Predictability in Major League Sports: Betting Odds versus Mathematical Models

April 16, 2026

Abstract

We investigate the efficiency of the sports betting market in predicting game outcomes using data from the seventeen most recent seasons of the four major professional sports leagues in the US, the NFL, NBA, MLB, and NHL. We focus specifically on moneyline bets, a type of bet that amounts to a probabilistic prediction on the outcome of a game and that represents a particularly interesting and rich environment for such a study, from a mathematically point of view as well as from a betting perspective. Using four different metrics for evaluating probabilistic forecasts, we quantify the accuracies of predictions based on the moneyline betting market and compare these accuracies to those of the Bradley-Terry model, a classical mathematical model for probabilistic predictions on game outcomes, as well as a baseline coinflip model that predicts the home team to win with a probability given by the home team team winrate over a given set of past games.

Our main findings are as follows: First, the betting market based model was consistently more accurate than the Bradley-Terry model, which in turn was significantly more accurate than the baseline coinflip model. Second, across all four leagues, the accuracies of betting market based predictions have remained relatively steady over the past several decades. Third, among the four leagues in our study the NBA and NFL had consistently higher accuracy rates than the MLB and NHL and, in this sense, are significantly more "predictable" than the latter two leagues. Finally, an analysis of calibration plots and of returns of investment from specific moneyline based betting strategies did not reveal any major inefficiencies of the moneyline betting market at any parts of the moneyline spectrum. In particular, we found no evidence of a bias for or against strong favorites or longshots in the betting market.

1 Introduction

In Week 15 of the 2025 NFL season, the Los Angeles Chargers were playing against the Kansas City Chiefs. On Monday before the game, the betting site *BetMGM* listed the *moneylines* for the two teams as follows [Bet25]:

Chargers: +165; Chiefs: −200

From a bettor's point of view, these numbers have a simple, and intuitive, interpretation: First, the signs of the moneylines indicate which team is predicted to win (the "favorite") and which team is predicted to lose (the "underdog"): The favorite is the team associated with the *negative* moneyline quote, while the underdog is the team associated with the *positive* moneyline quote[1]. Thus, in the above case the Chiefs are the favorite, while the Chargers are the underdog.

Second, the moneyline numbers themselves indicate how much a bettor will earn if the bet wins. For a bet on the underdog, the moneyline represents the (net) profit a bettor would earn on a $100 wager if the underdog wins. For a bet on the favorite, the absolute value of the moneyline indicates how much a bettor must wager in order to earn a $100 profit. In the above example, a $100 wager on the underdog Chargers will earn the bettor a profit of $165 should the Chargers win (and a loss of the $100 wager should the Chargers

[1]The above example illustrates the most common situation where one of the two moneylines is negative and the other is positive. If the teams are closely matched, it may happen that both moneylines are negative, in which case the team associated with the larger (in absolute terms) moneyline is the favorite; see Section 3.3 and Table 5.5 for further details.

lose). Conversely, a $200 wager on the Chiefs will earn the bettor a profit of $100 with a win by the Chiefs (and a loss of the $200 wager should the Chiefs lose).

From a mathematical point of view, moneylines can be interpreted in terms of win probabilities. In the above example, the moneylines +165 and −200 for the Chargers and Chiefs imply, by a routine calculation (see (3.4) and (3.5) below), the following win probabilities for the two teams:

Chargers: 36.15%; Chiefs: 63.85%

How accurate are such *probabilistic* predictions? How do such predictions compare with those obtained by mathematical models and with "naive" approaches such as always predicting the home team to win? How does the "predictability" of the NFL compare with that of other major professional sports leagues such as the NBA, the MLB, and the NHL?

These are the types of questions we seek to answer in this paper. To that end, we collected game results and moneyline betting odds for regular season games played in the four major professional US sports leagues (NFL, NBA, MLB, NHL) during the period 2009–2025. We computed the accuracies of these predictions using a variety of approaches for evaluating probabilistic forecasts, and we compared these accuracies to those of predictions obtained by a classical mathematical model, the *Bradley-Terry model*, as well as a baseline probabilistic prediction model where the home team is predicted to win with a probability equal to the proportion of home team wins in past games within the given season and league.

Outline of paper. The remainder of this paper is organized as follows. Section 2 provides some background and motivation for this research as well as an overview of some related work in the literature. In Section 3, we describe the probabilistic prediction models we used in our study, while in Section 4, we describe the metrics we applied to evaluate and compare the accuracy of these models. In Section 5, we describe the data that forms the basis of our analysis and provide some descriptive statistics for the collected data. Section 6 contains the results of our study. In Section 7, we summarize our findings and comment on some related questions and directions for further study.

2 Background

Questions such as those mentioned above were considered in a 1997 paper by Stern [Ste97] in which the author compared the accuracy of predictions on game outcomes in major professional and college sports using three types of approaches: (i) the naive approach of always predicting the home team to win; (ii) two mathematical approaches that use a least squares model to predict future game outcomes based on past game results or past game scores; and (iii) oddsmakers' predictions (i.e., predictions by the betting market). Stern's results for professional sports leagues are shown in Table 2.1.

Sport	**Data covered**	**Always predict home team**	**Least squares (win/loss only)**	**Least squares (using scores)**	**Oddsmakers' predictions**
Pro football	NFL 1988–1993	0.58	0.63	0.65	**0.67**
Pro basketball	NBA 1985–1986	0.66	0.69	0.70	**0.71**
Pro baseball	MLB NL 1986	0.53	0.56	**0.56**	0.55

Table 2.1: Proportion of games within the listed data set that have been correctly predicted by different approaches (excerpted from Table 2 of [Ste97]). The highest accuracy rates for each data set are shown in boldface.

Stern's analysis shows that predictions based on betting odds were generally the most accurate among the four approaches analyzed, beating out the other models in all but one case[2], thus supporting the notion

[2]The exception is the MLB where the betting odds model had a slightly lower accuracy rate than the score-based least squares model. This may be due to the small sample size of the MLB data and possibly also to rounding effects.

that the betting market is efficient. Also, not surprisingly, the naive approach of always betting on the home team had the lowest accuracy rate for all leagues analyzed, though still was significantly more accurate than a completely random prediction where the predicted winning team is decided by a coinflip. Finally, among the three professional leagues analyzed by Stern, the NBA had the highest accuracy rates, followed by the NFL and the MLB. In this sense, the NBA is the most "predicable" league, while the MLB is the least predictable.

Our goal in this paper is to provide an analysis similar in spirit to that of Stern, but based on a more recent, and much larger, set of data, and focusing on *probabilistic* instead of *binary* predictions. Such an analysis is particularly relevant and timely given the dramatic explosion of the sports betting market following a 2018 US Supreme Court decision that led to the legalization of sports betting in most US states[3].

We conclude this section with an overview of some related work in the literature. Aside from the work of Stern [Ste97] described above, closest in spirit to the present paper are studies by Boulier and Stekler [BS03] and Song, Boulier and Stekler [SBS07] from 2003 and 2007, respectively, that compared predictions on outcomes of NFL games derived from statistical models with predictions made by human experts such as sports journalists and predictions based on the betting market. The findings from these studies are in line with those from Stern's study in that the betting market outperforms both statistical models and human experts by a significant margin.

Lopez, Matthews, and Baumer [LMB18] used moneyline betting data from the 2006–2016 seasons of the four major professional sports leagues in the US (NFL, NBA, MLB, NHL) to develop models and metrics for the randomness inherent in different sports leagues. They found, in particular, that the MLB and NHL are significantly more random—and hence less predictable—than the NFL and NBA.

One of the earliest large scale studies of the efficiency of the sports betting market is that of Gray and Gray [GG97]. Using data from the 1976–1994 NFL seasons, these authors analyzed specific betting strategies such as the *home-underdog strategy* of betting on the home team team whenever the home team is an underdog. They found evidence that there exist strategies (e.g., the home-underdog strategy) that outperform a random betting strategy by a statistically significant margin. However, the margin was generally not sufficient to offset the bookmakers' profit. Moreover, any inefficiencies in the betting market tend to dissipate over time. Similar conclusions for the soccer betting market were obtained in a recent paper by Winkelman et al. [WODM24], which analyzed the 2006–2019 seasons of five major European soccer leagues.

A particular type of betting strategy that has received considerable attention in the literature is betting on an extreme favorite or an extreme underdog. In terms of moneylines, this amounts to placing a bet whenever a team has a very large negative moneyline or a very large positive moneyline. The *Favorite-Longshot (FL) Bias* states that longshots (i.e., extreme underdogs) tend to be overbet by bettors, thus resulting in moneylines for underdogs that are smaller than they should be, while the *Reverse Favorite-Longshot (RFL) Bias* states that longshots tend to be underbet. FL type biases have been documented in race track betting [Sau98] and some other sports such as professional tennis [FM07]. By contrast, studies by Woodland and Woodland [WW94, WW01] observed an RFL type bias in the MLB and NHL, though follow-up studies [WW03, WW11] showed that this bias has weakened or disappeared. For a recent survey on this topic, see Newall and Cortis [NC21]

3 Prediction Models

3.1 Coinflip model with home bias

Our first model is one that predicts the home and away teams to win with probabilities p and $1 - p$, respectively, where p is the proportion of past games won by the home team within the given season and league. This model represents a natural analog among probabilistic prediction models of the "always predict home team" model in Table 2.1, and it represents a simple baseline model against which one can compare more sophisticated models.

[3] The amount of money wagered on legal sports betting in the US increased from \$6.6 billion in 2018 to over \$165 billion in 2025 (cf. [Bis26]).

3.2 Bradley-Terry model

The Bradley-Terry model is a classic model for making probabilistic predictions on outcomes of paired comparisons. Originally conceived more than 70 years ago by Bradley and Terry [BT52], the model has stood the test of time and remains one of the most widely used, and well regarded, mathematical models of its kind. In the context of predictions on games between sports teams, the basic idea of the model is that the probability that team i beats team j is of the form

$$P(i \text{ beats } j) = \frac{\pi_i}{\pi_i + \pi_j}. \tag{3.1}$$

The numbers π_i and π_j here are positive real numbers representing the relative strengths of the teams involved and are calculated from a given set of past game results using a maximum likelihood estimate.

Our implementation of the Bradley-Terry model is based on the algorithm described in [Cod19], which applies an iterative procedure to compute the strength parameters π_i from a given set of game results. As input to this computation we use the results of all games within the given season that have occurred prior to the game to be predicted.

3.3 Betting odds

The most common types of bets on games in US professional sports leagues are bets against the spread, over/under bets, and moneyline betting ("odds betting").

The first two of these bets are examples of *binary* bets: The bettor bets on whether or not a team beats another team by (at least) a specified threshold ("spread"), or whether or not the sum of the two scores exceeds a specified threshold. The thresholds are set by the bookmaker in order to approximately equal the amount of betting done on either side of the threshold and thus represent ways to generate a "fair" binary betting proposition from a game between unevenly matched teams.

By contrast, moneylines correlate, in a natural way, with predicted win probabilities, and a moneyline bet can thus be interpreted as a *probabilistic* prediction on the game outcome. The conversion between moneylines and win probabilities is as follows:

First, given a moneyline m, define an associated "raw" probability $p^* = p^*(m)$ by

$$p^* = p^*(m) = \begin{cases} \dfrac{100}{100+m} & \text{if } m > 0, \\ \dfrac{|m|}{100+|m|} & \text{if } m < 0. \end{cases} \tag{3.2}$$

Now consider two teams, A and B, with moneylines m_A and m_B, respectively, and let p_A^* and p_B^* be the associated raw probabilities computed via (3.2). The sum of these raw probabilities will (generally) be greater than 1, with the difference representing the bookmaker's profit. To account for this, we normalize these probabilities by letting

$$p_A = \frac{p_A^*}{p_A^* + p_B^*}, \quad p_B = \frac{p_B^*}{p_A^* + p_B^*}. \tag{3.3}$$

The normalized probabilities p_A and p_B defined by (3.3) sum to 1 and have a natural interpretation as the probabilities for each of the two teams to win. In the example mentioned in the Introduction with moneylines $+165$ and -200 for the Chargers and Chiefs, respectively, the calculations are as follows:

$$p^*_{\text{Chargers}} = \frac{100}{100+165} \approx 0.37735, \quad p^*_{\text{Chiefs}} = \frac{200}{100+200} \approx 0.66666, \tag{3.4}$$

$$p_{\text{Chargers}} \approx \frac{0.37735}{0.37735+0.66666} \approx 0.3615, \quad p_{\text{Chiefs}} \approx \frac{0.66666}{0.37735+0.66666} \approx 0.6385. \tag{3.5}$$

Thus, the pair of moneylines $(+165, -200)$ translates into winning probabilities of approximately 36.15% for the Chargers and 63.85% for the Chiefs.

Moneyline bets of the form described here, and which we are focusing on exclusively in this paper, are so-called two-way moneyline bets, i.e., bets on events with two outcomes (e.g., "Team A wins" and "Team B

wins") and two associated moneylines[4]. These types of bets are the standard moneyline bets for three of our leagues, the NFL, NBA, and MLB. For NHL games, another popular type of moneyline bet is a three-way bet ("1X2 bet") that distinguishes between wins by either team in regulation time and the game going into overtime, with an associated a moneyline for each of these three outcomes. Our analysis is based on the traditional two-way bet that does not distinguish between wins in regulation time and wins in overtime.

4 Evaluation Metrics

For *deterministic* predictions on binary events (such as whether a teams wins or loses a game, or whether the score difference falls above or below a given threshold) there exists a simple and natural measure for the quality of the prediction, namely the proportion of correctly predicted outcomes, or *accuracy* of the prediction. By contrast, evaluating and comparing *probabilistic* forecasts on game outcomes—the main focus of this paper—is far from obvious, and no single metric exists that fully captures the quality of the prediction. For our analysis we will use four methods that measure the quality of a probabilistic prediction in very different ways and which are described in more detail in the following sections: Accuracy rate, Brier score, calibration plots, and ROI.

4.1 Accuracy rate

The accuracy rate of a prediction method is the fraction of "correct" predictions among all predictions made. In the case of deterministic prediction methods on binary events (i.e., methods that only predict whether or not an event occurs), this is the natural metric for evaluating a particular prediction method. Accuracy rates can also be defined for probabilistic predictions by interpreting a "correct" prediction to be one in which the occurrence of the event was predicted with probability greater than 50%. Obviously, this entails a significant loss of information as there is no difference between a 51% and a 99% win probability in terms of the resulting accuracy rate. Despite this drawback, accuracy rates can serve as simple and easy-to-understand baseline metric for probabilistic predictions, and they are also useful if one seeks to compare probabilistic predictions with deterministic predictions on binary events.

4.2 Brier score

The Brier score is a classical metric for evaluating probabilistic forecasts. Originally conceived more than 75 years ago by Glenn Brier [Bri50] to evaluate probabilistic weather forecasts, it has since been used in many other contexts. In particular, the "NFL forecasting game", a probabilistic prediction contest on NFL games organized by *FiveThirtyEight* [Fiv21], used the Brier score as metric to rank participants in this contest.

The Brier score essentially represents the mean square error in the probabilistic forecast. It is formally defined as follows: Given a set of binary events $i = 1, \dots, n$ with outcomes $o_i \in \{0, 1\}$, and a corresponding set of probabilistic predictions on these outcomes, $p_i = P(o_i = 1)$, the Brier score is given by

$$B = \frac{1}{n} \sum_{i=1}^{n} (p_i - o_i)^2. \tag{4.1}$$

It is clear from the definition (4.1) that the Brier score B is a real number in the interval $[0, 1]$, and that $B = 0$ if and only if $p_i = o_i$ for all i, i.e., if and only if the forecast probability, p_i, is 1 if the event i occurs, and 0 if the event i does not occur. In the context of wins and losses (encoded as $o_i = 1$ and $o_i = 0$, respectively), a Brier score of 0 means that the prediction probabilities for wins are either 100% or 0%, and that these probabilities correspond exactly to the actual outcomes. Thus, a prediction with a Brier score of 0 can be viewed as a "perfect" probabilistic forecast.

On the other hand, a probabilistic forecast in which all prediction probabilities are 50%, i.e., where $p_i = 0.5$ and therefore $p_i - o_i \in \{\pm 0.5\}$ for all i, results in a Brier score $B = 0.5^2 = 0.25$. The latter value, 0.25, can be interpreted as the Brier score of a "trivial", or no-skill, prediction, and it represents a natural baseline against which one can compare a given probabilistic prediction.

[4] We ignore the possibility of tied games, which are exceedingly rare in the major US sports leagues that form the subject of this paper.

4.3 Calibration plots

Another way to interpret a "perfect" probabilistic forecast is as a forecast for which the predicted probabilities of events coincide with the actual frequencies with which the events occur. Thus, in an ideal scenario, among all instances in which the predicted probability of an event was 20%, the event should occur in exactly 20% of these instances.

A natural way to visualize how close a given probabilistic forecasting method comes to this ideal is by graphing the actual frequencies of occurrence against the predicted probabilities. A perfect probabilistic forecast would be one in which this graph coincides with the diagonal line $y = x$. A graph of this type is called a *calibration plot*.

In practice, as the number of data points available is finite, this requires grouping predicted probabilities into bins, and comparing predicted probabilities and actual frequencies binwise. One common practice is to split the interval $[0, 1]$ representing the predicted probabilities into 10 subintervals of equal length, $I_1 = [0, 0, 1)$, $I_2 = [0.1, 0.2)$, ..., $I_{10} = [0.9, 1]$. If f_i denotes the proportion of events among those whose predicted probabilities fall into I_i that actually occur, then we can plot f_i, the actual frequency of occurrence among events that fall into I_i (and hence have predicted win probabilities between $(i-1)/10$ and $i/10$), against $(i-0.5)/10$, the midpoint of the interval I_i. The discrete points $((i-0.5)/10, f_i)$, $i = 1, \ldots, 10$, represent a discrete version of the calibration plot described above.

In addition to providing an intuitive, and easy-to-understand, visual representation of the quality of a probabilistic forecast, a calibration plot can reveal inefficiencies in such a forecast, such as overpredicting or underpredicting events at particular locations of the predicted probability spectrum.

4.4 ROI

Our final method of evaluating probabilistic predictions, and specifically those based on moneyline odds, is the Return on Investment (ROI). We again split the set of all events (i.e., games) into bins I_i, $i = 1, \ldots, 10$, based on the predicted probabilities implied by the moneylines, as described above. For each of these bins we calculate the ROI under two types of betting strategies:

- **Underdog betting on I_i:** Only bet if the event falls into bin I_i, and bet the underdog in this case.
- **Favorite betting on I_i:** Only bet if the event falls into bin I_i, and bet the favorite in this case.

This analysis can reveal subtle inefficiencies in the betting market that may be exploited by appropriate betting strategies.

5 Description of Data

5.1 Data summary

Our analysis is based on game results and betting odds data for the four major professional sports leagues in the US during the period 2009–2025. The data was obtained from the OddsPortal site, `oddsportal.com`, a website that aggregates and archives sports betting data from major bookmakers. Table 5.1 shows a summary of the data we used for our work.

League	Seasons	Total # of Games	# of Second-Half Games
MLB	2009–2025	38805	19301
NBA	2009–2025	19512	9695
NFL	2009–2025	4360	2124
NHL	2009–2025	19164	9478

Table 5.1: Summary of data used in our analysis. The second half of each season was used to evaluate the different prediction models, while game results from the entire season up until the game to be predicted were used to tune the parameters in the home bias coinflip and Bradley-Terry models.

Only regular season games were considered. The second column in Table 5.1 shows the seasons covered in the analysis. For leagues whose regular seasons span two years, seasons were indexed by the year in which they *ended*; for example, for the NBA the season range 2009–2025 covers all NBA seasons from 2008/2009 through 2024/2025.

Each season was split into two half-seasons, with the first half-season consisting of all games that occurred at or before a certain cutoff date D, and the second half-season consisting of the games that occurred after date D. The cutoff date D was chosen as the earliest date such that the total number of first-half season games is greater or equal to the total number of second-half season games. This choice ensures that the two half-seasons consist of *approximately* the same number of games[5].

The second-half season games were used to evaluate the various prediction models using the metrics described in Section 4. The home win proportion p for the home bias coinflip model and the team strength parameters π_i in the Bradley-Terry model were calculated from the results of first-half season games as well as the games in the second half of the season that have been played before the date of the game to be predicted.

Excluded data. Games for which moneyline data was not available at `oddsportal.com` were excluded from our analysis and from the game counts shown in Table 5.1. Also excluded were games that were played on a neutral location and games that ended in a tie. As Table 5.2 shows, the number of games affected by the various exclusions is very small, leaving at least 94% valid games for each of our leagues.

League	Missing moneyline data	Other reasons
MLB	850 (2.140%)	60 (0.151%)
NBA	789 (3.885%)	10 (0.049%)
NFL	3 (0.068%)	40 (0.908%)
NHL	1129 (5.562%)	8 (0.039%)

Table 5.2: Numbers and percentages of games excluded from our analysis for each league.

5.2 Calculation of moneyline based win probabilities

As described in Section 3.3, moneyline betting odds can be converted to win probabilities for the two teams involved. For most games, our data source, `oddsportal.com`, provided moneyline quotes from multiple bookmakers. For such games, we calculated a "consensus" pair of moneylines as follows: Given a set of individual moneyline pairs $(m_{A,i}, m_{B,i})$, $i = 1, \dots, k$, provided by different bookmakers, we first converted

[5]Since multiple games may be played simultaneously or on the same date, the numbers of first-half and second-half season games are not necessarily equal or differ by 1. However, for all of our leagues they are within one or two percentage points of each other.

each of these pairs into a corresponding pair $(p^*_{A,i}, p^*_{B,i})$ of *raw* (i.e., non-normalized) win probabilities via equation (3.2). We let $p^*_A = (1/k)\sum_{i=1}^k p^*_{A,i}$ and $p^*_B = (1/k)\sum_{i=1}^k p^*_{B,i}$ denote the averages of these probabilities, and converted the pair (p^*_A, p^*_B) back to a pair of moneylines $(m_A, m_B) = (m(p^*_A), m(p^*_B))$ by letting[6]

$$m = m(p^*) = \begin{cases} \langle 100 \cdot (1-p^*)/p^* \rangle & \text{if } p^* \leq 1/2, \\ \\ \langle 100 \cdot p^*/(1-p^*) \rangle & \text{if } p^* > 1/2, \end{cases} \tag{5.1}$$

where $\langle x \rangle$ denotes the value of x rounded to the nearest integer. The pair (m_A, m_B) obtained in this manner is a pair of integers in $(-\infty, -100] \cup (100, \infty)$ that has a natural interpretation as a "consensus" moneyline pair for the given game. Using this pair (m_A, m_B), we calculated an associated pair of win probabilities (p_A, p_B) via equations (3.2) and (3.3).

5.3 Some descriptive statistics

Home winrates. Table 5.3 shows, for each of the four leagues, the proportion of second-half season games within our data set that were won by the home team. As the table shows, the NBA enjoys the largest home advantage, followed by the NFL, the NHL, and the MLB. It is interesting to compare the data in this table (which covers the period 2009–2025) to the data compiled by Stern in Table 2.1 (which is based on the period 1988–1993). The home winrates for the NBA and NFL both have decreased significantly, though remain slightly higher than that of the MLB, which has remained relatively constant at 53%–54%. Possible explanations for this decrease include lower travel costs for fans and increased levels of parity among teams in a league.

League	**Home Winrate**
MLB	53.671%
NBA	57.060%
NFL	55.132%
NHL	54.146%

Table 5.3: Percentage of second-half season games won by the home team.

Bookmaker profits. Consider a game G between Teams A and B with an associated pair of moneylines $m(G) = (m_A, m_B)$. Let p^*_A and p^*_B denote the implied "raw" win probabilities for the two teams, calculated from m_A and m_B via equation (3.2)). The sum of these raw win probabilities will (generally) be greater than 1, with the excess, $P(G) = p^*_A + p^*_B - 1$, representing the (expected) bookmaker's profit ("vig") on game G.

Table 5.4 shows the average bookmaker profit for each league, i.e., the average value of the quantity $P(G)$, taken over all games G in the given league within our data set. The table shows that bookmaker profits are fairly consistent across all four leagues, averaging around $5\% - 6\%$.

[6] It is easy to check that the function $p^*(m)$ defined in (3.2) is a bijection from the set $(-\infty, -100] \cup (100, \infty)$ to the set $(0, 1)$, and that the function $m(p^*)$ defined in (5.1) represents the inverse of $p^*(m)$, rounded to the nearest integer.

League	Bookmaker Profit
MLB	5.698%
NBA	5.093%
NFL	5.456%
NHL	5.221%

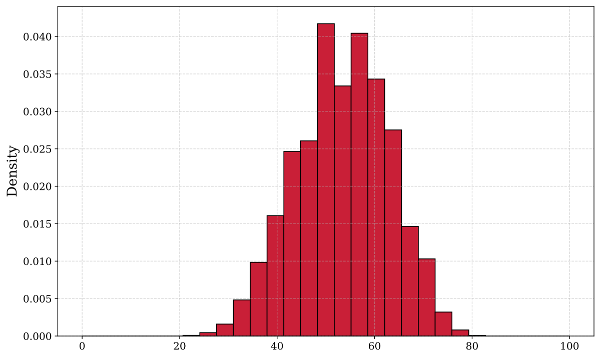
Table 5.4: Average bookmaker profit by league.

Games with two negative moneylines. Some interesting insight into the predictability of different leagues can be gleaned by examining the moneyline data. Due to the bookmakers' profit margins, if two teams are closely matched, it may happen that both moneylines are negative (so that both sides represent worse than even odds for a bettor). Table 5.5 shows how frequently this situation occurs, for each of our four leagues. The table reveals a striking difference between the MLB and NHL on the one hand, and the NFL and NBA on the other hand. For the former two leagues, two negative moneylines are around three times more frequent as for the latter two leagues. This indicates that there is a much greater level of uncertainty in predicting MLB and NHL games than in predicting NFL and NBA games.

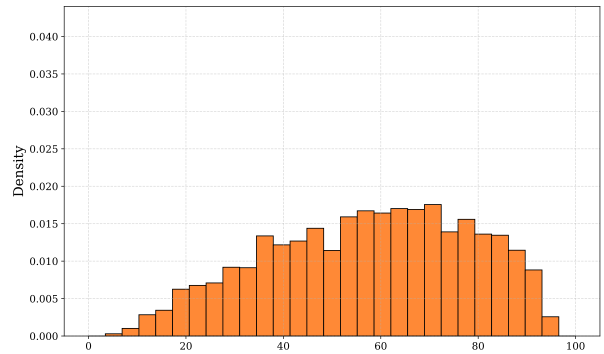
League	Number of Games	Proportion of Games
MLB	4436	22.983%
NBA	651	6.715%
NFL	148	6.968%
NHL	1778	18.759%

Table 5.5: Games with two negative moneylines

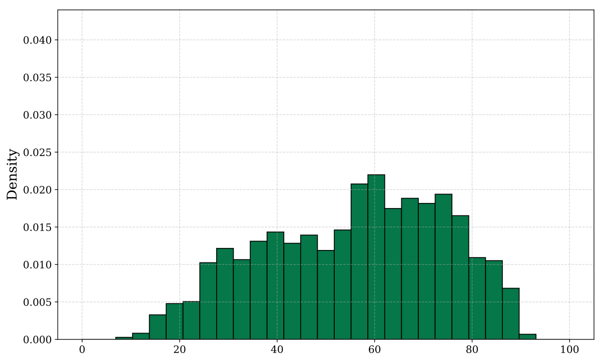
Distribution of win probabilities. Figure 5.1 shows, for each of the four leagues, the distribution of win probabilities for the home team implied by moneylines. All four of these distributions are slightly shifted to the right, reflecting the higher winrates by the home team observed in Table 5.3. More interestingly, there are some striking differences in the *shapes* of these distributions: For the MLB and NHL, the win probability distributions are mostly concentrated near the middle of the range, while for the NFL and NBA these probabilities occupy a much wider range within the full probability spectrum. This indicates that there is a much greater inherent uncertainty in predicting MLB and NHL games than in predicting NFL and NBA teams. In other words, among the four leagues, the MLB and NHL are much less predictable than the NFL and NBA. This is consistent with findings from earlier studies such as those of Stern [Ste97] (cf. Table 5.5) and Lopez, Matthews, and Baumer [LMB18].



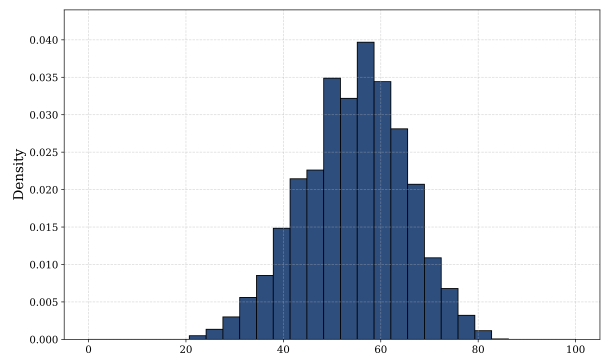
(a) MLB



(b) NBA



(c) NFL



(d) NHL

Figure 5.1: Distribution of home team win probabilities implied by moneylines.

6 Results

6.1 Accuracy rates

Recall that the accuracy rate of a probabilistic prediction model is the proportion of games whose outcomes were correctly predicted by this model, in the sense that the winner of the game was the team with a predicted win probability strictly greater than 50%. If both teams have a predicted win probability of exactly 50%, the probabilistic prediction does not identify a winner. We therefore excluded those cases when calculating accuracy rates; the number of games affected by this exclusion, shown in Table 6.1, is very small and does not materially affect the overall conclusions.

League	Home Bias Coinflip	Bradley-Terry	Moneyline
MLB	0 (0.000%)	0 (0.000%)	321 (1.663%)
NBA	0 (0.000%)	2 (0.021%)	18 (0.186%)
NFL	11 (0.518%)	0 (0.000%)	3 (0.141%)
NHL	0 (0.000%)	1 (0.011%)	196 (2.068%)

Table 6.1: Number of games in which both teams had a predicted a win probabilities of exactly 50% for one of our models. These games were excluded from the accuracy rate computations for the corresponding model.

Table 6.2 shows the accuracy rates of the predictions derived from the three probabilistic prediction models described in Section 3, after removing the games listed in Table 6.1. Figure 6.1 provides a visual representation of the same data. The table and figure show that, for each of the four leagues, moneyline based predictions have the highest accuracy rates, followed by predictions based on the Bradley-Terry model, with the home bias coinflip model coming in third among the three (nontrivial) prediction models, though still clearly outperforming a pure (i.e., non-biased) coinflip model. Moreover, within the Bradley-Terry and moneyline models the MLB and NHL had significantly lower accuracy rates than the NBA and NFL.

League	Home Bias Coinflip	Bradley-Terry	Moneyline
MLB	53.671%	56.557%	**58.625%**
NBA	57.060%	65.222%	**69.412%**
NFL	53.999%	63.183%	**69.496%**
NHL	54.146%	57.518%	**59.966%**

Table 6.2: Accuracy rates of binary predictions derived from different probabilistic prediction models. The highest accuracy rates for each league are shown in boldface.

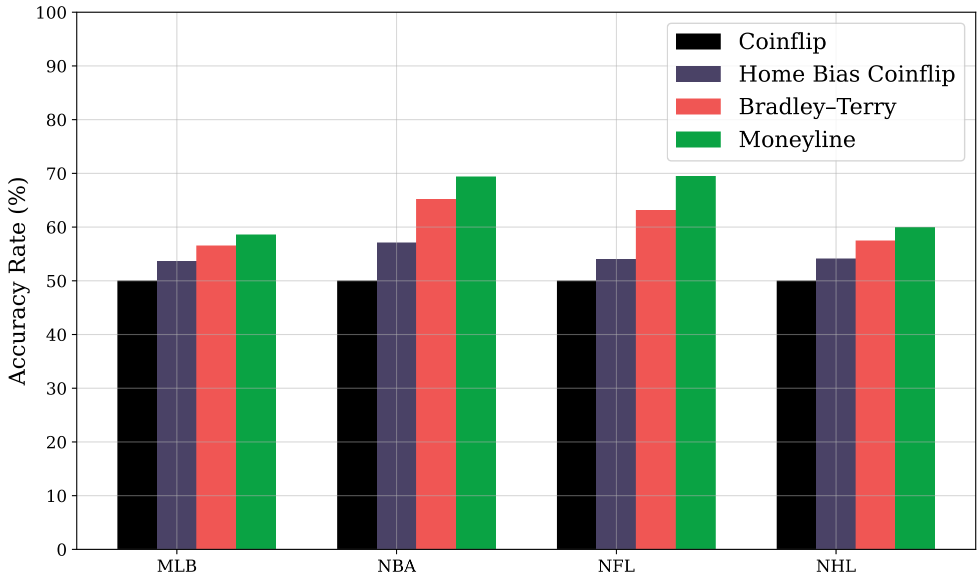
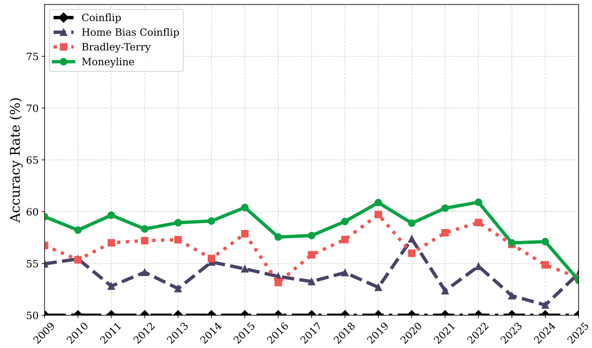
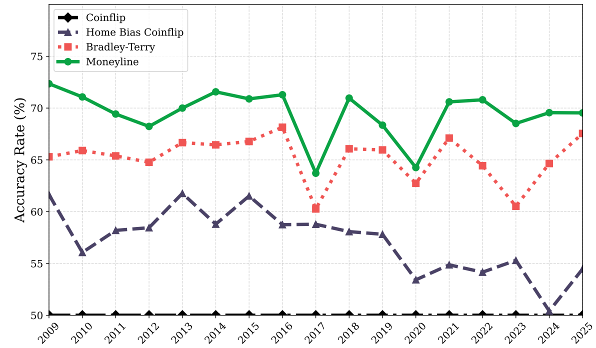


Figure 6.1: Accuracy rates by league and prediction model.

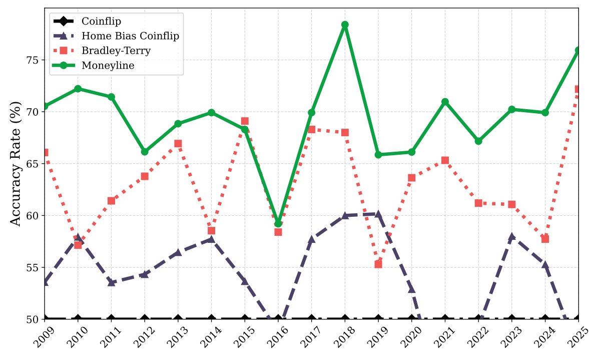
For further insight, we broke down accuracy rates by season. The results, shown in Figure 6.2, reveal that these rates can vary significantly from season to season. The largest seasonal variations can be observed in the NFL, and the smallest variations in the MLB. This may be explained by the fact that an NFL season is only 16–17 games long, compared to 162 games for a typical MLB season, resulting in a much smaller sample size of games to be predicted.



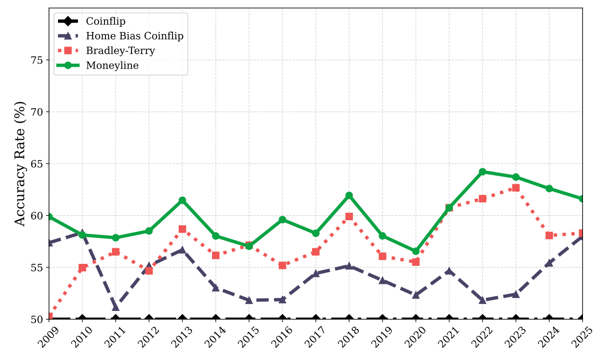
(a) MLB



(b) NBA



(c) NFL



(d) NHL

Figure 6.2: Accuracy rates by league, prediction model, and season.

It is interesting to compare the accuracy rates in our study with those from earlier studies such as the data compiled by Stern [Ste97], shown in Table 2.1. In terms of relative accuracies of different prediction models and across different leagues, our results are largely consistent with the earlier results: Among the three prediction models used, betting market based models had the highest accuracy rates, followed by mathematical models, with naive models based only on home advantage coming in last. This is, of course, not surprising, given the sophistication of the various models and the amount of information that the model predictions are based on. Among different leagues, the NBA and NFL had significantly higher accuracy rates than the MLB.

It is remarkable how close the accuracy rates of moneyline based predictions in our study are to those based on oddsmakers' predictions in Stern's study, even though the time periods covered in the two studies are several decades apart (2009–2025 versus 1985–1993) and the betting market is vastly different now than it was at the time of Stern's study. Indeed, the differences in accuracy rates between these two time periods are comparable to the seasonal variations shown in Figure 6.2, suggesting that these differences are due to random effects rather than indicative of any real improvement in prediction accuracies. This may be surprising given the increased sophistication of both bookmakers in setting moneylines and professional bettors in trying to exploit inefficiencies in moneylines set by bookmakers. A possible explanation might be that any improvements in prediction accuracies due to the increased sophistication of the betting market has been offset by a greater level of parity within the major sports leagues.

6.2 Brier scores

Table 6.3 and Figure 6.3 show that moneyline based predictions have the lowest Brier scores across all four leagues, followed by predictions based on the Bradley-Terry model, with the home bias coinflip model coming in last[7]. Moreover, among the different leagues, the NFL and NBA had significantly lower moneyline based Brier scores than the MLB and NHL, indicating that the former two leagues are significantly more predictable than the latter two leagues. These findings are in line with those based on accuracy rates (cf. Table 6.2) and with earlier cross-league studies such as those of Stern [Ste97] and Lopez, Matthews, and Baumer [LMB18].

League	Home Bias Coinflip	Bradley-Terry	Moneyline
MLB	0.249	0.245	**0.239**
NBA	0.245	0.218	**0.199**
NFL	0.249	0.245	**0.203**
NHL	0.248	0.245	**0.236**

Table 6.3: Brier scores by league and prediction model. The lowest Brier scores (corresponding to the most accurate predictions) for each league are shown in boldface.

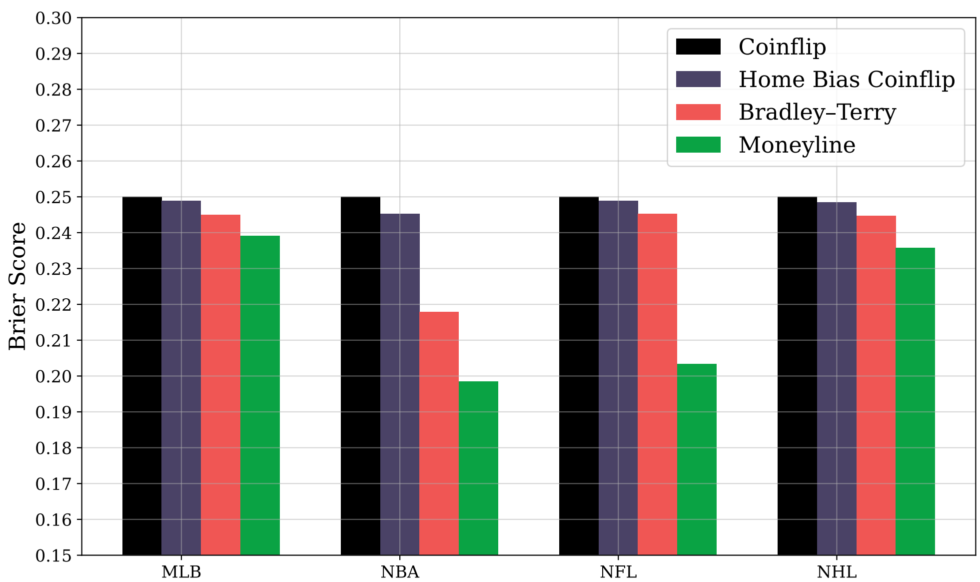
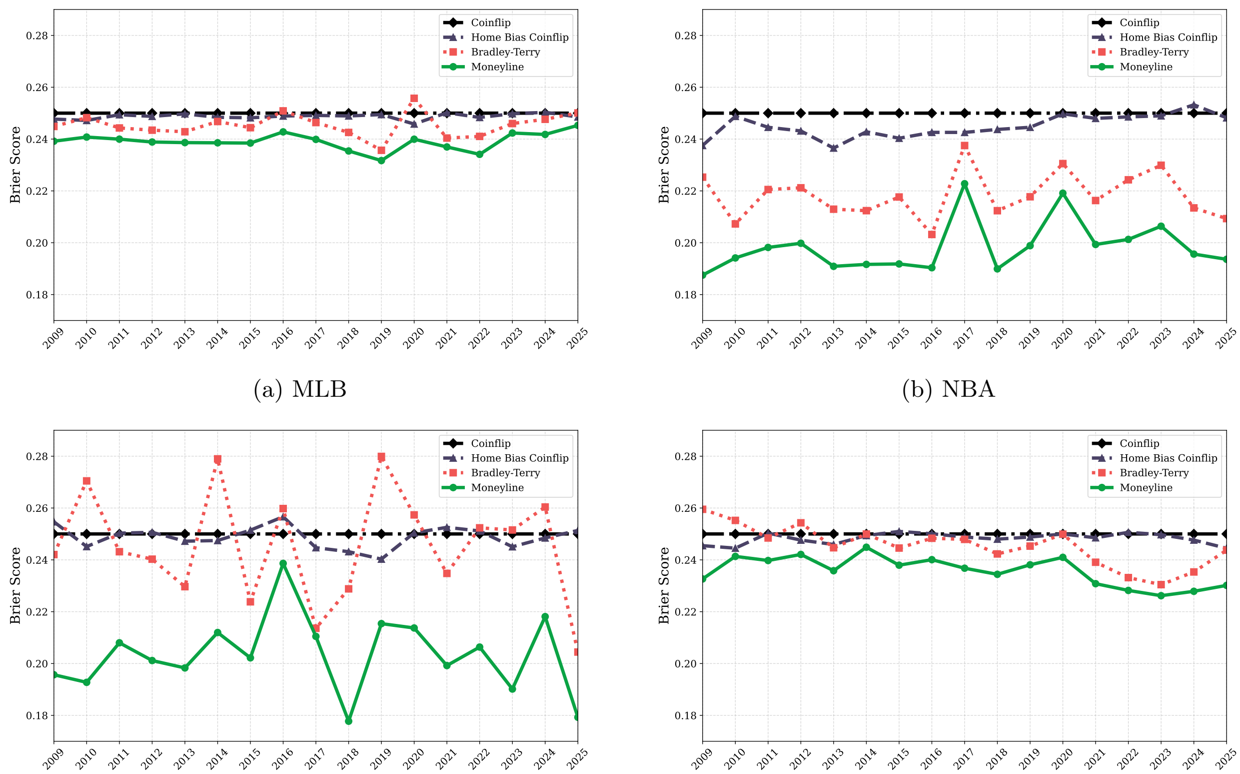


Figure 6.3: Brier scores by league and prediction model.

Figures 6.4 and 6.5 show the breakdown of the Brier scores from Table 6.3 by season. While there are substantial seasonal variations in the Brier scores, for most seasons the NFL and NBA had significantly lower Brier scores than the MLB and NHL. The difference, which is particularly noticeable in Figure 6.5, reflects the differences in predictability between these observed leagues.

[7]Recall (cf. Section 4.2) that smaller Brier scores correspond to more accurate predictions, and that a pure coinflip model (i.e., one in which each team is predicted to win with probability 0.5) has a Brier score of 0.25.



(a) MLB (b) NBA

(c) NFL (d) NHL

Figure 6.4: Brier scores by league, prediction model, and season.

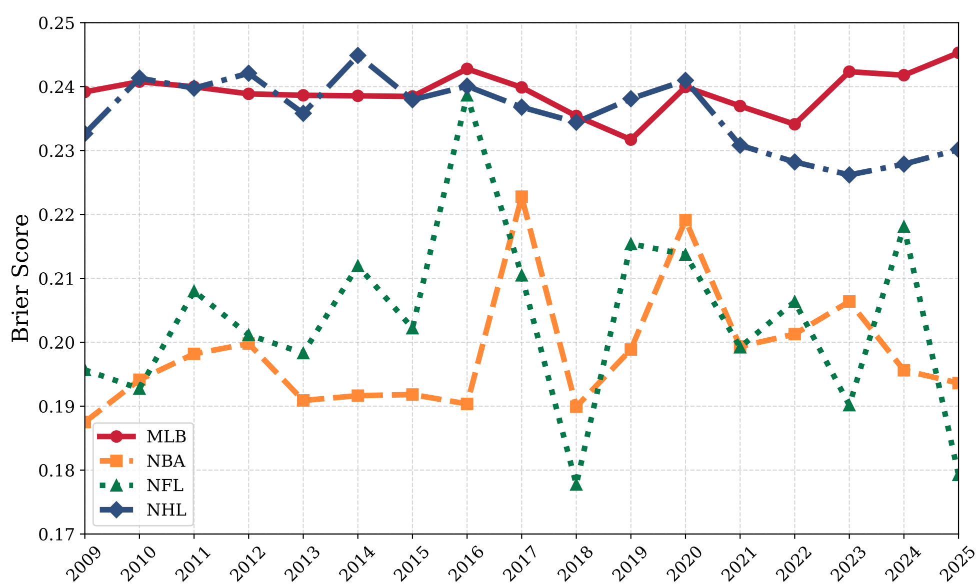
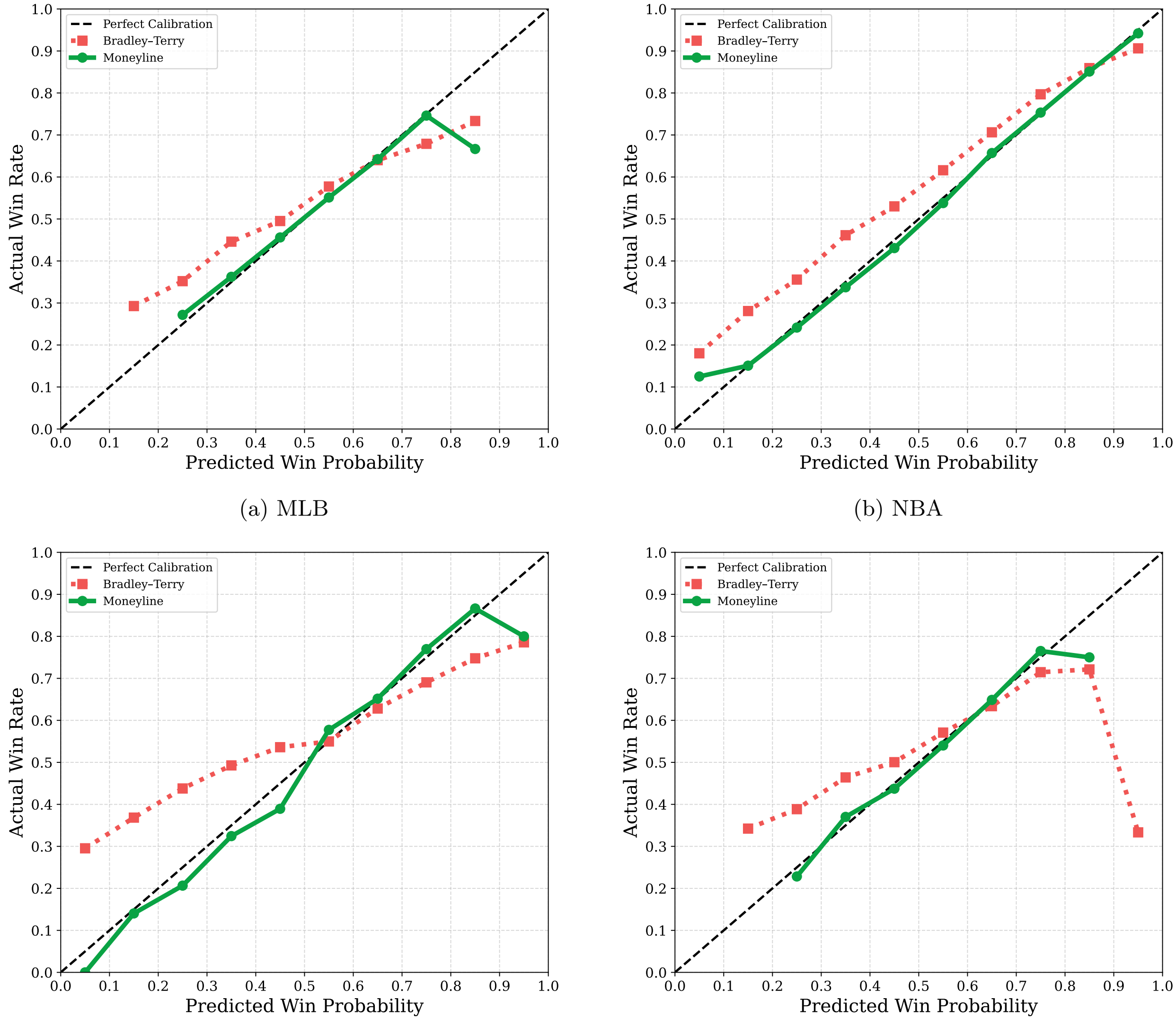


Figure 6.5: Brier scores of moneyline based predictions by league and season.

6.3 Calibration plots

Figure 6.6 shows calibration plots for moneyline and Bradley-Terry based predictions for each of the four leagues. The figure shows that the plots based on the moneyline model are very close to the diagonal line (representing a perfectly calibrated prediction model), indicating that the moneyline betting market is highly efficient.



(a) MLB

(b) NBA

(c) NFL

(d) NHL

Figure 6.6: Calibration plots for moneyline and Bradley-Terry based predictions.

6.4 ROI

Tables 6.4–6.7 show the ROI for each of the betting strategies described in Section 4.4: that is, bet on either the favorite or the underdog if the win probability implied by the moneylines falls into a given interval[8]. Nearly all ROI numbers are negative, meaning that the corresponding betting strategy is unfavorable to the bettor. The few instances in which there was a positive ROI are indicated in boldface. All of these are

[8]Except for the final interval, 90%–100%, which includes both endpoints, the probability intervals are to be interpreted as including the left endpoint and excluding the right endpoint. A few of the bins contained no data points so that ROI numbers could not be calculated; in those cases the corresponding entries in the ROI columns were left blank.

associated with relatively small sample sizes and thus do not yield a viable betting strategy[9]. In particular, there is no evidence of either a favorite-longshot bias or a reverse favorite-longshot bias in the moneyline data. That is, bets based on the extreme ends of the probability spectrum (i.e., bets on strong favorites or on heavy underdogs) do not have significantly better ROIs than bets based on other parts of the probability spectrum.

Probability Bin	**Bin Count**	**ROI (Bet on Favorite)**	**ROI (Bet on Underdog)**
0%–10%	0	–	–
10%–20%	0	–	–
20%–30%	92	-4.179%	-5.000%
30%–40%	1,565	-4.797%	-5.796%
40%–50%	5,397	-6.062%	-4.981%
50%–60%	7,234	-4.742%	-6.409%
60%–70%	4,360	-4.978%	-5.868%
70%–80%	650	-2.498%	-11.697%
80%–90%	3	-22.331%	**66.667%**
90%–100%	0	–	–

Table 6.4: **MLB:** ROI for favorite and underdog bets based on moneyline implied win probabilities.

Probability Bin	**Bin Count**	**ROI (Bet on Favorite)**	**ROI (Bet on Underdog)**
0%–10%	40	-8.701%	**59.400%**
10%–20%	378	-3.642%	-10.336%
20%–30%	733	-3.176%	-9.895%
30%–40%	1,058	-2.329%	-8.832%
40%–50%	1,295	-1.869%	-9.063%
50%–60%	1,475	-7.356%	-1.778%
60%–70%	1,621	-3.867%	-7.292%
70%–80%	1,508	-4.031%	-6.366%
80%–90%	1,242	-4.237%	-5.983%
90%–100%	345	-2.245%	-32.197%

Table 6.5: **NBA:** ROI for favorite and underdog bets based on moneyline implied win probabilities.

[9]All but three of the cases with a positive ROI involve sample sizes in the single or double digits. The three exceptions are the 20%–30% and 40%–50% bins in the NFL data and the 20%–30% bin in the NHL data, which contained 184, 288, and 127 data points, respectively. Keeping in mind that our data covers a 17 year period (2009–2025), this amounts to an average of between 8 and 17 data points per season, which seems far too small of a sample size to draw any conclusions from the positive ROI values.

Probability Bin	Bin Count	ROI (Bet on Favorite)	ROI (Bet on Underdog)
0%–10%	1	**5.995**%	-100.000%
10%–20%	57	-2.211%	-17.632%
20%–30%	184	**2.056**%	-23.745%
30%–40%	262	-1.238%	-13.939%
40%–50%	288	**4.905%**	-18.514%
50%–60%	359	-1.211%	-8.993%
60%–70%	399	-4.275%	-5.987%
70%–80%	382	-2.598%	-14.123%
80%–90%	187	-2.749%	-21.086%
90%–100%	5	-16.054%	**110.600%**

Table 6.6: **NFL:** ROI for favorite and underdog bets based on moneyline implied win probabilities.

Probability Bin	Bin Count	ROI (Bet on Favorite)	ROI (Bet on Underdog)
0%–10%	0	–	–
10%–20%	0	–	–
20%–30%	127	**1.754%**	-18.850%
30%–40%	751	-5.264%	-1.736%
40%–50%	2,277	-3.123%	-7.532%
50%–60%	3,362	-6.344%	-3.447%
60%–70%	2,337	-3.867%	-6.894%
70%–80%	596	-0.395%	-15.666%
80%–90%	28	-11.762%	**26.714%**
90%–100%	0	–	–

Table 6.7: **NHL:** ROI for favorite and underdog bets based on moneyline implied win probabilities.

7 Discussion

Our main purpose in this paper was to investigate the efficiency of the sports betting market in predicting game outcomes, using data from the most recent seasons of the four major professional sports leagues in the US, the NFL, NBA, MLB, and NHL. Such an analysis is of particular interest and relevance in light of the dramatic explosion of the sports betting market over the past several years.

Our study focused specifically on moneyline bets, a type of bet that amounts to a probabilistic prediction on the outcome of a game. Compared to more straightforward up-or-down bets such as over/under bets and bets against the spread, moneyline bets represent a richer environment within which to study the efficiency of the betting market, and they provide a wider range of opportunities for the bettor to exploit any inefficiencies.

Using four different metrics for evaluating probabilistic forecasts, we quantified the accuracies of predictions based on the moneyline betting market. We compared these accuracies to those of the Bradley-Terry model, a classical mathematical model for probabilistic predictions on game outcomes, as well as a baseline model that predicts the home team to win with a probability given by the home team winrate over a given set of past games.

Our main findings are as follows: First, and not surprisingly, among the three prediction models we analyzed, the betting market based model was consistently more accurate than the Bradley-Terry model, which in turn was significantly more accurate than the home bias coinflip model. More interestingly, across all four leagues, the accuracy rates of moneyline based predictions over the period covered in our study (i.e., the most recent 17 seasons) were comparable to those from studies of oddsmakers' predictions from the 1980s and 1990s. In this sense, the predictability of games has remained relatively steady over the past several decades. Among the four leagues, the NBA and NFL were significantly more predictable than the MLB and NHL. This is also in line with findings from earlier studies.

An analysis of calibration plots did not reveal any significant inefficiencies of the moneyline betting market at any parts of the moneyline spectrum. In particular, we found no evidence of a bias for or against strong favorites or longshots in the betting market. An analysis of returns on investment (ROIs) resulting from specific betting strategies based on moneyline implied win probabilities confirmed these conclusions: There is no such strategy that can be expected to consistently yield a positive ROI for the bettor.

We conclude with some comments and suggestions on possible extensions of our study and directions for future work. First, it would be interesting to see to what extent our conclusions remain valid in betting markets for other sports leagues such as global soccer leagues and college sports leagues. The four leagues covered in our analysis are the most prominent sports leagues in the US and they make up the lion's share of the sports betting market in the US. This may, in part, explain the relatively high efficiency of the betting market for these leagues observed in this study: A market of this size can be expected to be highly self-regulating, with any inefficiencies being eliminated in short order due to market forces. Such a high level of self-regulation may not be present in smaller betting markets such as those for less visible leagues and for niche sports.

A key feature of the four leagues in our study is that they are relatively homogeneous due to strict regulations such as salary caps and draft systems that favors the weaker teams. As a result, games in which one team is an extreme favorite or an extreme underdog are relatively rare within the four leagues covered in our study. This may not be the case in smaller sports leagues or in leagues that are less regulated, thus opening up additional opportunities for inefficiencies in the betting market. It would be interesting to explore this further.

Another area for future studies is to compare the betting market to more sophisticated mathematical models than the Bradley-Terry model we used here, which is based only on the *results* of past games, i.e., on the win/loss records of the teams involved. In particular, it would be interesting to study models that take into account additional quantifiable information from past games, such as the individual *scores* earned by each of the two teams. The data from Table 2.1 suggests that taking into account scores might provide a small boost in the accuracy of the model, but not nearly enough of a boost to overcome the gap to the accuracy of betting market based models.

References

[Bet25] BetMGM. Chargers at Chiefs: Betting trend for week 15. `https://www.chiefs.com/news/chargers-at-chiefs-betting-trend-for-week-15`, 2025. Accessed 2026-1-29.

[Bis26] James Bisson. US sports betting revenue & handle: Tracking betting market data 2026. `https://www.sportsbookreview.com/news/us-betting-revenue-tracker/`, 2026. Accessed 2026-4-15.

[Bri50] Glenn W Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3, 1950.

[BS03] Bryan L Boulier and Herman O Stekler. Predicting the outcomes of national football league games. *International Journal of Forecasting*, 19(2):257–270, 2003.

[BT52] Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, 39:324–345, 1952.

[Cod19] One-Off Coder. The Bradley-Terry-Luce model. `https://datascience.oneoffcoder.com/btl-model.html`, 2019. Accessed 2026-1-29.

[Fiv21] FiveThirtyEight. How to play our NFL forecasting game. `https://fivethirtyeight.com/methodology/how-to-play-our-nfl-forecasting-game/`, 2021. Accessed 2025-12-21.

[FM07] David Forrest and Ian McHale. Anyone for tennis (betting)? *The European Journal of Finance*, 13(8):751–768, 2007.

[GG97] Philip K Gray and Stephen F Gray. Testing market efficiency: Evidence from the NFL sports betting market. *The Journal of Finance*, 52(4):1725–1737, 1997.

[LMB18] Michael J Lopez, Gregory J Matthews, and Benjamin S Baumer. How often does the best team win? A unified approach to understanding randomness in North American sport. *The Annals of Applied Statistics*, 12(4):2483–2516, 2018.

[NC21] Philip WS Newall and Dominic Cortis. Are sports bettors biased toward longshots, favorites, or both? A literature review. *Risks*, 9(1):22, 2021.

[Sau98] Raymond D Sauer. The economics of wagering markets. *Journal of economic Literature*, 36(4):2021–2064, 1998.

[SBS07] ChiUng Song, Bryan L Boulier, and Herman O Stekler. The comparative accuracy of judgmental and model forecasts of American football games. *International Journal of Forecasting*, 23(3):405–413, 2007.

[Ste97] Hal S Stern. How accurately can sports outcomes be predicted? *Chance*, 10:19–23, 1997.

[WODM24] David Winkelmann, Marius Otting, Christian Deutscher, and Tomasz Makarewicz. Are betting markets inefficient? Evidence from simulations and real data. *Journal of Sports Economics*, 25(1):54–97, 2024.

[WW94] Linda M Woodland and Bill M Woodland. Market efficiency and the favorite-longshot bias: The baseball betting market. *The Journal of Finance*, 49(1):269–279, 1994.

[WW01] Linda M Woodland and Bill M Woodland. Market efficiency and profitable wagering in the National Hockey League: Can bettors score on longshots? *Southern Economic Journal*, 67(4):983–995, 2001.

[WW03] Linda M Woodland and Bill M Woodland. The reverse favourite–longshot bias and market efficiency in major league baseball: an update. *Bulletin of Economic Research*, 55(2):113–123, 2003.

[WW11] Linda M Woodland and Bill M Woodland. The reverse favorite-longshot bias in the National Hockey League: Do bettors still score on longshots? *Journal of Sports Economics*, 12(1):106–117, 2011.