshot Bias and two lesser-known biases, home-

team and league-winner scoring advantage, to make, as stated in the conclusion, 353 bets with an average return of $r = 9.31\%$ across 19 seasons of Italian top-flight soccer. This gives a total of $19 \cdot 380$ =7,220 games, and hence a bet selection $s = 3{,}533/7{,}220 = 4.89\%$. The author achieves this by betting at mean odds, and states that an additional 3-4% could have been made by betting at the longest odds. Therefore, we assign $r = 12.31\%$ to this paper.

13.2.1.4 Alesina (2022)

In his second paper presented here, "Exploiting the Main Biases in American Soccer Betting Markets", Alesina managed to provide strong returns while betting not at longest odds but at mean odds, with linear staking, achieving, as stated in their abstract, $r = 6.8\%$ over 567 bets. With their dataset consisting of 14,229 matches in total, this gives a bet selection of $s = 567/14{,}229 = 3.98\%$.

13.2.1.5 da Costa et al. (2022)

Over a dataset of 19,122 total games, the authors of "Forecasting football results and exploiting betting markets: The case of 'both teams to score'" present various betting models which achieve large profits. Table 5 in their paper presents these models, their per-bet returns and the number of bets placed. We select the model (classifier) with the highest value in the PRF column (equivalent to the $n \cdot r$ condition applied above, see Figure 5), giving $r = 3.25\%$, $s = 1{,}904/19{,}122 = 9.96\%$.

13.2.1.6 Hassan and Londoño (2017)

In their paper "Profiting From the English Premier League: Predictive Elicitation, the Kelly Criterion, and Black Swans", Hassan et al. use a Categorical Dirichlet model to forecast probabilities for match outcomes. They primarily use Kelly staking, but compare this to linear staking, particularly in Table 2, where, under the "Posterior" heading, they give the number of bets made and mean returns for the three Premier League seasons covered. Calculating total returns as $\sum_{i=1}^{3} n_i r_i$ for each of the three seasons, we get a total profit of $0.04 \cdot 220 + 0.02 \cdot 211 + 0.012 \cdot 188 = 15.276$. Dividing this by the total number of games $\sum_{i=1}^{3} n_i = 220 + 211 + 188 = 619$, gives $r = 15.276/619 = 0.0247\%$. With Premier League season consisting of 380 matches, bet selection is $s = 619/(3 \cdot 380) = 54.3\%$.

13.2.1.7 Constantinou (2022)

Focusing on the under-explored Asian Handicap market (where odds are set according to a scenario where a number of goals has been assigned to one of the teams as a head start), in "Investigating the efficiency of the Asian handicap football betting market with ratings and Bayesian networks", the author analyses Premier League odds from thirteen seasons, achieving a total profit of 393.36 units over 2,265 games, giving the $r = 393.36/2{,}265 = 17.37\%$ figure presented in Table 11. Bet selection is of course $s = 2{,}265/(13 \cdot 380) = 45.85\%$.

13.2.1.8 Wheatcroft (2020)

Following in the footsteps of seminal sports-betting papers by Maher (1982) and Dixon and Coles (1997), in "A profitable model for predicting the over/under market in football" Wheatcroft creates a team rating system based on goal-scoring and defending attributes. Where the aforementioned pioneers used their distributions of goal scoring to further predict match outcomes, the author directly applies them in the goal-based Over/Under market. In his abstract, he states

that his best strategy generates a per-bet return of 0.8%, which in Table 3 leads to a total accumulated profit of roughly 535, which (assuming the 0.8% figure is linear), gives $n = 535/0.008 =$ 66,875 bets. However, there are only 54,437 games in Wheatcroft's dataset. Therefore, although he does apply linear staking, we must assume that 0.8% is a geometric rate of return. Hence, we get that $n = log_{1.008}535 = 788.4$ bets, and a bet selection of $s = 788.4/54{,}437 = 1.45\%$. Finally, linear returns are $r = 535/788.4 = 69.7\%$.

13.2.2 American Football

13.2.2.1 Shank (2018)

The paper "NFL Betting Biases, Profitable Strategies, and the Wisdom of the Crowd" presents a betting strategy based on identified behavioural biases among regular punters. Developing a contrarian strategy, betting against the crowd in cases when there is a slight but not overwhelming consensus (Table 4, Panel A, first two rows), gives $n = 72 + 87 + 59 + 82 = 300$ bets which, with the total available bets equal to 527, is a bet selection of $s = 300/527 = 56.9\%$. Total profit is $780 + 1,710 =$ 2,490, linearly, with 110-unit bets. Therefore, per-bet returns are $r = (2{,}490/110)/300 = 6.63\%$.

13.2.2.2 Shank (2019)

In his second (as presented here) extensive study "NFL Betting Market Efficiency, Divisional Rivals, and Profitable Strategies", Shank bets against the NFL spread markets by considering a vast range of quantitative and qualitative data (not exclusively sports-related) in a probit model. From Table 6, Panel A, the best model generates total profit of 3,470 from $n = 599 + 513 =$ 1,112 bets. Stakes are 110 units (see Tryfos et al. (1984) explanation above), meaning that per bet returns are $r = (3{,}470/110)/1{,}112 = 2.84\%$. There were 3,582 games played across the 15 seasons between 2003 and 2017 (the covered span), while bets were made for the 10 seasons between 2007 and 2016, meaning that the test dataset was two thirds of the total dataset. Therefore, the bet selection is $s = 1{,}112/(3{,}582 \cdot 2/3) = 46.6\%$.

13.2.2.3 Arscott (2023)

Analysing data on college football, in his paper "Market Efficiency and Censoring Bias in College Football Gambling" Arscott uses a probit model on the team totals market with great effect. Splitting his data into five non-overlapping groups, each acting as test data with the other four as the corresponding training data, Table 6 provides both linear and geometric (Kelly) returns. Total returns are $\sum_{i=1}^{5} n_i r_i$ for each of the five groups, equal to $149 \cdot 0.013 + 117 \cdot 0.77 + 163 \cdot 0.066 + 146 \cdot 0.033 + 125 \cdot 0.11 = 40.712$. The total number of bets is $\sum_{i=1}^{5} n_i = 149 + 117 + 163 + 146 + 129 = 704$, giving $r = 40.712/704 = 5.78\%$. Table 6 also gives the total number of games, which sums to 13,276, giving bet selection of 5.3%.

13.2.3 Basketball

13.2.3.1 Song et al. (2020)

"Making real-time predictions for NBA basketball games by combining the historical data and bookmaker's betting line" presents a live-betting strategy, where multiple bets are placed during a game. This complicates the calculation of per-bet returns (since multiple bets are placed per game) and particularly the bet selection rate. Across three NBA seasons, 32,886 bets were placed in total. In one sample game where multiple bets were placed (Figure 6, Panel A),

seventeen bets were placed in total. If we assume that this game was chosen due to being representative, then we take seveneteen as the average number of bets placed per bet game. Hence, $32{,}886/17 = 1{,}934.5$ games were bet, in total. A single NBA (regular) season has 1,230 games. This gives bet selection $s = 1{,}334.5/(3 \cdot 1{,}230) = 52.4\%$. Table 1 gives total bet amounts, total profits and the rate of return, which after calculation is linear, with unit stakes. We get that $r = (0.1044 + 0.0859 + 0.134)/1{,}934.5 = 0.017\%$

13.2.3.2 Walsh and Joshi (2023)

In their paper "Machine learning for sports betting: should forecasting models be optimised for accuracy or calibration?", the authors present a machine learning model trained on team and player-data, which is trained by maximising not accuracy but calibration. There is no explicit mention of the total number of bets placed, but inspecting Figure 3 visually, we estimate that they made $n = 1{,}157$ bets. From Table 4, the final bankroll was 13,244.51, having grown from 10,000 with 100-unit bets, a per-unit profit of $(13{,}244.51 - 10{,}000)/100 = 32.4451$. With our estimate of n, this gives $r = 32.4451/1{,}157 = 2.8\%$. These bets were made over a single NBA season, giving a bet selection of $s = 1{,}157/1{,}230 = 94.06\%$.

13.2.4 Baseball

The recent papers relating to sports-betting in baseball are all of one kind: machine learning models with high accuracy scores in classifying game outcomes. These are not betting papers, per se, rather, they present models which can be leveraged by a bettor to create remarkable returns. What is particularly distinct about this class of model is that there is no bet-selection, rather, the models are extremely accurate across all games, meaning that, if bets are placed against the entire collection of matches, profits will be made. These models are defined by a single number, their accuracy, the proportion of game outcomes they correctly classify. To evaluate what a certain accuracy translates to in terms of betting returns, we collect odds for ten MLB games, presented in Table 13.

Date	Home	Away	Home	Home
29-3-24	Milwaukee Brewers	New York Mets	1.95	1.92
29-3-24	Atalanta Braves	Philadelphia Phillies	1.8	2.1
29-3-24	Toronto Blue Jays	Tampa Bay Rays	2.08	1.83
30-3-24	Pittsburgh Pirates	Miami Marlins	2.15	1.8
30-3-24	New York Yankees	Houston Astros	2.04	1.85
30-3-24	Boston Red Sox	Seattle Mariners	2.2	1.73
30-3-24	Cleveland Guardians	Oakland Athletics	1.75	2.2
30-3-24	Colorado Rockies	Arizona Diamondbacks	3	1.45
30-3-24	San Fransisco Giants	San Diego Padres	2.25	1.7
30-3-24	St Louis Cardinals	Los Angeles Dodgers	3.05	1.45

Table 13: Odds for the subsequent ten Major League Baseball games at the time of collection, gathered the 29th of March 2024 (Sweden) from oddschecker.com, with the dates being American. Moneyline market, longest odds available.

Given that bets are placed for every game with these accuracy-based, non-selective models, we take a pessimistic approach, assuming that the model only bets on favourites. In reality, this

is not true: for the model to be more calibrated than bookmaker prices, it will also have to identify misclassified underdogs, which from a betting perspective entails betting on these. But assuming that only favourites are backed, for the ten games listed in Table 13, the mean price for the favourite is 1.728. A model with an accuracy score of Acc would then have expected per-bet returns of $r = Acc \cdot 1.728 - 1$. Given this framework, we present the following baseball papers, in ascending order of accuracy.

13.2.4.1 Cui (2020)

"Forecasting Outcomes of Major League Baseball Games Using Machine Learning" develops a model with $Acc = 61.77\%$, giving $r = 0.0.6177 \cdot 1.728 - 1 = 6.74\%$

13.2.4.2 S.-F. Li et al. (2022)

The best model presented in "Exploring and Selecting Features to Predict the Next Outcomes of MLB Games" had an accuracy score of 65.75%. Hence, $r = 0.6575 \cdot 1.728 - 1 = 13.62\%$.

13.2.4.3 Chen (2014)

"Construction of the Winner Predictive Model in Major League Baseball Games: Use of the Artificial Neural Networks" describes a model with 73.68% accuracy. This model therefore has per-bet returns of $r = 0.7368 \cdot 1.728 - 1 = 27.32\%$

13.2.4.4 Huang and Y.-Z. Li (2021)

Finally, these authors improved on a previous-year paper, by demonstrating 94.18% accuracy in "Use of Machine Learning and Deep Learning to Predict the Outcomes of Major League Baseball Matches". This gives $r = 0.9418 \cdot 1.728 - 1 = 62.7\%$

13.2.5 Tennis

13.2.5.1 Ramirez et al. (2023)

Presenting an ingenious betting strategy, in "Betting on a buzz: Mispricing and inefficiency in online sportsbooks" the authors use data for Wikipedia-page visits for two opposing players to forecast the match outcome. In Table 4, they present two distinct sub-strategies: betting at longest odds and Bet365 (the world's single largest bookmaker) odds, respectively. For the former, the total return was 3.05% over 2,350 bets, giving $r = 0.0305/2{,}350 = 0.0013\%$, while bet selection is $s = 2{,}350/(5{,}188/2) = 90.6\%$ (5,188 is the number of odds, twice the number of games, there being two outcomes in a tennis match). For the Bet365 strategy, the equivalent calculations give $r = 0.0553\%$ and $s = 12.02\%$.

13.2.5.2 Lyócsa and Vỳrost (2018)

In an extremely extensive study "To bet or not to bet: a reality check for tennis betting market efficiency", the authors outline forty different betting strategies and test them against odds from tournaments on the three different competitive tennis surfaces: grass, clay and hard. Tables 2, 3 and 4, respectively, provide s and r for each strategy. We choose the best strategy for each surface: C05 for grass, with $s = 8.26\%$ and $r = 0.7\%$; A07 for clay, with $s = 20.63\%$ and $r = 3.44\%$; and C05, again, for hard, with $s = 5.06\%$ and $r = 5.16\%$.

13.2.6 Cricket

13.2.6.1 Norton et al. (2015)

In their paper "Yes, one-day international cricket 'in-play' trading strategies can be profitable!", the authors describe a vast range of live-betting strategies for cricket. As they restrict themselves to one bet per game (although, they note, they could easily increase this), this makes calculating r and s straight-forward. Considering Tables 7 and 8, we choose the strategy with the lowest p-value, which is the strategy in the ninth row of Table 7. From 73 bets (out of 186 matches, giving $s = 73/186 = 39.2\%$) the total return is 36.3%, which leads to $r = 0.363/73 = 0.497\%$.

13.2.6.2 Bhattacharjee and Talukdar (2020)

The paper "Predicting outcome of matches using pressure index: evidence from Twenty20 cricket" describes another machine-learning strategy. For international twenty-over matches, accuracy was $Acc = 78.4\%$. As above, we collect odds to see what this accuracy converts to in practice. The odds in Table 14 are not from international matches, but from the major twenty-over franchise league, the Indian Premier League.

Date	Home	Away	Home	Away
30-3-24	Lucknow Super Giants	Punjab Kings	1.75	2.08
31-3-24	Gujarat Titans	Sunrisers Hydrabad	1.82	1.99
31-3-24	Delhi Capitals	Chennai Super Kings	2.22	1.66
1-4-24	Mumbai Indians	Rajasthan Royals	1.79	2.01
2-4-24	Royal Challengers Bengaluru	Lucknow Super Giants	1.85	1.94
3-4-24	Delhi Capitals	Kolkata Knight Riders	1.95	1.84
4-4-24	Gujarat Titans	Punjab Kings	1.76	2.05
5-4-24	Sunrisers Hyderabad	Chennai Super Kings	2.04	1.77
6-4-24	Rajasthan Royals	Royal Challengers Bengaluru	1.87	1.92
7-4-24	Mumbai Indians	Delhi Capitals	1.8	1.99

Table 14: Odds for the subsequent ten Indian Premier League games at the time of collection, gathered the 30th of March 2024 from oddschecker.com. Win market, longest odds available.

As argued above, we take the mean odds for favourites, in this case 1.791, and calculate $r = 0.784 \cdot 1.791 - 1 = 0.4041\%$.

13.2.7 Ice Hockey

13.2.7.1 Gu et al. (2019)

In another machine-learning, accuracy-based paper "A game-predicting expert system using big data and machine learning", this model achieves an accuracy score of $Acc = 91.8\%$. With the same methodology, Table 15 shows odds for ten National Hockey League games. With mean odds on favourites of 1.791, this equates to $r = 0.784 \cdot 1.791 - 1 = 40.41\%$.

Date	Home	Away	Home	Away
30-3-24	Buffalo Sabres	New Jersey Devils	2.12	1.84
30-3-24	Florida Panthers	Detroit Red Wings	1.53	2.95
30-3-24	Minnesota Wild	Vegas Golden Knights	2.05	1.96
30-3-24	Edmonton Oilers	Anaheim Ducks	1.26	5
30-3-24	Arizona Coyotes	New York Rangers	2.73	1.54
30-3-24	Colorado Avalanche	Nashville Predators	1.61	2.53
31-3-24	Buffalo Sabres	Toronto Maple Leafs	2.32	1.71
31-3-24	Colombus Blue Jackets	Pittsburgh Penguins	2.53	1.61
31-3-24	Montreal Canadiens	Carolina Hurricanes	3.15	1.49
31-3-24	Philadelphia Flyers	Chicago Blackhawks	1.46	3.1

Table 15: Odds for the subsequent ten National Hockey League games at the time of collection, gathered the 30th of March 2024 from odds-comparison site oddsportal.com. Moneyline market, longest odds available.

13.3 Games and Odds by Sport

13.3.1 American Football

We collect odds data from aussportsbetting.com, which provides odds for the NFL since 2006. We limit ourselves to the last ten years, taking closing odds (prices when the game started) for the 2014-15 season through 2023-24. The game data for the NFL in Table 16 were taken from the same source. The CFL (Canadian Football League) data was taken from the official CFL website (Canadian Football League, 2024) and covers the upcoming 2024 season.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
NFL	29								49	73	59	75
CFL						15	17	16	19	14	5	
Total	29					15	17	16	68	87	64	75

Table 16: American football games per league-month pair.

13.3.2 Baseball

Odds data was collected from sports-statistics.com, taking all regular-season and playoff matches from the 2021 Major League Baseball season. The games data in Table 17 is for the largest baseball leagues across the world, primarily in the Americas North and South as well as Asia. The MLB games data is from the same dataset as the odds (for the 2021 season). For the remaining leagues, we make loose approximations using general information from the Wikipedia page for the most recent season, for instance, for the most recent Cuban National Series season (Wikipedia, 2023). This page lists the duration of the regular season, the duration of the post-season, the number of regular-season games per team, the number of teams, and a full playoff-bracket diagram, listing the number of games played in each series. The regular season ran from March 29th to July 3rd, with 75 games per team and 16 teams. This means that there were $75 \cdot 16/2 = 600$ regular season games (although in this particular case, inspecting the league standings, fewer games were played for certain teams, which we account for). We

assume that these games were uniformly distributed across the span of the regular season. In this case, to simplify the distribution of games per month, we pretend that the regular season ran from April 1st to June 30th. Accounting vaguely for the varying number of days in a month, we distribute these games across the regular-season months. Then we inspect the playoff-bracket, summing up all games played, and add this value to the relevant month (July, in this case). This same general approach was applied for the remaining leagues as well (and for most sports below). Although not as detailed as would ideally be the case, this method manages to capture the vague range of dates when games are played, which, on a monthly basis, is sufficient for our purposes.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Cuban National Series				145	146	145	146	57				
Mexican League				110	171	170	171	171	86			
Nippon Professional Baseball				139	140	139	140	140	139	42		
KBO League				111	112	111	112	112	111	65		
MLB				382	416	396	371	413	401	81		
Total				887	985	961	940	893	737	188		

Table 17: Baseball games per league-month pair.

13.3.3 Basketball

We collect basketball odds data from sportsbookreviewsonline.com, covering the NBA 2018-19 through 2022-23 seasons. The games data in Table 18 for the NBA is from the 2021-22 season from this dataset, there being an issue with the dates attached to the 2022-23 season odds (the odds are intact). For the remaining leagues in the table, the same Wikipedia-page method described above was applied.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Liga ACB	39	35	39	38	39	21				39	38	37
Basketball Super League	33	30	33	32	33	18				36	32	33
VTB United League	24	101	91	25	12					36	35	37
Basketball Bundesliga	39	37	39	38	39	20				40	38	39
National Basketball League	37	6								38	37	36
Chinese Basketball Association	124		94	27						93	225	209
NBA	231	163	229	129	38	6						
Total	527	372	525	289	161	65				282	405	391

Table 18: Basketball games per league-month pair.

13.3.4 Cricket

Cricket, as a sport, is distinct in that the game is played over multiple different time formats: similar to as if soccer games were played not only over 90 minutes per game, but also 30 minutes per game and five hours per game. We choose to collect odds and games data only for the shortest of these formats, twenty-over cricket, the most common in franchise cricket and equal to the others in international play. For the odds, we accessed aussportsbetting.com, taking prices from the Australian T20 league, the Big Bash, from 2011-12 through 2023-24. The data for the franchise leagues in Table 19 are taken familiarly from Wikipedia. Note that The

Hundred is played with a fourth time format (outside of the conventional three, Test, fifty-over and twenty-over cricket), which is only slightly shorter than T20s. For the T20 internationals, we used the official International Cricket Council website (icc-cricket.com, 2023) . In taking all international twenty-over matches played during 2023, removing the under-19s category games.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
SA20	20	10										
Caribbean Premier League								14	20			
Major League Cricket							19					
IPL				37	37							
Pakistan Super League		21	13									
Lanka Premier League								24				
International League T20	14	20										
Big Bash League	22											22
Women's Big Bash League										18	41	
Super Smash (Men+Women)	44											20
Vitality Blast						38	39	39	17			
The Hundred (Men+Women)								136				
European Cricket League			123									
T20I	8	18	32	11	32	39	73	43	41	91	37	49
Total	108	69	168	48	69	77	131	256	78	109	78	91

Table 19: Cricket games per league-month pair.

13.3.5 Tennis

Tennis odds were collected from checkbestodds.com, taking five seasons of the Australian Open, from 2018 through 2022. In tennis, as mentioned above, there is one main tour, the ATP tour, which organises weekly-to-monthly tournaments across the year. They have released a pdf file with the full calendar for the 2024 season, which provides all necessary data for Table 20 (ATP Tour, 2024).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Grand Slam	127				127		127	127				
ATP Masters 1,000			190	150	95			110		150		
ATP 500		93		47		31	68		31	31		
ATP 250	81	27		27	27	54	27	47	27	27	27	
Total	208	120	190	224	249	85	222	284	58	208	27	

Table 20: Tennis games per competition-month pair.

13.3.6 Soccer

For soccer, we use the dataset described in the dataset section. The distribution, from the most recently completed season, 2022-23, of games played across leagues and months is shown in Table 21.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Belgium 1	44	37	28	35			18	35	29	53	18	9
Germany 1	27	36	28	44	36			36	28	44	27	
Germany 2	9	36	29	43	36		18	36	29	43	27	
England 1	35	40	31	62	48			49	18	59	20	18
England 2	39	68	49	83	15		11	70	37	87	44	49
England 3	49	68	58	86	14		12	59	45	79	29	53
England 4	49	71	62	85	13		12	59	50	75	30	45
England 5	52	74	69	80				72	48	73	49	35
France 1	40	50	31	47	42	10		49	32	49	20	10
France 2	40	39	31	49	41	9	9	51	29	41	20	20
Greece 1	35	27	22	34	17			14	21	35	21	14
Italy 1	50	40	30	50	50	10		38	32	50	30	
Italy 2	30	43	37	42	38			30	31	49	30	50
Netherlands 1	44	37	26	37	36			35	28	43	20	
Portugal 1	35	38	27	45	36			36	29	34	18	8
Scotland 1	26	22	15	24	30		6	24	12	33	18	18
Scotland 2	18	19	18	28	5		5	20	10	30	15	12
Scotland 3	18	17	19	27	5		5	20	10	30	15	14
Scotland 4	16	22	20	28	5		5	20	13	22	16	13
Spain 1	38	42	31	57	52	10		30	31	59	20	10
Spain 2	44	44	45	51	47			33	45	65	45	43
Turkey 1	47	17	21	42	29	14		36	27	45	17	18
Total	785	887	727	1079	595	53	101	852	634	1098	549	439

Table 21: Soccer games per league-month pair.

13.4 Assets

13.4.1 Stocks

We chose the S&P500 as our stock index (S&P Dow Jones Indices, 2024), which tracks the performance of the stocks belonging to the 500 largest US-listed companies. Available across the entire period (February 2014 through March 2024), its mean monthly return is 0.0109, with a corresponding volatility of 0.0564.

13.4.2 Bonds

Also compiled by Standard and Poor, we choose the S&P Global Develloped Aggregate Ex-Collateralized Bond Index (S&P Dow Jones Indices, 2024), it being the first index listed on their website of Composite/Global Bonds indices. It tracks how government and corporate investment-grade debt performs, excluding collateralised and securitised bonds. Also available over the entire period, it returned -0.0006 per month, fluctuating with a volatility of 0.0226.

13.4.3 Real Estate

For Real Estate we chose the Dow Jones U.S. Real Estate Index (S&P Dow Jones Indices, 2024), which returned 0.0023 per month with a volatility of 0.0671 over the entire period. This index

follows real estate investment trusts and similar entities which invest solely in real estate in the USA.

13.4.4 Commodities

We chose the S&P GSCI for commodities, which is, per Standard and Poor themselves, a commonly used benchmark (S&P Dow Jones Indices, 2024). Throughout the studied period, it generated a monthly mean return of 0.0069, possessing a volatility of 0.763.

13.4.5 Cryptocurrency

The first index without data fully since 2014, we choose the S&P Cryptocurrency LargeCap Ex-MegaCap Index (S&P Dow Jones Indices, 2024), available since March 2018 and returning 0.0478 per month with a volatility of 0.2858. It excludes the largest individual cryptocurrency assets, but has a broad exposure to a greater variety of such assets.

13.4.6 Art

As outlined by Worthington and Higgs (2004), art has historically performed reasonably well as an alternative investment. This is reflected in how the MAB100, our chosen art index, provided monthly returns of 0.0198 since April 2019, varying with a volatility of 0.0821. Found at myartbroker.com, the index tracks the 100 prints which over the preceding five years accumulated the highest transaction volume.

13.4.7 Wine

As demonstrated by, amongst others, Chu (2014) wine is another well-returning alternative investment. We choose the Cult Wines Global Index, provided at cultx.com and available since 2014. It tracks several regional indices, each region being either a large geographical area or else a notorious wine region. Its performance over the covered period was consistent with a low volatility of 0.111, while returning 0.0046.

13.4.8 Automobiles

Laurs and Renneboog (2019) showed that classic cars can be a complementary asset for investing. While k500.com provides numerous car indices, we chose the K500 Market Average, which unsurprisingly tracks multiple sub-indices, ranging from pre-war European cars to affordable classics to the Japanese domestic market. Its monthly return over the target period was 0.0019, while its volatility was 0.01147.

13.4.9 Rare Coins

Discussed in general as a tradeable asset by Dickie et al. (1994), rare coins are a low-volatility alternative investment, with our chosen index, the Collectors Rare Coin Index, fluctuating only by 0.0096 per month, in terms of volatility, while returning 0.0048. Compiled and collected at rarecoins101.com, it tracks a selection of famous US coins, providing data which includes 2014.

13.4.10 Watches

Another alternative asset, Kylämarkula (2023) outlined the case for investing in luxury watches. We include the ChronoPulse watch index (chrono24.com), which tracks the ten most successful watches from the fourteen most successful brands over the past three years (rolling), having done so since 2019. Over this period, its mean return was 0.0054 per month, with corresponding volatility of 0.0213.

13.4.11 Collectibles

Tracked only since July 2021, the Total Collectibles index (tag CLCTBLS.REGA), gathered at Yahoo Finance, covers various collectible assets, including trading cards and memorabilia, objects which per Burton and Jacobsen (1999) are not uncommonly coveted for financial returns. Since its inception, the index has yielded -0.0181 per month, with a volatility of 0.0447.

13.4.12 Risk-Free Asset

We use the conventional proxy for the risk-free asset, the US-treasury bills, in our case, the one-month maturity class. We access data from fred.stlouisfed.org, from March 2014 on a monthly basis, where the risk-free return was, on average 1.36%.