On the Efficiency of Sports Betting Markets

Conference Paper · November 2022

1 author:

Thomas Robbins
East Carolina University

DECISION SCIENCES INSTITUTE

On the Efficiency of Sports Betting Markets

Thomas R. Robbins
East Carolina University
Email: robbinst@ecu.edu

ABSTRACT

Betting on sports is increasingly popular, and legal in the United States. Ever since the Supreme Court removed the PASPA ban on sports betting many states have moved to legalize sports betting both in-person and on-line. In this paper we evaluate the sports betting market and assess its efficiency in the financial sense. Using a large dataset of betting odds and outcomes we evaluate the weak-form efficiency across major North America sports. We find statistically significant inefficiencies in professional and collegiate football, college basketball, and MLB. We are however unable to demonstrate statistically significant inefficiencies in the NBA or NHL.

KEYWORDS: Sports Betting, Market Efficiency

INTRODUCTION

Gambling has been popular throughout human history, and gambling on sports has been popular as long as there have been sports. But the legality of sports betting in the United States has varied considerably over time. The Professional and Amateur Sports Protection Act of 1992, also known as PASPA or the Bradley Act, effectively outlawed betting on most sports throughout most of the US, making exceptions only for licensed sports pools in Nevada as well as lotteries in Oregon, Delaware and Montana. Excluded from the reach of PASPA were jai alai, as well as parimutuel horse and dog racing.

The situation changed dramatically in 2018 when the United States Supreme Court ruling in Murphy vs. NCAA struck down the PASPA law and returned the regulation of gambling to the states. In the years since many states have moved to legalize sports gambling, both in-person and on-line. Total betting numbers are uncertain, but in 2019 betting in the Vegas sportsbook alone exceeded $5 billion. In the first quarter of 2022 DraftKings Inc., one of the larger on-line sportsbooks, reported quarterly revenue of $417 million dollars, a 34% increase from the prior year (Jones 2022). Sports betting is now a large and growing financial market, increasingly legal and in the open.

In this paper we evaluate the sports betting market and assess its efficiency in the financial sense. Using a dataset of odds and results on over 140 thousand sporting contests across the major sports in North America that covers 15 seasons, we evaluate the accuracy and efficiency of the odds to assess if the markets can be considered weak-form efficient. Our analysis finds that inefficiencies exist in the odds to the degree that we conclude that some of the markets are not weak-form efficient. These inefficiencies vary from sport to sport, but are reasonably consistent over time. However, the level of inefficiency is small enough that it remains difficult for a bettor to exploit them profitably in the short term or on an after-tax basis.

LITERATURE REVIEW

Capital Market Efficiency

The focus of our analysis is on the efficiency of the sports betting markets. Market Efficiency is a concept first developed in the economics and finance literature as defined by the efficient market hypothesis. (EMH). The efficient-market hypothesis (EMH) is a hypothesis in financial economics that states that asset prices *fully reflect* all available information. A direct implication is that it is impossible to "beat the market" consistently on a risk-adjusted basis since market prices should only react to new information (Wikipedia 2022).

A key review of the theoretical and empirical literature on the Empirical Market Hypothesis is provided in Fama (1970). Fama analyzes efficient markets relative to three board information sets. Weak form efficiency is based on the use of historical prices, semi-strong efficiency on publicly available information, and strong efficiency is based on all information public and private. In this framework the possibility of trading systems based on the relevant information set generating excess returns, returns in excess of equilibrium expected profits are ruled out. Stated simply, excess profits are not possible from a trading system in a market that is efficient. The empirical analysis in this paper provides reasonably strong support for the weak, and semi-strong forms of market efficiency. The analysis however identified exceptions to the strong form of market efficiency whereby market makers and corporate insiders can exploit their monopolistic access to information to earn returns in excess of the expected risk-adjusted rate.

Sports Betting Market Efficiency

A large body of research published beginning in the 1980s and 90s examined efficiency in sports betting markets. Much of this early research focused on pari-mutual and fixed odds systems in horse racing. A comprehensive review of this literature is provided in (Kuypers 2000). Kuypers reviews 5 papers that examine pari-mutual systems, 6 papers on odds-based systems, and 4 papers on spread based systems. Seven of the reviewed papers assess weak form efficiency, while six examine semi-strong efficiency and two assess strong-form efficiency.

Kuypers uses the following definitions of efficiency in the sports betting context:

- Weak form: no abnormal returns, either to the bookmaker or the bettor, can be achieved solely from price information. An abnormal return is defined as a return different from the bookmaker’s expected take.
- Semi-strong: no abnormal returns can be achieved from odds or any publicly available information.
- Strong: no abnormal returns can be achieved by any group in society incorporating odds publicly available and privately available information.

Our analysis will focus on weak form efficiency. The implication of weak form efficiency is that the return on bets in any odds range to a bettor will be negative, consistent across odds ranges, and equal to the bookmaker’s average hold.

Kuyper tests weak-form efficiency in the betting market for UK football (soccer) analyzing 3882 matches from 1993-95. He divides the bets into 20 bins and calculates the expected after-tax return from taking all bets in each bin and compares those returns to the expected after-tax return of -18.5% implied by the fixed hold. The analysis finds that all returns are negative. The best return is -3.13% in the odds bin where the mid-point implied probability is 49%; the slight underdog. His analysis concludes that while market inefficiencies exist, no formula based simply on betting an odds range will yield a positive return. He further concludes that there is no systematic bias in the odds by regressing the actual win probabilities against the implied win probabilities for each group and failing to reject the null hypothesis that the slope is equal to one.

A more recent examination of English football is performed in Deschamps and Gergaud (2007). They analyze 8,377 matches between 2002 and 2006 with odds from six different bookmakers. They also find considerable variation across odds groups, but no positive returns. Another assessment of English football odds is provided in (Direr 2011). Direr evaluates 11 years of odds (200-2011) from 6-10 odds makers, for a total of nearly 80,00 games and 2.8 million betting opportunities. He finds that positive returns are available in the range of 2.8% for average odds, and 4.4% for best odds by betting on overwhelming favorites.

Other papers perform similar analysis on other sports. Levitt (2003) examines NFL games. (Hickman 2020) – looks at the NCAA "March Madness" basketball tournament. Gandar, Zuber et al. (2004) examines the National Hockey League (NHL) while Gandar, Zuber et al. (1988) looks at point spreads in NFL games. They implement two tests and come up with mixed results. A statistical test fails to reject rationality, while an economic test does reject rationality.

Longshot Bias

A specific type of inefficiency, and a frequent topic of analysis in betting markets, is the so-called longshot, or favorite-longshot, bias (FLB). An early review and assessment of this phenomenon was presented in the inaugural Anomalies series in the Journal of Economic Perspectives (Thaler and Ziemba 1988). This paper analyzed parimutuel betting and re-asserts the criteria that in a weekly efficient market no best should have a positive expected value, and in a strongly efficient market all bets have an expected value of (1-*t*), where *t* is the racetrack's fixed take. The review demonstrates that the returns are systematically associated with the odds. Bets on favorites earn an above average return, while bets on longshots earn below average returns. Parimutuel odds are directly set by the amount bet so the longshot bias indicates that bettors systematically overestimate the probability that a longshot will pull off the upset. The longshot bias has at least two alternative explanations; risk seeking behavior by the bettor, or misestimation of the odds in extreme scenarios. Odds misperception is consistent with Prospect Theory's assertion that individuals overestimate the likelihood of rare events (Kahneman and Tversky 1979). A detailed comparison of these two possible explanations is presented in Snowberg and Wolfers (2010), who argue the data support the misperception hypothesis. A textbook level description of the phenomenon is provided in (Ottaviani and Sørensen 2008). They review the two explanations discussed above, as well as several others. A more recent review of the literature on the longshot bias is presented in Newall and Cortis (2021).

One additional explanation of the longshot bias is that some bettors posses private (inside) information. This approach, sometimes known as the Shin model, was developed by Hyun Song Shin in the early 1990s (Shin 1991, Shin 1992, Shin 1993). It has been further explored in subsequent papers (Cain, Law et al. 2003). Empirical analyses of the longshot bias have been published for sports such as UK football (Peel, Cain et al. 2000) and major league baseball (Gandar, Zuber et al. 2002).

THE DATA SET

Our data set includes odds and results on major professional and collegiate sports. Data has been collected from the website Sports Book Review (TopSportsbooks 2022). Odds are provided for professional football (NFL), college football (CFB) professional basketball (NBA), college basketball (CBB), major league baseball (MLB) and professional hockey (NHL).

Sport Book Review provides a single file for each sport for each season. The type of data varies from sport to sport and even season to season, and the data is not without issue. Considerable effort was required to merge and clean the data. After eliminating records without the required odds we were left with the following data set.

Table 1- Data Set Summary

Data Set Summary
Through April 2022

League	From	To	Games	Beg	End	Seasons	Teams
CBB	2007-11-05	2022-02-28	58,874	2007-08	2021-22	15	385
CFB	2007-08-30	2022-01-10	13,023	2007-08	2021-22	15	263
MLB	2010-04-04	2022-04-24	27,776	2010	2022	13	30
NBA	2007-10-30	2022-04-24	19,098	2007-08	2021-22	15	30
NFL	2007-09-06	2022-02-13	4,025	2007-08	2021-22	15	32
NHL	2007-09-29	2022-04-24	18,814	2007-08	2021-22	15	33

The data set includes 141,610 games from August 2007 through April 2022.

BETTING ODDS AND PROBABILITIES

The menu of bets that can be made on sports is very large. Bets can be placed on almost anything related to a game, a team, or even individual performances of players. With minor variations from sport to sport, the main betting options have three different components: totals, spreads and moneyline.

- Totals: the total, or over/under, is a bet on the total points scored in the game. Bettors can bet the total points will be over, or under the stated line.
- Spreads: a bet on a team to win be a certain margin. The underdog is bet with plus points, the favored with negative points.
- Moneyline: a straight bet on what team will win the game. Moneyline bets are made with differential payouts such that a bet on a favorite will risk more than can be won, while a bet on an underdog will return more than the amount risked.

Note that both totals and spread bets are quoted along with moneyline odds so that the payout to a winner is less than the amount risked. Odds are stated in different equivalent formats in different locations and different settings. In the United States odds are most often quoted in American Odds format.

In the American format the odds can be expressed as either a positive number or a negative number. A positive number shows the profit a successful wager will return on a $100 bet. So, for example, a bettor who wagers $100 at +110 odds and wins, will earn a profit of $110, plus the original wager of $100 for a total payout of $210. Positive odds typically imply the team is an underdog. Conversely, negative odds show how much a bettor must risk to earn a $100 profit. So, for example if a bet is made for $120 at -120 odds, the successful bettor will receive a profit of $100, plus the original wager of $120 for a total payout of $220. The favorite team is given negative odds, but in some evenly matched games both teams may have negative odds. More formally the Payout *P* to a wager of stake *S*, at odds *M* are given by equation (1).

$$P = \begin{cases} S \times \dfrac{M}{100} + S & \text{for M} > 0 \\ \dfrac{S}{-M/100} + S & \text{for M} \le 0 \end{cases} \tag{1}$$

Odds of +100- and -100 are equivalent. In practice *M* is always quoted as a number with an absolute value greater than or equal to 100. So, while odds of -125 and +80 would both return a profit of $80on a $100 bet, the odds are always quoted as -125.

Moneyline odds carry an implied probability of success. The implied probability is the probability at which a bettor is indifferent to taking either side of the bet. The probability calculation in the American odds format again depends on whether the odds are positive or negative. So, for a bet with odds *M*, the implied probability *p* is given by

$$p = \begin{cases} \dfrac{100}{M+100} & \text{for M} > 0 \\ \dfrac{-M}{-M+100} & \text{for M} \le 0 \end{cases} \tag{2}$$

While equation (2) gives the odds on one side of a bet, the bookmaker quotes odds in pairs. So, for example, a bookmaker might quote odds of -120 for a favorite and +110 for the underdog. Converting each of these to implied probabilities gives probabilities of 54.5% and 47.6%. These odds are not fair, in the sense that they add up to more than 100%. The excess probability, in this example 2.1%, is the book margin (*k*), sometimes referred to as the *vig* or the *juice*. The book margin exists so that the bookmaker is guaranteed a profit as long as bets are made in the appropriate proportion. Book margins in the range of 3%-5% are common.

In order to convert the bookmaker's odds into meaningful probability estimates the odds must be converted to consistent probabilities. Draws are rare in the sports we are evaluating. So, if the contest ends in a draw all win-lose bets are effectively cancelled and bettors are returned their original stake. The most common way to convert the implied probabilities is a simple normalization process. So, for a contest with implied probabilities of p_1 and p_2, the normalized probability that team 1 will win the game and bets will pay is

$$p_{1_n} = \frac{p_1}{p_1 + p_2} \tag{3}$$

THE SPORTSBOOK'S MARGIN

Because the implied odds are unfair, they add up to more than one, the sports book has a built-in advantage. The excess probability gives the sportsbook a built-in margin, appropriately allocated bets on either side will guarantee the book a profit. The sportsbook's profit margin is proportional to the book sum, the excess implied probability in the stated odds. If we have a two-way bet with implied odds p_1 and p_2, then the booksum *k*, is given by

$$k = p_1 + p_2 - 1 \tag{4}$$

The bookmaker's margin (*m*), also known as the *hold*, is the sportsbook's average profit and can be shown to be

$$m = \frac{k}{k+1} \tag{5}$$

The booksum and margin varies from game to game, and league to league, but typically averages in the 3% range. Summary metrics for our dataset by league are shown in Table 3

Table 2- Average Margin

Booksum and Margin by League

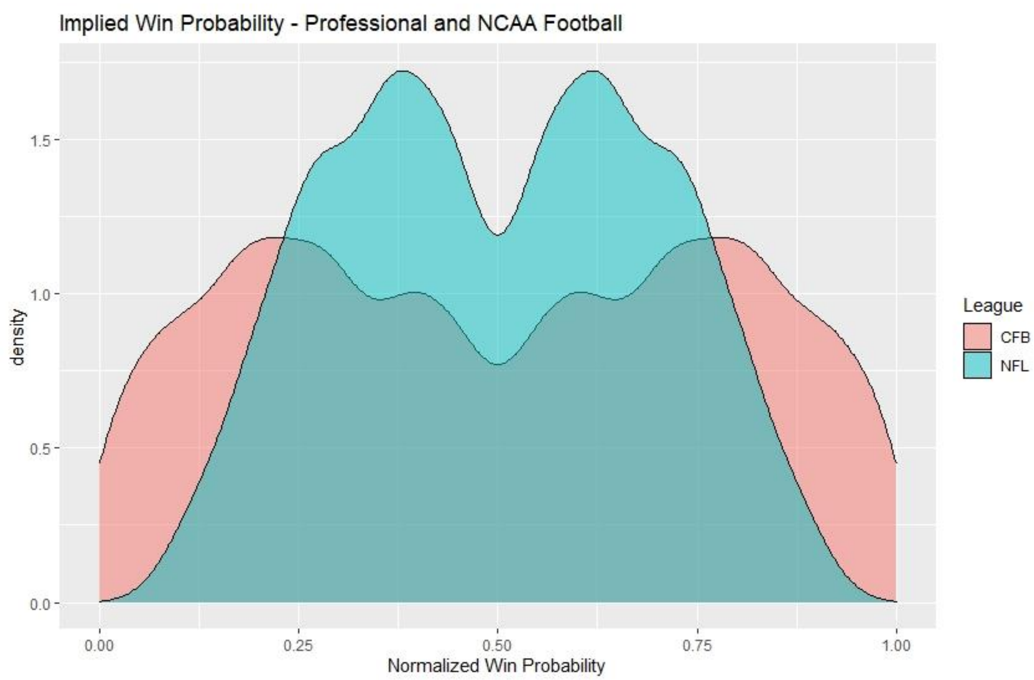
2007- 2022

League	K	M
CBB	4.06%	3.90%
CFB	3.50%	3.38%
MLB	2.92%	2.84%
NBA	3.77%	3.64%
NFL	3.79%	3.65%
NHL	3.43%	3.32%

EMPIRICAL ODDS DISTRIBUTIONS

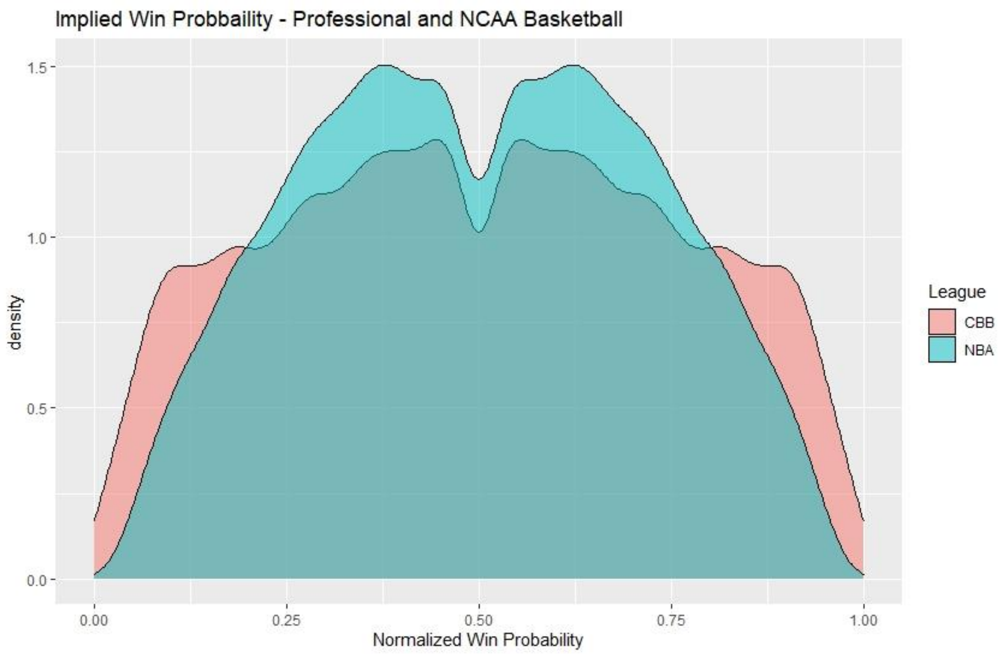
Before we investigate efficiency in detail, let us examine the distribution of odds for each major sport. The following graphs represent the normalized odds for each sport, plotted as a density graph using the ggplot library in R. Not that each graph is by design symmetrical since each game is represented by the normalized odds of the favorite and the underdog which by definition must add to one. Figure 1 shows the density plot for football, both professional and collegiate.

Figure 1-Football Odds



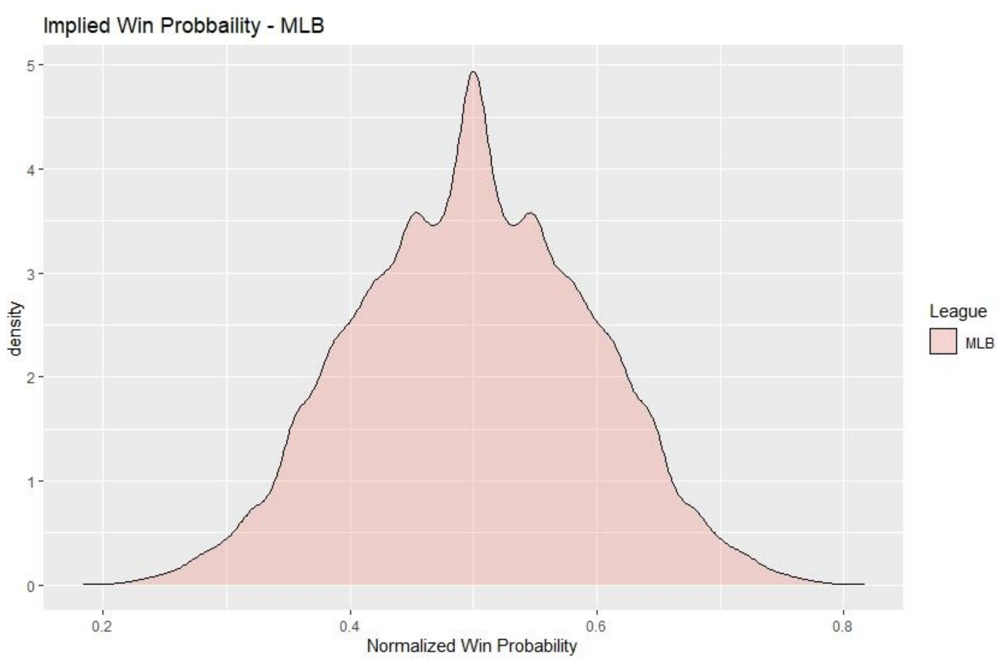
This graph reveals a few interesting properties about this data. First the odds for college football are more lopsided than for the NFL implying a higher level of parity at the professional level. Secondly, slight favorite-underdog matchups are more common then even (pick'em) odds. The NFL odds are bimodal with a peak around 65%, and due to symmetry a corresponding peak near 35%. The modal odds for NFL underdogs is +170, corresponding to an implied probability of 37%. The density drops of sharply with a trough at 50%. It appears the odds are defined so that having one team as a small favorite is more common than an even-money bet.

Figure 2-Basketball Odds



Odds for professional and college basketball have a similar distribution to football. College odds are more dispersed, and the slight favored effect is in place for both leagues. A similar trough exists for both basketball leagues with pick'em odds being less popular than slight favorite matchups.

Figure 3-MLB Odds

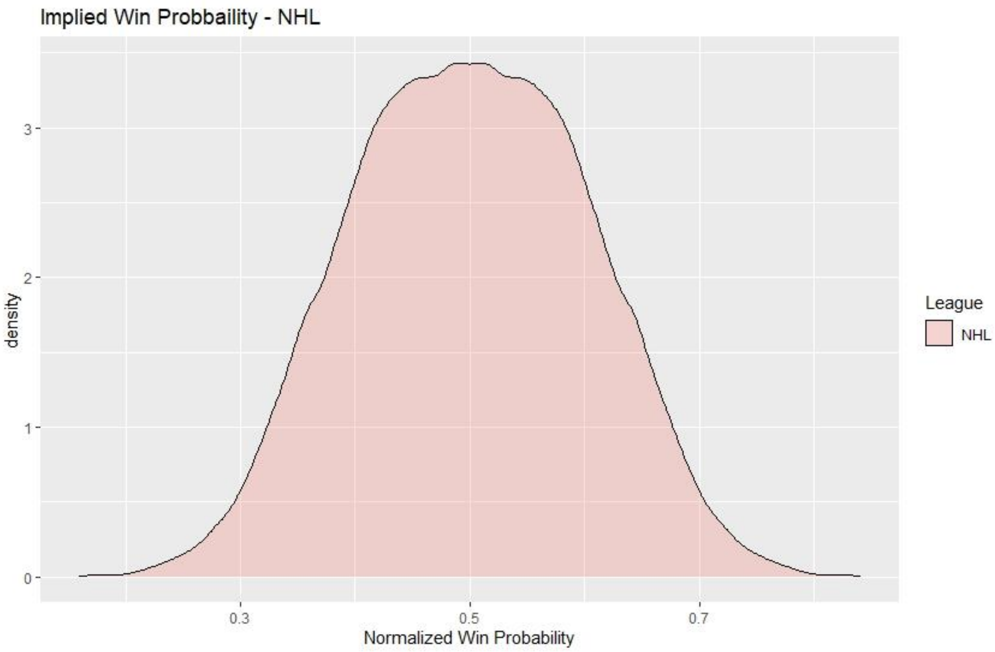


Odds for professional hockey appear quite different. The pick-'em trough is absent and the modal probability is 50%. There is in fact a small bump in the distribution near the 50% level.

Extreme odds are also much less common. Whereas football and basketball both had extreme probabilities, at or near 100%, odds greater than 75% appear quite rare.

The odds for professional hockey are similar to baseball, but with a pronounced peak near parity. Odds can be a little more extreme in hockey with the density extending to nearly 80%.

Figure 4-NHL Odds



RETURNS BY ODDS GROUPS

We now examine the return on bets made at different odds ranges. To do this we take the entire data set of games in each sport sorted by the implied probability from low to high. Note once again this implies each contest is represented by two records, one for each team. But, unlike the normalized odds this data is not symmetric since we are using the quoted odds which includes the bookmaker's margin. For the purpose of this analysis, we divide the odds into twenty bins of approximately equal count. We assume we bet $100 on every contest and determine the average profit earned on those bets. Recall that in a weak-form efficient market the returns on all these bets would be negative, equal to each other, and equal to the bookmaker's average hold. If a longshot bias exists, we would expect to see higher returns for the favorites and lower returns for the underdogs.

National Football League

In Figure 5 we see the returns for bets on NFL games. Recall that this is based on a data set of over 4,000 games over 15 seasons. The graph reveals several issues that nominally support the notion of inefficiencies in the market.

- The return on bets made in each group does not appear to be constant by group.
- The return on bets in some odds categories is positive.
- While not readily apparent from the graph, the average return across all bets is -4.13% which is lower than the expected profit of -3.65% indicated by the average hold.

The graph does seem to indicate a longshot bias, returns on bets on the lower probability teams have lower returns than bets on the highly favored teams.

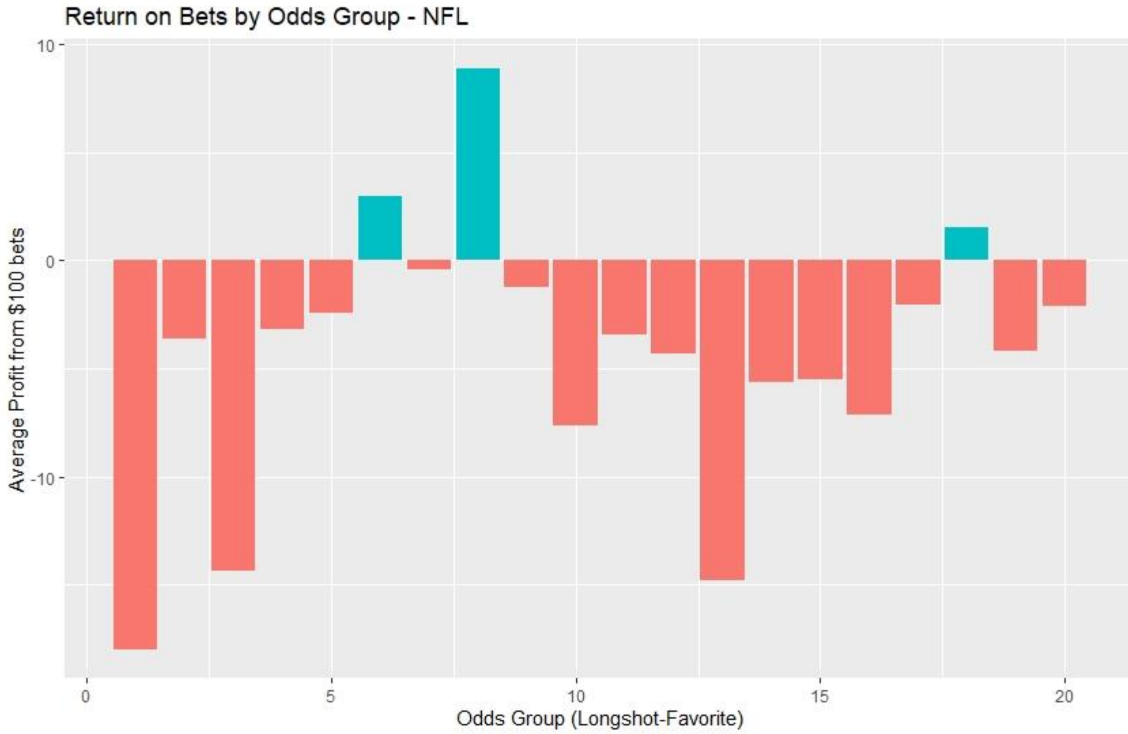


Figure 5-NFL Returns

The odds group with the strongest positive returns corresponds to a bin 8 where odds range from +125 to +144, corresponding to implied win probabilities of 41% to 44%. The two bins below that level also offer either a very small loss, or a positive return. It is worth noting that these games correspond roughly to the bump in the density of implied probabilities in Figure 1.

In Table 3. We examine the consistency of these returns over time.

Table 3- Returns on NFL Bets by Season

Return on NFL Moderate Underdog Bets by Season

Season	[+160,+178]	[+145, +159]	[+125, +144]
2007-08	3.5%	−31.0%	43.8%
2008-09	−46.2%	15.4%	12.6%
2009-10	13.2%	12.7%	20.5%
2010-11	3.4%	20.0%	3.2%
2011-12	−6.7%	−5.0%	10.7%
2012-13	37.2%	19.8%	5.0%
2013-14	−25.7%	−61.1%	−13.0%
2014-15	−24.5%	−8.3%	29.0%
2015-16	−11.4%	10.4%	−22.1%
2016-17	19.5%	−49.0%	3.9%
2017-18	−19.9%	−28.8%	5.9%
2018-19	3.9%	5.5%	−11.9%
2019-20	26.4%	19.1%	24.9%
2020-21	18.6%	25.6%	7.8%
2021-22	16.9%	25.5%	35.0%

The returns in these 3 bins, representing 30% of the games in each season, vary considerably. Each bin has positive returns in some seasons, and negative returns in others. Bin 8, odds in the range of [+125-+144] has the highest long-term return and the most consistently positive returns; 12 out of 15 seasons.

While the odds appear to indicate inefficiencies, the statistical validity of inefficiency is marginal. A test of a systematic bias as per Kuypers we regress the predicted win probability in each bin against the actual win probability. This test failed to reject the null hypothesis that the slope of that line is one; the 95% confidence interval is [.95, 1.07]. So, we cannot conclude that there is a systematic bias across the range of odds. While the returns appear different across the bins, an ANOVA test of the null hypothesis that the returns are the same for each bin, has a p-value of .311, so our ability to reject the hypothesis of equal returns is marginal at best. If we examine the return on the most profitable bin, bin 8 with odds of [+125,+144], and test the null hypothesis that these returns are negative, we can reject that hypothesis with a p-value of .043. We cannot reject the null hypothesis for negative returns on bin 6, the p-value is .316. Finally, to perform a more formal test of the longshot bias we perform a two-sample hypothesis test on the returns in bin 1 and bin 20, the biggest underdogs and biggest favorites. While there appears to be a strong difference in the graph, the p-value of this test is .164; again, too high to confidently reject the null hypothesis with much confidence.

In summary there is evidence that would suggest that the odds on NFL games are weak form inefficient. There appears to be a variation in return against different odds groups and heavy longshot bets have historically performed the worst, but these conclusions are tentative due to the variable state of the returns. The one test that does meet the 5% conventional threshold, barely, is the return on slight underdogs in the range of [+125,+144]. We can reject the null hypothesis that returns on these bets have a negative return and therefore conclude that the weak form efficiency does not strictly hold.

College Football

In
Figure 6 we see the graph for college football. The CFB graph is similar to the NFL graph; returns are uneven, heavy longshots yield very low returns relative to other groups. There is also a set of profitable bets in the slight underdog range, in the case of CFB those odds are in the range of +285 to +150. These represent implied win probabilities of about 26% to 40%.

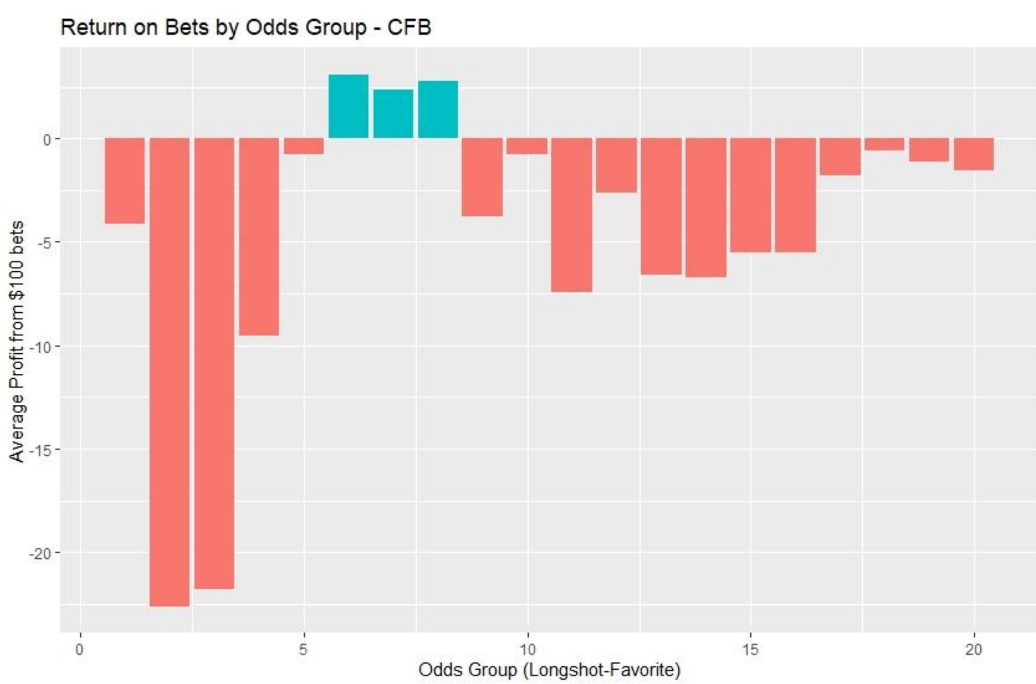


Figure 6-CFB Returns

As was the case with professional football, the returns in these bins are sometimes positive and sometimes negative over the course of a season, thought there is less consistency than there was for the NFL.

Table 4- Returns on CFB Bets by Season

Return on CFB Moderate Underdog Bets by Season

Season	[+235,+289]	[+185, +234]	[+120, +149]
2007-08	−4.9%	−8.2%	2.3%
2008-09	23.5%	−30.1%	7.8%
2009-10	−8.5%	−6.8%	3.5%
2010-11	3.0%	3.7%	15.8%
2011-12	8.0%	23.6%	14.3%
2012-13	−10.2%	11.8%	5.2%
2013-14	−10.9%	9.9%	−19.6%
2014-15	23.4%	−3.8%	0.1%
2015-16	3.6%	−7.8%	−32.8%
2016-17	18.2%	17.2%	−2.9%
2017-18	3.0%	−6.5%	9.9%
2018-19	−4.4%	5.9%	10.2%
2019-20	−17.7%	5.9%	20.6%
2020-21	8.6%	12.2%	−5.1%
2021-22	15.1%	11.2%	−4.4%

There is some weak evidence that there is a general difference across all bins, the ANOVA test has a p-value of 0.108. While the returns on bins 6-8 have been positive over an extended period of time, we have limited evidence that this is a statistically significant difference. If we test the null hypothesis that the returns are negative the p-values on these tests are 0.316, 0.281, and 0.219 respectfully. To test for a longshot bias we compare bins 2-19, and 3-18. For the bin 2 bin

3 test we can safely reject the null hypothesis with a p-value of 0.0068. For the bin 3-18 test we can reject the null with a p-value of 0.00077.

So, in summary we have some evidence to reject the weak form efficiency of CFB odds. The strongest evidence to reject efficiency is the longshot bias of bets on teams with odds in the range of [+480, +1495] have statistically significantly lower returns that betting the favorites in those contests at [-652, -3000]. While these are better bets, they have negative returns. The odds that do show positive returns are positive with questionable significance.

Professional Basketball

The returns for professional basketball are shown and Figure 5. (Figure 6 shows the returns for college basketball.) These graphs are quite different than what we saw for football. Pro basketball returns are generally negative, with a very small positive return in bin 8. The p-value for the ANOVA test of equal returns is 0.866. A small positive return is shown in bin 8, but the p-value of 0.463 provides very low confidence that this is a meaningfully positive return. The longest odds have a negative return of about -9.2% as compared to its complement -3.01%. The p-value for the difference being non-zero is only 0.347.

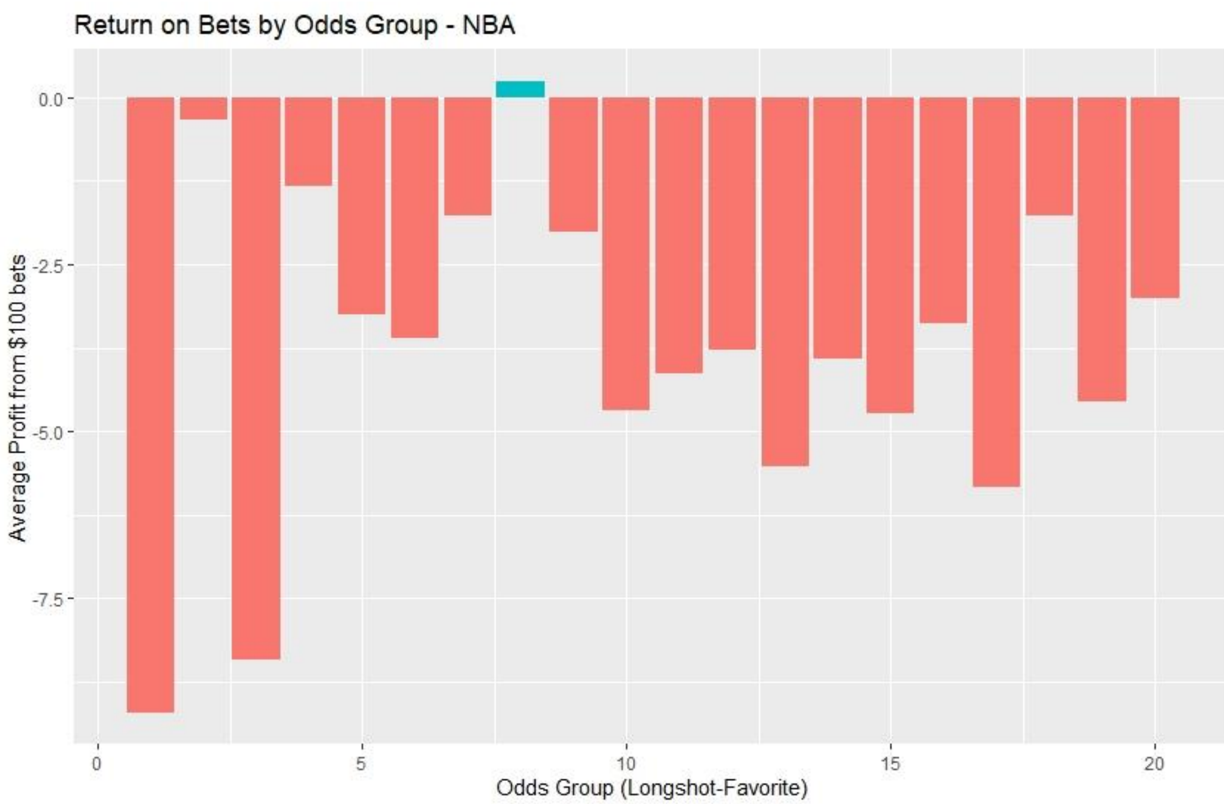


Figure 7-NBA Returns

So, while the graph for NBA odds shows some variation, a small positive return for some bet ranges, and a modest longshot bias, none of these claims can be substantiated at a reasonable level of statistical significance. We therefore cannot conclude that NBA odds are not weak form efficient.

College Basketball

The returns on best in college basketball show significant variation. A formal ANOVA test confirms this with a p-value of less than 2E-16. College basketball returns also show the most significant longshot bias of all the sports examined in this paper. While returns are negative on all bins, the biggest longshot bin has a negative return of approximately 50%. Betting on the corresponding favorite has a return of -0.002. The returns are different with a p-value of less than 2.2E-16. Recall that the odds are more lopsided in college vs. pro basketball and there are more mis-matches. Betting on the heavy underdog to pull of the big upset in college basketball is on average, the worst bet among the major sports. The longshot bias also holds for bins 2-19 (p-value 2.84 E-05), bins 3-18 (p-value 5.52 E-05), bins 4-17 (p-value 1.56 E-05) and bins 5-16 (p-value .00013).

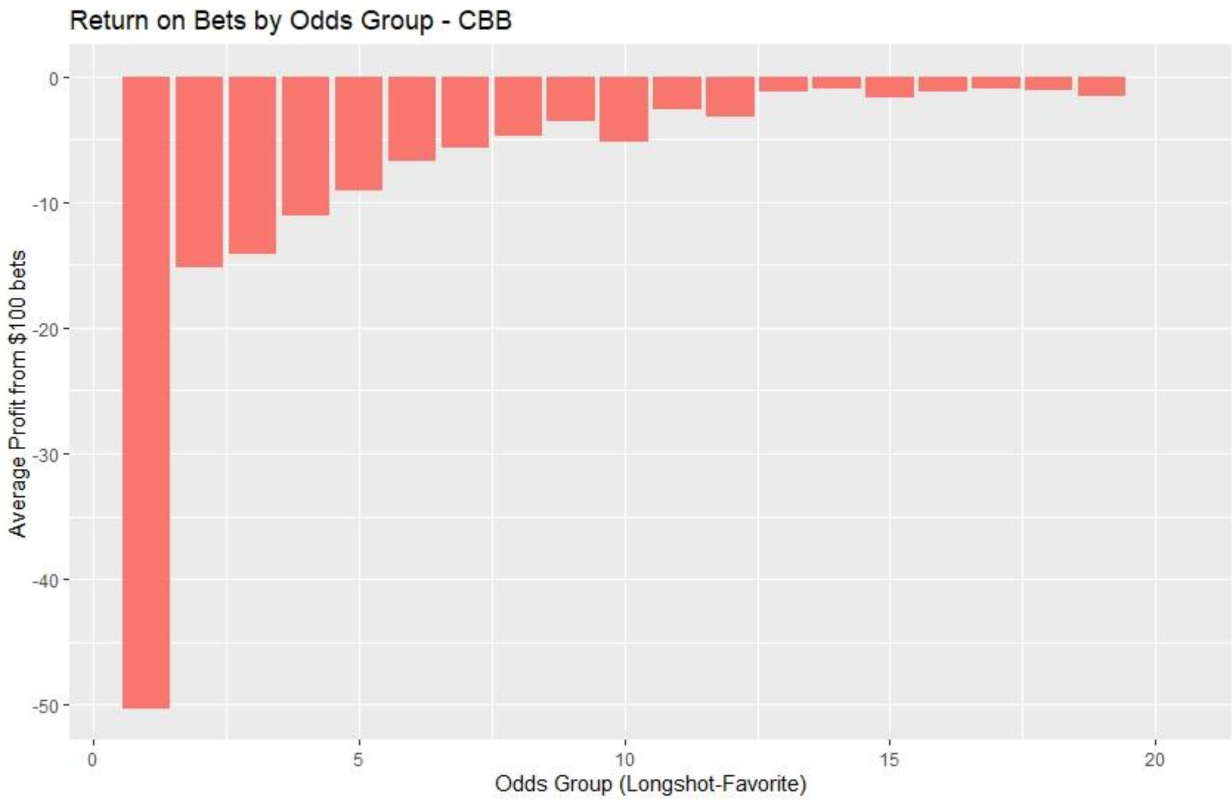


Figure 8-CBB Returns

So, in conclusion, while all odds ranges in college basketball have negative returns there are some bets clearly worse than others. College basketball shows a very strong favorite-longshot bias. Bets on extreme underdogs have very poor returns, and in general bets on underdogs perform significantly worse than bets on favorites. So based on the inconsistency of returns on odds groups, and longshot bias we can reject the hypothesis that college basketball odds are strictly weak-form efficient.

Major League Baseball
Bets in baseball appear generally consistent, at least more consistent than the other sports we have looked at. The returns on bets in each odds group is shown in Figure 9. The p-value on the ANOVA test that all returns are equal is 0.203. Generally, have a negative return over time, though there is a small positive return on bet in bin 14 corresponding to odds of [-125, -131] a slight favorite. The return is 1.48%, but the p-value associated with the test that it less than zero is 0.196.

A longshot bias is clear with bets on the biggest underdogs yielding a return of -7.51, with the corresponding favorites yielding a -0.44% return. The returns are different with a p-value of 0.016.

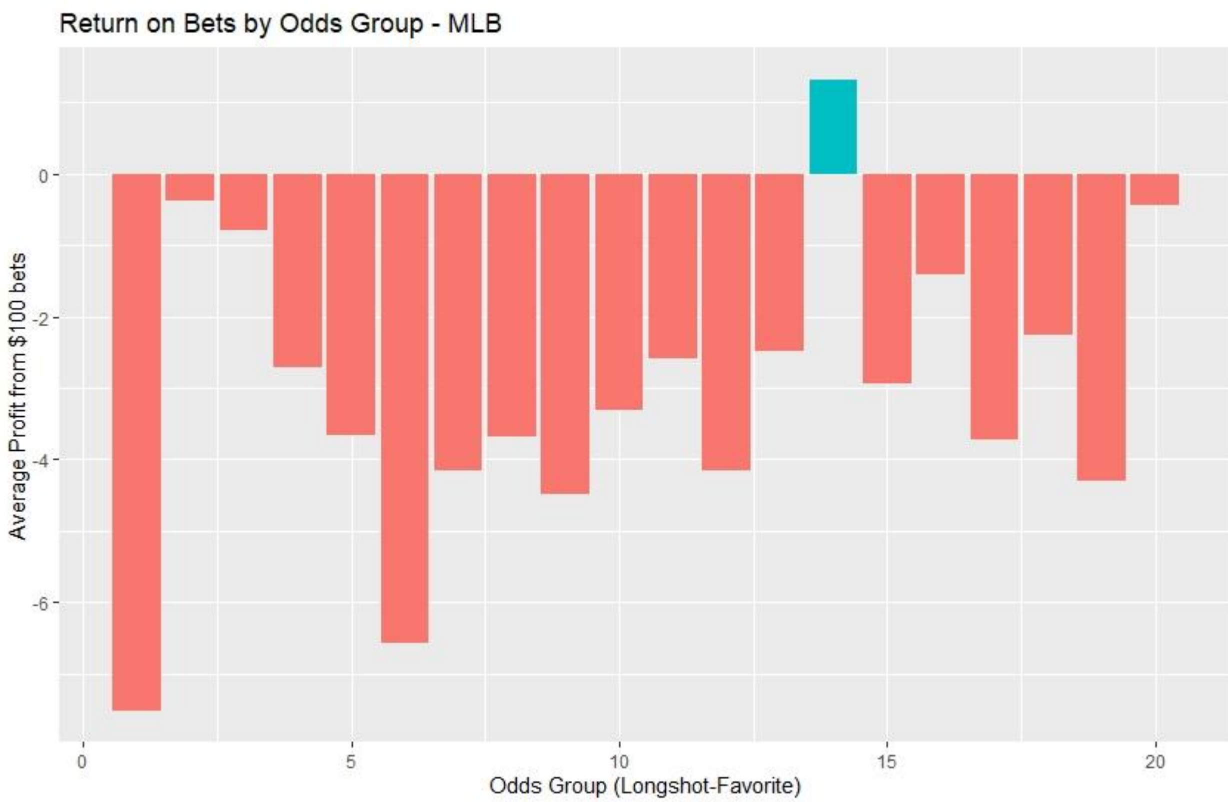


Figure 9-MLB Returns

So, again we can reject the hypothesis that MLB odds are strictly weak form efficient. In this case the rejection is based on a statistically significant longshot bias. While one band of favorite odds has a nominally positive return, we can not conclude that these returns are non-negative at a reasonable level of statistical significance.

Professional Hockey
Professional hockey returns are shown in Figure 10. They are similar to baseball in that they are all mostly negative and reasonably consistent. The p-value for an ANOVA test of the hypothesis that all returns are equal is 0.876, so we can not reject the null hypothesis.

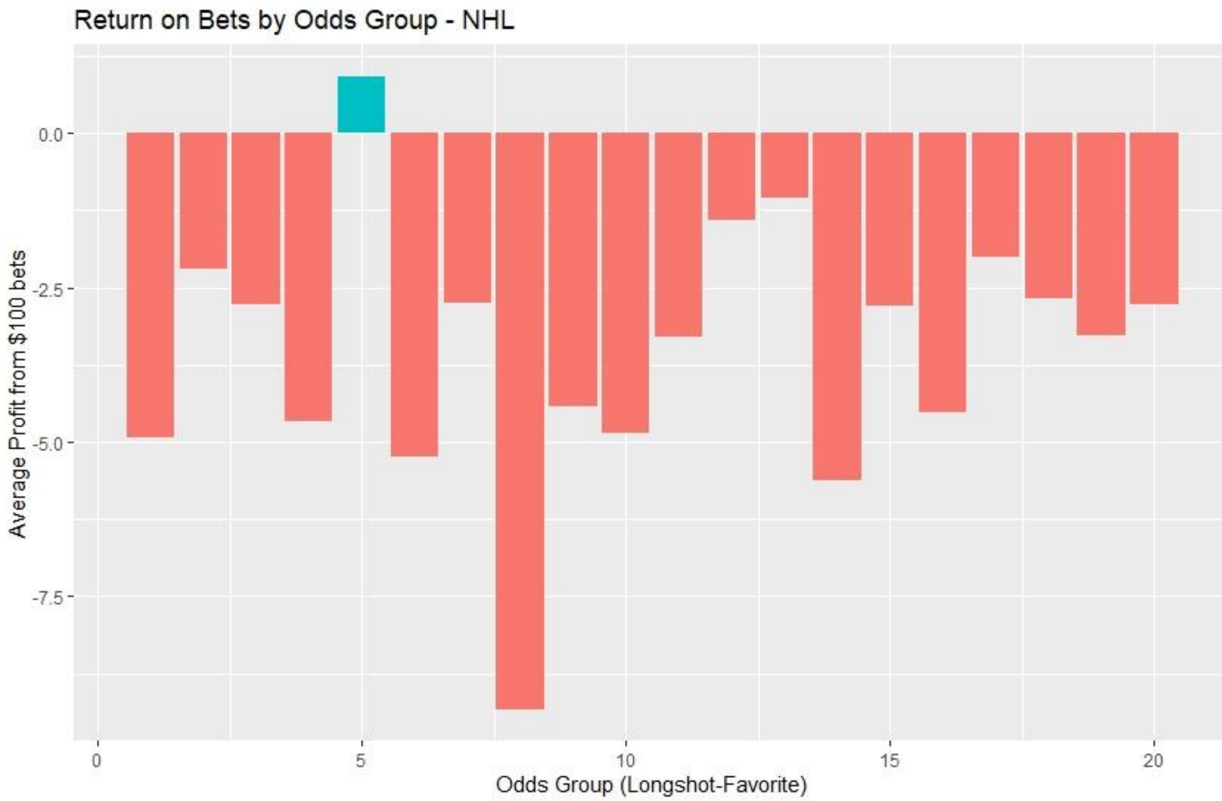


Figure 10-NHL Returns

The returns in bin 5 [+130, +136] are nominally positive at 0.929%. The p-value with associated with the hypothesis test that the returns are negative is 0.364, so we cannot reject the null hypothesis. There does not appear to be any clear longshot bias in the hockey odds. The return on the biggest longshots, bin 1 [+190, +505] are -4.92%, slightly worse than the returns on the complementary odds in bin 20 [-700, -219] which are -2.77%. We fail to reject the null hypothesis that these returns are different with a p-value of .546. The returns on the set of bins 2 and 19, show a nominal reverse longshot bias. Bin 2 returns are -2.19% and bin 19 returns are -3.28%, but the null hypotheses that these returns are different is 0.76%.

NHL odds appear to be the most efficient of the odds set we have examined. While there are some nominal differences in the historical returns, and one bin is slightly positive, none of the differences can be shown to be the result of anything but statistical noise.

CONCLUSION

Betting on sports is becoming increasingly popular, and legal, in the United States. Profitable betting is however very difficult. The sportsbook has several advantages, the most significant of which is their ability to offer unfair bets. These unfair bets create a margin for the book which allows them to profit regardless of the outcome as long as bets are placed in the appropriate proportion.

With the odds literally stacked against the bettor, it is difficult to win. That being said, our analysis has shown that not all bets are equally risky, and the odds markets are not weak form efficient across all sports. With the exception of the NBA and NHL, we have been able to identify inefficiencies in the odds markets using a large set of bets over an extended period of time. In that sense the moneyline betting market on major North American sports is inefficient.

At the most basic level the returns on bets placed on different ranges of odds yield different returns. But many of these differences are indistinguishable from random variation, while others are meaningful. Bets on slight underdogs in the NFL have a statistically significant positive return. Bets on extreme long shots in college basketball yield very poor returns while bets on corresponding heavy favorites in college basketball also yield negative returns, but these returns are significantly better. Our data indicates a statistically significant longshot bias in college basketball that extends over much of the odds range. Betting underdogs in general, and longshots in basketball is in general a losing proposition

In summary, with the exception of the NBA and the NHL, we have identified some level of statistically significant inefficiency in each of the sports analyzed. In summary:

- Professional football: bets on slight underdogs [+125,+144] have statistically significant positive returns.
- College football: a statistically significant longshot bias exists. Bets on underdogs in the range of [+480, +1495] have statistically significantly lower returns that betting the favorites in those contests at [-652, -3000].
- College basketball: an ANOVA test strongly rejects the hypothesis that returns are equal across all odds ranges. A statistically significant longshot bias exists with returns on underdogs up to +254 having statistically lower returns than bets on their opponents.
- Major League baseball: a statistically significant longshot bias holds where bets on the most extreme underdogs, up to +180, have lower returns than bets on the corresponding opponents.

While our analysis shows that inefficiencies exist in some betting markets, it does not imply that making money betting on sports is in anyway easy. Where positive returns exist, the pre-tax returns are small. And while they are positive over the long run, they are punctuated with long periods of negative returns. The strategy analyzed in this paper, bet on all opportunities in a certain odds range, is not recommended, nor is it likely to be profitable after tax in the short to medium term. But what our analysis does show is that some bets are better, or worse, then others. Betting on longshots in college basketball, is for example a strategy very unlikely to be successful. Bets on slight underdogs in football, on the other hand, are more likely to be successful.

A limitation of this paper is that we only analyzed a strategy of placing bets based on the odds, and therefore tested for weak form efficiency. An open issue, and an area for further research, is semi-strong efficiency. In a world of big data, AI, and machine learning, is it possible to build models that will predict outcomes successfully enough to overcome the sportsbook's hold and yield profitable results? While models may be very accurate in terms of predicting outcome, the efficient market hypotheses suggests that the output of those models would be quickly reflected in the price of the bets and profitable opportunities would be removed.

REFERENCES

Cain, M., D. Law and D. Peel (2003). "The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets." Bulletin of Economic Research **55**(3): 263-273.
Deschamps, B. and O. Gergaud (2007). "Efficiency in Betting Markets: Evidence From English Football." The Journal of Prediction Markets **1**: 61-73.
Direr, A. (2011). "Are betting markets efficient? Evidence from European Football Championships." Applied Economics **45**(3): 343-356.
Fama, E. F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work." The Journal of Finance **25**(2): 383-417.
Gandar, J., R. Zuber and R. Johnson (2004). "A Reexamination of the Efficiency of the Betting Market on National Hockey League Games." Journal of Sports Economics - J SPORT ECON **5**: 152-168.
Gandar, J., R. Zuber, R. Johnson and W. Dare (2002). "Re-examining the Betting Market on Major League Baseball Games: Is There a Reverse Favourite-Longshot Bias?" Applied Economics **34**: 1309-1317.
Gandar, J., R. Zuber, T. O'Brien and B. Russo (1988). "Testing Rationality in the Point Spread Betting Market." The Journal of Finance **43**(4): 995-1008.
Hickman, D. C. (2020). "Efficiency in the madness? examining the betting market for the NCAA men's basketball tournament." Journal of Economics and Finance **44**(3): 611-626.
Jones, R. (2022). "DraftKings reports US$417m revenue for Q1 2022." Retrieved 5/26/2022, from https://www.sportspromedia.com/news/draftkings-q1-2022-revenue-financials-sports-betting-us/.
Kahneman, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." Econometrica **47**(2): 263-291.
Kuypers, T. (2000). "Information and Efficiency: An Empirical Study of a Fixed Odds Betting Market." Applied Economics **32**: 1353-1363.
Levitt, S. (2003). "How Do Markets Function? An Empirical Analysis of Gambling on the National Football League."
Newall, P. W. S. and D. Cortis (2021). "Are Sports Bettors Biased toward Longshots, Favorites, or Both? A Literature Review." Risks **9**(1): 22.
Ottaviani, M. and P. N. Sørensen (2008). Chapter 6 - The Favorite-Longshot Bias: An Overview of the Main Explanations. Handbook of Sports and Lottery Markets. D. B. Hausch and W. T. Ziemba. San Diego, Elsevier**:** 83-101.
Peel, D., M. Cain and D. Law (2000). "The Favourite-Longshot Bias and Market Efficiency in UK Football Betting." Scottish Journal of Political Economy **47**: 25-36.
Shin, H. S. (1991). "Optimal Betting Odds Against Insider Traders." The Economic Journal **101**(408): 1179-1185.
Shin, H. S. (1992). "Prices of State Contingent Claims with Insider Traders, and the Favourite-Longshot Bias." The Economic Journal **102**(411): 426-435.
Shin, H. S. (1993). "Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims." The Economic Journal **103**(420): 1141-1153.
Snowberg, E. and J. Wolfers (2010). "Explaining the Favorite-Long Shot Bias: Is it Risk-Love or Misperceptions?" Journal of Political Economy **118**: 723-746.

Thaler, R. H. and W. T. Ziemba (1988). "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries." Journal of Economic Perspectives **2**(2): 161-174.

TopSportsbooks. (2022). "Sports and Odds Archives." Retrieved 5/18/2022, from https://sportsbookreviewsonline.com/scoresoddsarchives/scoresoddsarchives.htm.

Wikipedia. (2022). "Efficient-market hypothesis." from https://en.wikipedia.org/wiki/Efficient-market_hypothesis.