A Computational Analysis of Arbitrage Opportunities in Sports Gambling

Brian Matejek

Abstract

Arbitrage opportunities appear in sports gambling when bookkeepers disagree about the probabilities of different outcomes. Some bookkeepers actively try to prevent arbitrage gamblers from using their services. This paper addresses the various issues with arbitrage gambling by considering the problems with betting on individual games as well as betting on many games over a long period of time. I also evaluate the methods discussed in this paper against the industry standard and discuss the feasibility of arbitrage as a long-term investment strategy.

1. Introduction

Examining sports gambling methods that bookkeepers employ and that gamblers should use, this paper addresses the two major issues with sports gambling - micromanaging bets on individual games and macromanaging multiple financial accounts over the course of a few years. This paper then discusses arbitrage opportunities that arise when analyzing multiple bookkeepers because arbitrage is the only gambling strategy that may guarantee a profit for the gambler.

The goal of this independent work is to improve on the sports gambling arbitrage methods that are currently in existence. RebelBetting specializes in exploiting arbitrage opportunities and offers users an expected investment growth of 10% a month[4]. Sports Arbitrage World also claims that they can grow users accounts by 10% per month[5]. This project seeks to manage accounts with sixteen different bookkeepers and provide a better expected return than the other online arbitrage tools.

Bookkeepers have mixed thoughts on arbitrage. Some bookkeepers are fervently against arbitrage gamblers and actively seek to blacklist those gamblers from their bookkeeping services. These bookkeepers fear losing vast sums of money because of potential mistakes in their algorithms. With confidence in their probability models and in their ability to react to changing markets, other bookkeepers openly welcome arbitrage gamblers. Regardless of a bookkeeper's views on arbitrage,

arbitrage gamblers tend to correct probability errors in the market by forcing bookkeepers to change extreme odds.

The problems discussed in this paper are grouped into two large categories, focusing on micromanaging and macromanaging the arbitrage process. The micromanaging aspect of gambling is deterministic since all of the odds are known at the time of placing the bet. The macromanaging aspect of this project represents some online algorithms that attempt to estimate the correct long term strategy. Later in the paper I focus on a simulation that tests the various algorithms created in this paper.

The final goal of this paper is to address the issues with arbitrage gambling. I hope to analyze the logistics of arbitrage gambling as well as the claims of RebelBetting and Sports Arbitrage World. In the end, I hope to conclude whether or not arbitrage gambling is a realistic and sound gambling and long-term investment strategy.

2. Existing Research

There has not been much accessible research in the field of sports gambling and arbitrage. Sports gambling is illegal in many countries and states, and thus it doesn't attract considerable outside academic research. Most of the research pertinent to sports gambling comes from other fields. Many arbitrage gamblers employ minimax strategies to ensure an equal payoff regardless of the outcome of an event. Bookkeepers also use the minimax theorem for the same purpose. The minimax theorem first appeared in academics when John von Neumann devised a strategy for zero-sum games[9]. The theory was developed over the last half century and is now used in a variety of fields.

Crowd sourcing and wisdom of the crowd is one of the major theories cited in this paper. There has been substantial research into wisdom of the crowd and how it relates to predicting sporting events. Mathematician Brian Galebach ran a contest in 2003 asking each participant to assign probabilities for football teams to win various games[18]. Most participants did not predict the probabilities well at all. However, when averaging together all of the participants, the average predictor did astonishingly well. In 2003 ProbabilityFootball, the average of everyone's predictions would have finished 7^{th} out of 2,231 participants[18]. Even the very worst participants averaged together would have finished in 62^{nd} place[18].

Online gambling is a multi billion dollar industry[2], and attracts a lot of interest from various sources online. Thus, there is a significant amount of data on each particular bookkeeper, including several years of archived results on the website OddsPortal.com[13]. OddsPortal.com acts as a center for all sports gambling, providing various odds for several different bookkeepers and even indicating arbitrage opportunities that are currently available. However, OddsPortal.com does not manage accounts and does not claim that users will receive a certain growth rate on their investment.

One particular paper by Egon Franck, Erwin Verbeek, and Stephan Nüesch focused on *Inter-market Arbitrage in Sports Betting*. Their paper hoped to combine exchange rates with arbitrage opportunities to create more guaranteed bet scenarios[7]. However, the paper doesn't consider arbitrage gambling by itself citing previous research that dismissed the possibilities for traditional arbitrage opportunities. The previous research focused on three bookkeepers and one sport[7]. In this paper, I analyze 14 different bookkeepers from around the world in 13 different sports.

3. Web Scraping

In total, the simulation I created analyzes 1,966 teams in 90 sports leagues around the world. I created and ran various scripts that scraped data from OddsPortal.com for 105,065 games with dates ranging from Aug 25, 2006 to April 26, 2013. Each game required scraping a different web page for all of the individual closing odds offered by the different bookkeepers. Games ranged from having two to fourteen different odds all offered by different bookkeepers. I scraped data for 13 sports: baseball, basketball, soccer, hockey, American football, water polo, bandy, Australian rules football, handball, cricket, volleyball, rugby league, and rugby union.

I scraped all of the game results from OddsPortal. OddsPortal had the closing odds for many different games in convenient tables, but did not provide the opening odds as conveniently. This is discussed further in Section 9. However, the tables were generated using JavaScript that queried their database. `urllib` and other python libraries did not adequately scrape the data because these libraries do not wait for JavaScript to finish loading all the content. Thus, I had to write the scripts in JavaScript, using `phantomjs`, a “headless WebKit" that emulates all of the functionalities of a browser without the GUI[8]. `phantomjs` waited for the JavaScript to load before scraping the information, giving me access to all of the archived results I needed to create an adequate simulation of the past six years. However, `phantomjs` took between 5 and 7 seconds to load and scrape each individual web page. Thus, the CPU time required to handle all of the sporting events was somewhere in the range of 145 to 205 hours.

After scraping all of the web pages, I wrote multiple python scripts to analyze the collected data and put the data into an easy format for my simulation to read and process. The scripts totaled over 300 lines and created several large text documents that my simulation would read to create the proper environment to run.

4. Bookkeepers

There are many bookkeepers based in several countries. My simulation uses the data collected from 16 different bookkeepers, listed in Table 5 in Appendix A. Some bookkeepers provide maximum bets and maximum winnings to protect themselves from last minute large bets that could threaten their probability models and minimax strategies, as discussed in Section 5. Some bookkeepers have a maximum total amount of losses that they allow any one gambler to sustain over a period of time. These rules are in place to protect those with gambling addictions who would otherwise waste all of their money (as well as trying to limit arbitrage gamblers who are offsetting huge losses with another bookkeeper).

Each bookkeeper has a different outlook on arbitrage. Some bookkeepers, such as Pinnacle-Sports, openly welcome arbitrage gamblers, stating on their website, “In theory all bookmakers shouldn’t care about the motivation for placing a bet, but should simply look to balance the bet volume...[Others’] limiting of arbitrage players is a reflection of a bookmaker’s short-comings, such as posting ‘bad odds’, or an inability to move odds fast enough to avoid being the focus of arbitrage players."[19] Other bookkeepers, such as Ladbrokes, actively seek to remove arbitrage gamblers, and thus arbitrage gamblers need to disguise their betting patterns[24]. Bet-at-Home and BetVictor are two other bookkeepers that try to blacklist arbitrage gamblers since they react slowly to changing odds in the market place[6][23].

Some bookkeepers have maximum bets that they place on gamblers to protect themselves from having to payout too much in any given day. 10Bet has a maximum bet of $1,000, so while they are valuable early on in my simulation, they soon becoming the limiting factor in the total amount I could bet on any individual game[3][10]. myBet has a maximum bet of $1,400 so likewise it is useful early on but gradually begins to limit the amount that I could bet on any one arbitrage opportunity[11]. Some other bookkeepers have maximum bets that are large enough to not affect the simulation, such as NordicBet, which refuses to pay more than $31,000 to any one person on any one game[12].

Some of the bookkeepers are not particularly useful in arbitrage gambling, since they tend not to produce many odds that can be used in arbitrage. Unibet offered over 80,000 odds between the seven year period but only contributed to roughly 1,000 arbitrage opportunities. TonyBet is a more recent bookkeeper and offered very view odds during the course of the simulation[16].

5. Micromanaging Bets

There are two major components to the analysis of sports gambling and in particular arbitrage opportunities. The first component focuses on the micromanagement aspects of gambling, or how to best bet on individual games. The second component focuses on the macromanagement aspects of gambling, or how to manage finances when pursuing arbitrage opportunities. Both micro and macromanagement create difficulties when creating a simulation with a gambling algorithm. This next section discusses the micromanagement issues of arbitrage gambling, first focusing on minimax strategies, then how bookkeepers determine odds, and lastly different techniques on how to gamble per individual game.

Most of the micromanagement aspects of gambling stem from variations of the minimax problem. A bookkeeper and a good gambler both want to reduce their risk of losing a lot of money. To this end, good gamblers try to maximize their minimum payoff and good bookkeepers minimize their maximum payout so that regardless of the outcome they receive or pay the same amount of money. Bookkeepers and arbitrage gamblers do not want to rely on emotions or luck, and want to remain indifferent about the outcomes of any given game. This allows arbitrage gamblers to bet smartly without considering personal fan loyalties or feelings.

Consider matches that have only two outcomes, either the home team wins or the away team wins. o_H and o_A are the payoffs if the home or away teams wins, in European decimal format. European decimal odds show the amount the bookkeeper will pay if a given team wins. For example, if the decimal odds are 2.00, and the gambler bets \$100 on that outcome, the bookkeeper will pay the gambler \$200. This represents paying back the initial bet and the winnings. b_H and b_A are the bets made on the home and away teams. C is the total amount that the algorithm will spend on a given game. The payoff if the home team wins is $o_H * b_H$ and $o_A * b_A$ if the away team wins.

Lemma 1. *One can maximize the minimum payoff given any outcome by setting* $o_H * b_H = o_A * b_A$.

Proof. Consider any change δ to b_H or b_A. The new amount bet on the opposing team is $b_A - \delta$ or $b_H - \delta$ respectively. Since both of these values are less than the previous amount bet, and the payoff does not change, the amount won from an away or home win decreases if the amount bet on the home or away team increases. $\square$

By Lemma 1, to find the home and away bets to maximize the minimum payoff:

$$
\begin{aligned}
o_H * b_H &= o_A * b_A & \qquad b_H + b_A &= C \\
b_H + \frac{o_H * b_H}{o_A} &= C & \qquad b_A + \frac{o_A * b_A}{o_H} &= C \\
b_H &= \left(\frac{o_A}{o_H + o_A}\right) C & \qquad b_A &= \left(\frac{o_H}{o_H + o_A}\right) C
\end{aligned}
$$

Consider matches that have three outcomes, either the home team wins, or the away team wins, or there is a draw. There is the same notation as before but o_X is the payoff of a draw in European decimal format. b_X is the bet made on the outcome of the two teams drawing. The payoff if the teams draw is $o_X * b_X$.

Corollary 2. *One can maximize the minimum payoff given any outcome by setting* $o_H * b_H = o_A * b_A = o_X * b_X$.

Proof. The argument follows exactly as above, but either other outcome (or both) may change by δ_1 or δ_2. □

To find the home, draw, and away bets to maximize the minimum payoff:

$$\begin{array}{lll} o_H * b_H = o_A * b_A & o_A * b_A = o_X * b_X & o_X * b_X = o_H * b_H \\ b_H = \dfrac{b_A * o_A}{o_H} & b_A = \dfrac{b_X * o_X}{o_A} & b_X = b_X \\ b_H = \dfrac{b_X * o_X * o_A}{o_H * o_A} & & \end{array}$$

$$b_H + b_A + b_X = C$$

$$\frac{b_X * o_X}{o_H} + \frac{b_X * o_X}{o_A} + b_X = C$$

$$b_X = C\left(\frac{o_H * o_A}{o_H * o_A + o_X * o_A + o_X * o_H}\right)$$

$$b_A = C\left(\frac{o_H * o_X}{o_H * o_A + o_X * o_A + o_X * o_H}\right)$$

$$b_H = C\left(\frac{o_X * o_A}{o_H * o_A + o_X * o_A + o_X * o_H}\right)$$

Lemma 3. *The bookkeeper will try to present a set of closing odds where payouts will be equal regardless of outcome.*

Analysis. Bookkeepers will use the minimax strategy since they want to remain indifferent between the two outcomes in a sporting event. Bookkeepers assign probabilities to teams winning and do not want to concern themselves with the actual outcome of a given game. In sporting events, there is always a chance that the underdog will win, no matter what two teams are matching up against each other. The bookkeeper wants to protect himself from this risk by utilizing the minimax theorem so that regardless of outcome, he receives the same level of profit. If the payouts were not equal for a game, the bookkeeper could find himself paying out to a large number of people a sum that exceeds the amount of money he took in from the losing betters. This could end a bookkeeper's career in short order if the amount paid out is too large. □

Lemma 4. *In a fair bet, the bookkeeper will assign payoffs based on the probabilities that they assume to be correct.*

Analysis. There are two cases: one with two outcomes and one with three outcomes.

Case 1 - Two Outcomes: Suppose the bookkeeper thinks the home team will win with probability p'_H and the away team will win with probability p'_A. The bookkeeper wants the expected payoff to be the same regardless of the outcome by Lemma 3. Thus, the bookkeeper will pay:

$$p'_H * o_H = p'_A * o_A$$

where o_H and o_A are the fair bet payoffs given by the bookkeeper. Also, $p'_H * o_H + p'_A * o_A = 2$, since this is a fair bet.

$$2 = 2 * p'_H * o_H$$

$$o_H = \frac{1}{p'_H} \qquad o_A = \frac{1}{p'_A}$$

Regardless of who wins, the expected payoff for any bet that the bookkeeper receives is:

$$E[\text{Home Win}] = p'_H * \frac{1}{p'_H} = 1 \qquad E[\text{Away Win}] = p'_A * \frac{1}{p'_A} = 1$$

Case 2 - Three Outcomes: Suppose the bookkeeper thinks the home team will win with probability p'_H, the away team with probabilitiy p'_A, and there will be a draw with probability p'_X. The bookkeeper wants the expected payoff to be the same regardless of the outcome for each dollar bet by Lemma 3. Thus, the bookkeeper will pa:

$$p'_H * o_H = p'_X * o_X = p'_A * o_A$$

where o_H, o_X, and o_A are the payoffs for each outcome possible. Also, $p'_H * o_H + p'_X * o_X + p'_A * o_A = 3$, since this is a fair bet.

$$3 = 3 * p'_H * o_H$$

$$o_H = \frac{1}{p'_H} \qquad o_X = \frac{1}{p'_X} \qquad o_A = \frac{1}{p'_A}$$

Regardless of who wins, the expected payoff is equal and is:

$$E[\text{Home Win}] = p'_H * \frac{1}{p'_H} = 1 \qquad E[\text{Draw}] = p'_X * \frac{1}{p'_X} = 1 \qquad E[\text{Away Win}] = p'_A * \frac{1}{p'_A} = 1$$

In this situation, the bookkeepers want to minimize the maximum expected value of paying (from Lemma 3), and this occurs when all of the values are equal. □

A gambler will make his own probability for each outcome before betting. Assume the gambler believes the probabilities of the home team winning, away team winning, and a draw are p_H, p_A, and p_X respectively. In a fair bet, the opening odds of the bookkeeper represent the probabilities that the bookkeeper believes each team has of winning or drawing. Label these probabilities p'_H, p'_A, and p'_X. If the bookkeeper and the gambler disagree on these probabilities, the gambler will bet on the outcome which deviates most from the gambler. Since the bookkeeper's payoffs are correlated with the probabilities, the gambler will not bet on outcomes that he believes to be more likely than the bookkeeper. Consider the following expected return functions for the gambler:

$$E[\text{Bet H}] = \frac{p_H}{p'_H} \qquad E[\text{Bet X}] = \frac{p_X}{p'_X} \qquad E[\text{Bet A}] = \frac{p_A}{p'_A}$$

The gambler's highest expected return comes from betting on the outcomes that he believes are more likely than the bookkeeper since $\frac{p_i}{p'_i}$ will be greater than 1 so his expected return is greater than

1. As gamblers begin betting on these games, the bookkeeper can begin to judge how the public values the probabilities of the different outcomes. By Lemma 3, the bookkeeper's end goal is to pay the same amount of money regardless of the outcome in the game. If a disproportionate number of gamblers bet for one outcome, the bookkeeper will reduce the payoff for that outcome and raise the payoffs for the other outcomes so that the bookkeeper can return to the equilibrium of paying an equal amount no matter the outcome.

Lemma 5. *The closing odds represent a weighted average of the probabilities that many gamblers and the bookkeeper have made on a given team winning.*

Analysis. Since the bookkeepers want to pay out the same amount regardless of the outcome, they will adjust their odds before the game starts to reflect the trends of the gamblers. Thus, the probabilities of the gamblers are weighted into the value of the closing odds. By wisdom of the crowd, this adjusted probability is most likely better than the probability originally set by the bookkeeper. Even a weighting of the worst probability pickers produces a decent estimate for the actual probability of a game, as shown in ProbabilitiySports[18]. □

No bookkeeper would present a set of odds that allow for arbitrage using only his odds, since be would be guaranteed to lose money. Arbitrage opportunities come when different bookkeepers disagree on the probabilities of different outcomes and a gambler can hedge his bets to guarantee making money. Bookkeepers also do not provide fair bet odds since they would make no money in the best case (where they correctly provide odds that encourage gambling in equal payoffs per outcome). Instead, bookkeepers charge an amount for gambling with them, also known as vigorish. Vigorish is the derivation from a fair bet. First, consider the case where there are two outcomes, the home team wins or the away team wins. By Lemma 1, the gambler can maximize his minimum guaranteed payoff with bets b_H, b_A:

$$b_H = \frac{o_A}{o_H + o_A} \qquad b_A = \frac{o_H}{o_H + o_A}$$

If a gambler bets on these outcomes, the amount he will win regardless of the outcome of the match is:

$$b_H * o_H = b_A * o_A = \frac{o_H * o_A}{o_H + o_A}$$

If the bet were a fair bet, the gambler would make the amount that he bet. However, the bookkeeper charges the gambler the difference between a fair bet and the actual amount offered:

$$1 - \frac{o_H * o_A}{o_H + o_A}$$

Next consider the case where there are three outcomes, where the home team can win, the away team can win, or the teams can draw. By Corollary 2, the gambler can maximize his minimum guaranteed payoff with bets b_H, b_X, b_A:

$$b_H = \frac{o_X * o_A}{o_H * o_A + o_X * o_A + o_X * o_H} \qquad b_X = \frac{o_H * o_A}{o_H * o_A + o_X * o_A + o_X * o_H}$$

$$b_A = \frac{o_H * o_X}{o_H * o_A + o_X * o_A + o_X * o_H}$$

If a gambler bets on these outcomes, the amount he will win regardless of the outcome of the match is:

$$b_H * o_H = b_X * o_X = b_A * o_A = \frac{o_H * o_X * o_A}{o_H * o_A + o_X * o_A + o_X * o_H}$$

If the bet were a fair bet, the gambler would make the amount that he bet. However, the bookkeeper charges the gambler the amount of the vigorish:

$$1 - \frac{o_H * o_X * o_A}{o_H * o_A + o_X * o_A + o_X * o_H}$$

Another way to think about vigorish percentage is that $100 - v$ is the maximum percent return one can guarantee by only betting on this bookkeeper. This is inversely related with the maximum arbitrage opportunity given one bookkeeper. One bookkeeper would never present a set of odds where a gambler could hedge his bet and be guaranteed to make money, since the bookkeeper would then have a negative vigorish. Lemma 3 would fail and the bookkeeper would rely on chance to remain in business.

Theoretically, bookkeepers come up with fair odds for a game where each payoff represents a probability of a team winning. The bookkeeper then decreases the payoffs until he achieves the desired vigorish. Although bookkeepers do not publish how they calculate the modified odds from their original probability functions, they most likely divide the vigorish proportionately to the probabilities of the different outcomes. If the bookkeepers did not calculate the new odds in this fashion, one of the odds relative to the other would be larger than the corresponding probability ratio.

Theorem 6. *Bookkeepers reduce the odds from the fair bet probabilities proportionately from the original fair bet odds.*

Discussion. Suppose the bookkeepers believe the home team will win with probability p_H, the away team with probability p_A, and the teams draw with probability p_X. They expect people to bet such that the payoffs for each outcome remain relatively equal. If there is a great disparity in the amount of bets that one outcome receives, the bookkeeper will adjust his odds. For each outcome, the bookkeeper expects to pay the probability of the team winning times the payoff for that game. The payoffs were discussed in Lemma 4:

$$E[\text{Home Win}] = p_H * \left(\frac{1}{P_H} - v_H\right) \qquad E[\text{Draw}] = p_X * \left(\frac{1}{P_X} - v_X\right)$$

$$E[\text{Away Win}] = p_A * \left(\frac{1}{P_A} - v_A\right)$$

where v_H, v_X and v_A are the amount by which the bookkeeper reduces those odds in order to guarantee a profit. If the vigorish is not taken evenly from the three probabilities, the odds given to the gamblers will not reflect the probabilities that the bookkeeper believes each team has of winning. The probabilities minus the vigorish will not produce equal expected returns (as the bookkeeper wants) since there is a higher expected payoff for one of the outcomes - the one that is lessened by the vigorish disproportionately. To make all of the expected payouts equal, the bookkeeper wants:

$$E[\text{Home Win}] = p_H * \left(\frac{1}{p_H}\right)(1-v) \qquad E[\text{Draw}] = p_X * \left(\frac{1}{p_X}\right)(1-v)$$

$$E[\text{Away Win}] = p_A * \left(\frac{1}{p_A}\right)(1-v)$$

These values are all equal to $(1-v)$, and the corresponding v_H, v_X, and v_A values are all proportional to the odds given by the bookkeeper. $\square$

However, the only reason that arbitrage exists is because there is an outlier among the bookkeeper payoffs. Bookkeepers are in the business of creating probabilities for sports games, and thus the average of all the bookkeepers is a better estimate for the actual probability of each team winning. Thus, it is better, when analyzing the expected return function, to think about the probability of each team winning as the average of all the bookkeeper probabilities. By Lemma 6, the bookkeepers determine the probabilities of each team winning and then take cuts proportional to the odds they offer. Thus, it is easy to construct the probabilities that the bookkeepers believe each outcome has of occurring. First, where there are two outcomes, define o_H^* and o_A^* as the fair bet odds a bookkeeper would give:

$$o_H^* = \frac{1}{p_H} \qquad p_H = \frac{1}{o_H^*} \qquad\qquad o_A^* = \frac{1}{p_A} \qquad p_A = \frac{1}{o_A^*}$$

$$o_H^* = \frac{o_H}{1-v} = \frac{o_H}{\frac{o_H * o_A}{o_H + o_A}} = \frac{o_H + o_A}{o_A} \qquad o_A^* = \frac{o_A}{1-v} = \frac{o_A}{\frac{o_H * o_A}{o_H + o + A}} = \frac{o_H + o_A}{o_H}$$

$$p_H = \frac{o_A}{o_H + o_A} \qquad p_A = \frac{o_H}{o_H + o_A}$$

For three outcomes, first again define o_H^*, o_X^*, and o_A^* as the fair bets a bookkeeper would give:

$$o_H^* = \frac{1}{p_H} \quad p_H = \frac{1}{o_H^*} \qquad o_X^* = \frac{1}{p_X} \quad p_X = \frac{1}{o_X^*} \qquad o_A^* = \frac{1}{p_A} \quad p_A = \frac{1}{o_A^*}$$

$$o_H^* = \frac{o_H}{1-v} = \frac{o_H}{\frac{o_H * o_X * o_A}{o_H * o_A + o_X * o_A + o_X * o_H}} = \frac{o_H * o_A + o_X * o_A + o_X * o_H}{o_X * o_A}$$

$$o_X^* = \frac{o_X}{1-v} = \frac{o_X}{\frac{o_H*o_X*o_A}{o_H*o_A+o_X*o_A+o_X*o_H}} = \frac{o_H*o_A+o_X*o_A+o_X*o_H}{o_H*o_A}$$

$$o_A^* = \frac{o_A}{1-v} = \frac{o_A}{\frac{o_H*o_X*o_A}{o_H*o_A+o_X*o_A+o_X*o_H}} = \frac{o_H*o_A+o_X*o_A+o_X*o_H}{o_H*o_X}$$

$$p_H = \frac{o_X*o_A}{o_H*o_A+o_X*o_A+o_X*o_H} \qquad p_X = \frac{o_H*o_A}{o_H*o_A+o_X*o_A+o_X*o_H}$$

$$p_A = \frac{o_H*o_X}{o_H*o_A+o_X*o_A+o_X*o_H}$$

Next, consider an instance where there are arbitrage opportunities, where a gambler can hedge his bets over multiple bookkeepers and guarantee making a profit. On November 26, 2009, the Denver Broncos hosted the New York Giants. The Denver Broncos entered the game with 7 wins and 4 losses and the New York Giants had 6 wins and 5 losses. The Broncos won the game 26 to 6. The following bookmakers offered these odds in decimal format (Table 1). These odds were translated into probabilities in the same table.

Bookkeeper	Home Payoff	Away Payoff	Vigorish	Home Prob.	Away Prob.
10Bet	2.86	1.45	3.8%	.664	.336
bet-at-home	2.85	1.40	6.1%	.671	.329
bet365	2.80	1.44	4.9%	.660	.340
BetVictor	2.75	1.40	7.2%	.663	.337
bwin	2.95	1.42	4.1%	.675	.325
Ladbrokes	3.00	1.42	3.6%	.679	.321
myBet	3.50	1.30	5.2%	.729	.271
NordicBet	3.05	1.40	4%	.685	.315
Paddy Power	3.10	1.40	3.6%	.689	.311
Pinnacle Sports	2.94	1.45	2.9%	.670	.330
Unibet	2.95	1.40	5.1%	.678	.322
William Hill	2.75	1.48	3.8%	.650	.350

Table 1: Payoffs provided by various bookkeepers for the Denver Broncos, New York Giants football game with the corresponding probabilities.

The probability of the each team winning can be represented as a random variable and all of these bookkeepers are offering their best estimates of this random variable. The average probability offered by these bookkeepers is .676 for the Denver Broncos winning and .324 for the New York Giants winning. The standard deviations are both .020 (since all of the averages are related, summing to 1). Figure 1 is a diagram of the probabilities offered by the bookkeepers. The blue vertical line is the average probability of all the bookkeepers and the green vertical lines represent standard deviations from the mean. The black rings represent the various probabilities that the bookkeepers presented through their closing odds, and the gold rings are the optimal arbitrage opportunities, assuming a minimax strategy.

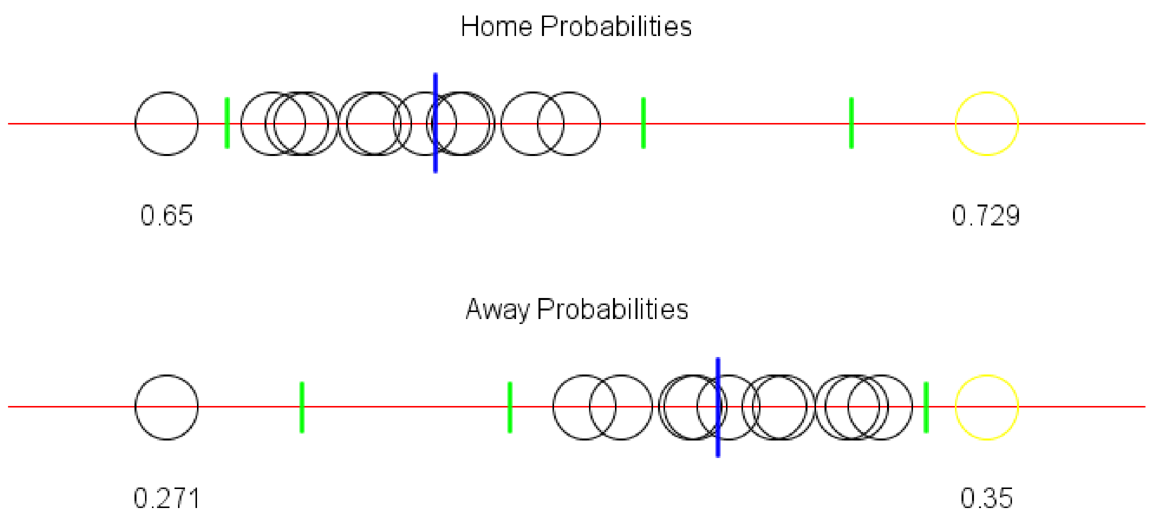


Figure 1: A graph showing the various probabilities different bookkeepers found for the Denver Broncos versus New York Giants football game.

This figure shows that the odds arbitrage gamblers bet on are on the fringe among all of the bookkeepers. If the odd creates an arbitrage opportunity, it is definitely greatly deviated from the average, since if all of the bookkeepers agreed, there would be very little variation in the odds (the only variation would come from the vigorish). myBet is over two standard deviations away from the average probability among the bookkeepers for the Broncos winning, and William Hill is over one standard deviation away from the average probability for the Giants winning. Even though William Hill does not believe the Giants are going to win (the probability is still 0.35), the probability is enough greater than the other bookkeepers to create an arbitrage opportunity.

Although the majority of the time my algorithm uses the standard minimax method, I occasionally focus more on the probabilities. First, define p_H, p_X, p_A for the probabilities of the home team winning, drawing, and losing as determined by averaging the probabilities given the odds. Next, bet

$p_H * C$ on the home team, $p_X * C$ on the team drawing, and $p_A * C$ on the away team winning. For the following analysis, all bets are normalized to sum to 1. The expected value for this bet is:

$$E[x] = p_H^2 * o_H + p_X^2 * o_X + p_A^2 * o_A$$

and the guaranteed bet from the minimax method is:

$$E[x] = \frac{o_H * o_A * o_X}{o_H * o_X + o_H * o_A + o_X * o_A}$$

I only use the new method when $p_H * o_H > 1$, $p_A * o_A > 1$, and $p_X * o_X > 1$ and the expected value of the first equation is larger than the expected value of the second equation. Subtracting the two expected values and substituting in the equations $p_H * o_H = 1 + \delta$, $p_A * o_A = 1 + \varepsilon$, $p_X * o_X = 1 + \gamma$, and $\frac{o_H * o_X * o_A}{o_H * o_A + o_H * o_X + o_X * o_A} = 1 + \alpha$:

$$E[\text{diff}] = (1+\delta) * p_H + (1+\varepsilon) * p_A + (1+\gamma) * p_X - (1+\alpha)$$

$$E[\text{diff}] = \delta * p_H + \varepsilon * p_A + \gamma * p_X - \alpha$$

When running the simulation, there were 967 instances when the constraints $p_H * o_H > 1, p_X * o_X > 1, p_A * o_A > 1$ were satisfied. In every instance, the expected value of the difference was positive. The average difference of the expected values over the 967 instances was 0.139 (or nearly 14% greater expected value) with a standard deviation of 0.557.

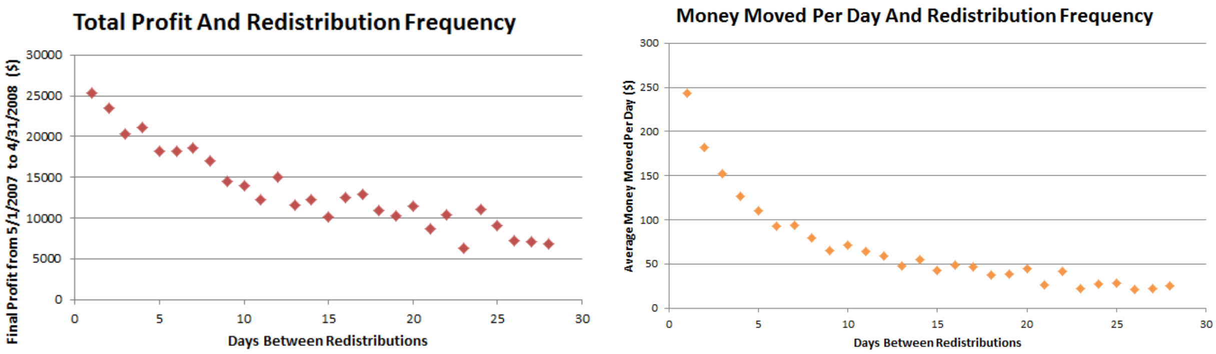
The other big issue with individual bets is how much to bet given the different account amounts with each bookkeeper. Obviously, no account may go below zero for any bookkeeper, since this signifies inadequate funding availability. One way to avoid this scenario is to always determine the amount bet by ascertaining the account with the lowest balance. Since an arbitrage bet is by definition guaranteed to make money, the highest possible return would come from betting the entire account amount. For a few reasons specified in the next section, this is not a viable option. The actual determination of the amount becomes a macromanagement problem since betting a lot on one game will significantly reduce the ability to bet on the next set of outcomes. Although the best return comes from betting all of a gambler's money on every arbitrage opportunity, it is simply not an option.

6. Macromanaging Accounts

One of the major issues with managing multiple accounts for many different bookkeepers is keeping track of where to put all of the money. Many of these bookkeepers use third party banking and other transaction services to allow customers to quickly withdraw money. With the free market, the time of withdrawal and the transaction fee will converge to instantaneous and 0%. Once one bookkeeper allows a free service, all of the bookkeepers will be forced to follow suit as gamblers switch their accounts to utilize the free services. Already, bookkeepers have switched to these free services to encourage gamblers to bet with them.[23] However, in my simulation I assumed that transactions would take at least one night to process and charge a 2% transaction fee for all money transferred. This seemed like a reasonable conservative estimate to a complex real world financial system.

One of the big problems with arbitrage gambling is ensuring that all of the bookkeeper accounts have money before making a bet, and that losing the bet will not put any account underwater. Bookkeepers will quickly blacklist any gambler who they believe cannot pay his debts. Originally I was going to double the previous deposit amount to ensure a constant competitive ratio with the total amount ever needed to deposit into an account (a ratio of 2). However, if an account nearly hits 0, it requires more than $200, $400, etc. to keep functioning. If every account has $20,000, except for one which has $400, the betting amount is greatly limited. A few accounts could limit the amount of money bet every single game. Thus, I implemented a procedure of redistributing the money in all of the accounts at regular intervals. This guaranteed that no account remained nearly drained for extended periods of times. This greatly increased the algorithm's ability to increase bets as the total amount won accumulated.

There was no simple algorithm that would correctly predict the best time to redistribute the accounts. The trouble is that future outcomes and arbitrage opportunities are simply not known.



(a) Negative correlation between final profit and number of days between redistributing accounts.

(b) Negative correlation between the amount of money moved daily and the number of days between redistribution.

Figure 2: All data between the dates of May 1, 2007 to April 30, 2008.

The three largest accounts might all belong to an arbitrage opportunity in two days, but I might even out all of the account balances that night. The goal is to keep each bookkeeper's account high enough that the account will not severely limit the amount one can bet on the arbitrage opportunity. Figure 2a shows various timescales for redistributing the accounts when gambling on the interval from May 1, 2007 to April 31, 2008. The more often the algorithm redistributes the better. However, part of the decision process must include the thought that bookkeepers will not accept account withdrawals every day. Thus, the algorithm readjusts accounts every 4 days. Figure 2b shows the average amount of money transfered between the accounts depending on the number of days before redistribution.

The amount of money transferred decreases over time because with time account balances will naturally converge. I am not going to win every arbitrage opportunity that I bet on with a given bookkeeper. Over time, the wins and losses of each bookkeeper will become roughly equal, since it is mostly random with which bookkeeper I win. Although the bookkeepers are presenting disproportionate probabilities of a given team winning, both bookkeepers are extreme in opposite ways, and both are nearly equally likely to be wrong, except in extreme circumstances. However, my macromanagement model evens out the accounts rather than wait for the accounts to naturally converge. The amount transfered is not that great even redistributing every 4 days, and it averages to less than $10 per bookkeeper per day (although this is an amortized cost that is accrued on every fourth day).

As stated in the previous section, betting every dollar on an arbitrage opportunity would yield the best returns. Even if an account only had a fraction of the total reserves across all bookkeepers, by betting the maximum a gambler is ensuring the greatest return. However, betting the entire amount in an account is a guaranteed way to alert bookkeepers that the gambler is engaging in arbitrage practices. Although putting the entire account on one team beating another in one game might be the best financial option in an arbitrage setting, it alerts bookkeepers that the gambler is leveraging the outcome by betting on another bookkeeper and guaranteeing a positive outcome. Thus, my algorithm always bets a percentage of the entire account. The amount bet during a game decreases if the account sinks below the initial $100 investment. The amount bet decreases then to ensure the account remains above zero and that the account can still be used if another arbitrage opportunity presents itself before the next redistribution of accounts.

Although I do not want to bet the entire account on one game, the goal should be for one account to be close to zero before redistributing the accounts. If one account is close to zero, I could not have bet much more on the series of games. There is no set number about how much is appropriate to bet on a given game, but I analyzed the data from May 1, 2007 to April 31, 2008 and have plotted the data of changing the percent of an account bet on a 2 outcome game in Figure 3 and the percent of a bet on a 3 outcome game in Figure 4 with the total spent. Looking at these figures, it appears that the percent bet on two outcome games does not greatly affect the overall money made during the course of this year. However, the percent bet on three outcome games does affect the overall money made during the simulation. Although it is better in terms of payoff to bet the entire amount, there are diminishing returns seen in the graph. It appears that around 60 - 70 % the second derivative becomes negative. I decided to bet 67% of the total amount in the account on any one game (up to the maximums constrained by the bookkeepers), since it was a reasonable trade off between making more money and still avoiding the turbulence of betting 100% of the account every arbitrage opportunity. If an account gets close to 0, I switch over and bet only a small amount of money, holding out until the next account redistribution.

Some bookkeepers blacklist gamblers that they believe are betting solely on arbitrage opportuni-

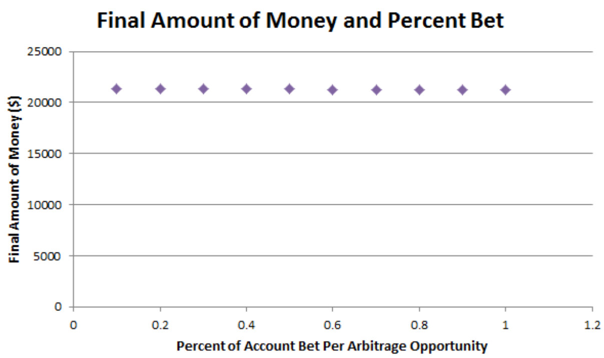


Figure 3: The ending amount of money depending on the percent of an account bet (for 2 outcome games).

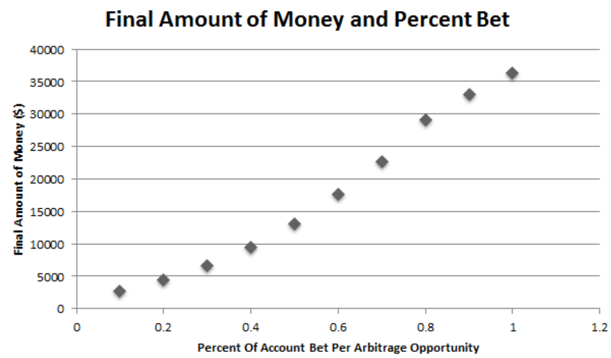


Figure 4: The ending amount of money depending on the percent of an account bet (for 3 outcome games).

ties. If a bookkeeper sees an account that has lost several thousand dollars very quickly, they will either assume the individual is arbitrage gambling and thus is offsetting the losses with another bookkeeper or that gambler has a serious problem. Either way, the bookkeeper will most likely blacklist the gambler and force him to take his money elsewhere. In order to prevent this from happening, my algorithm limits the amount that I can lose with any one bookkeeper in a given period of time. The trouble with arbitrage gambling is that there are no obvious bets (especially considering that the bookkeepers are really producing probabilities of each team winning). There is no guaranteed method for increasing an account without some risk. Thus, before I become blacklisted, I stop all betting with a given bookkeeper after having a net loss of $10,000. I ban trading with the bookkeeper for 6 months the first time, 12 months the second time, and 18 months after that, etc. I only consider the net gains and/or losses between bans so that one loss after a 6 month ban does not lead to 12 month ban. The goal is to amortize losses over an extended period of time. Starting with $100, the following bookkeepers in Table 2 were banned when running the simulation on the data before 12/31/2008. Around half of the bookkeepers were banned at some point during the simulation. The algorithm banned the first bookkeeper after 617 days, and adding in 183 days, at the end of the ban, it appears that the accounts had an amortized loss of slightly more than $12.50 per day.

Bookkeeper	Date Banned
William Hill	Day 617
bwin	Day 618
Unibet	Day 623
Paddy Power	Day 630
NordicBet	Day 637
bet365	Day 640
Pinnacle Sports	Day 644

Table 2: This shows the bookkeepers banned during the simulation that included all games before 12/31/2008 along with the date first banned.

7. Simulation and Design Decisions

I began writing my simulation in Python. However, I quickly switched to Java because I needed a language that clearly differentiated between types. This became incredibly clear when writing Python code that parsed my data and determined which team won. In Python, strings can be compared with equals, less than, and greater than signs. When comparing the outcomes of games to determine the winner, I forgot to cast the variables to integers, and so games with an outcome "10:5" would compare the strings 10 and 5 and find 10 to be smaller. This skewed some of my early data until I figured out the error. After finding out what was happening, I switched over to Java where the compiler would not allow me to make such mistakes.

In total, my simulation is roughly 3,000 lines of Java code and has classes representing individual teams, bookkeepers, odds, matches, dates, as well as the simulation itself. The simulation implements all of the features specified in the previous sections and in addition also has the ability to toggle features to more easily compare my simulation model with existing models. I used `StdIn.java` and `StdDraw.java` produced by Kevin Wayne and Robert Sedgewick to help with the file reading and the drawing of some of the graphs[20][21]. All of the other code is my original source code.

The simulation can take in no parameters in which case it simulates all of the data that I have collected. If two arguments are given, the simulation restricts betting on games only in that time period, but still analyzes all games before the starting date to create good starting data on which to assess games and bets. Some additional arguments are used to help create the graphs in the analysis section.

My program simulated an actual financial environment with multiple bookkeepers. It did not bet any real money or interact in any other way with the real online sports gambling market. Sports gambling is not legal in New Jersey and the simulation just emulated the market without betting with the real market.

8. Analysis

There are a few websites devoted to finding and exploiting arbitrage opportunities for users. One such website is RebelBetting, which claims that the investments of users can grow 10% per month. My algorithm starts with \$100 in the accounts for every bookkeeper. Assuming that my algorithm runs for roughly 36 months, with a goal of 10% growth per month, the final outcome that RebelBetting predicts its users will achieve is:

$$A = P(1.1)^t$$

where A is the account balance, P is in the principal, and t is the time in months.

$$A = (\$1,400) * (1.1)^{36} = \$43,277.75$$

RebelBetting also charges users a monthly fee of £129 (≈ \$200.38)[4]. Over this 36 month period, this turns out to be around 17% of the total amount made and thus the final expected return when using RebelBetting is:

$$A = \$36,064.07$$

Sports Arbitrage World also claims to offer users an investment growth of 10% per month[5]. Sports Arbitrage World charges a monthly fee of £196 (≈ \$304.74). Over this 36 month period, users of

Sports Arbitrage World would receive:

$$A = \$32,307.89$$

For my analysis, I decided to save the data from January 1, 2009 to December 31, 2011. I choose this three year period because it was long enough to allow my algorithm to test the bookkeeper blacklisting features as well as the effects of maximum betting. I did not want to go from January 1, 2010 to December 31, 2012 since many of the games from the 2012/2013 seasons were not included because the season had not yet finished. Thus, the seasons that ended in 2012 were included when I scraped the archived results from OddsPortal.com but the seasons that began in 2012 and end in 2013 were mostly not included. Using these dates, my simulation began with \$1,400 (since two bookkeepers did not present any arbitrage opportunities in this range and there is no reason to start an account with a bookkeeper until the first arbitrage opportunity) and ended with \$82,725.01. Bookkeepers should make it easy and quick to open accounts since they want to encourage gambling with them. This value includes \$114,533.95 from arbitrage gambling minus \$31,808.94 for transaction fees that come with redistributing the account balances between all of these bookkeepers. Solving for x in the following equation gives:

$$82,725.01 = 1400(1+x)^{36}$$

$$x = 0.120$$

This is an increase of 20% over the traditional methods of RebelBetting and Sports Arbitrage World. This increase represents the techniques developed in both the micromanagement and macromanagement sections. By using the standard minimax model and keeping my macromanagement aspects to my algorithm, an arbitrage gambler would receive a net profit of \$41,497.96, winning \$59,659.82 through arbitrage gambling and paying \$18,161.87 in transaction fees. The micromanagement model produced in the above section outperforms the minimax model currently used by most arbitrage websites by nearly 2 to 1 during this three year period. This entire increase occurs in only a handful of games where the expected value is much higher than the standard minimax method.

Theoretically, if gamblers could move all of their money without transaction costs and bank delays, they could put the entire savings into an arbitrage opportunity and be guaranteed to increase their savings. They could bet their entire savings on every arbitrage opportunity. Out of the 105,065 games considered over the course of 7 years, over 5000 provided arbitrage opportunities. The average arbitrage opportunity yielded a guaranteed return of 4.2%. If a gambler started with $1 and was able to bet on every arbitrage opportunity, he would finish the time period with $2.18e+89. Even if the gambler could only bet on one arbitrage opportunity for every ten, he would end with $858,735,628.16 This would quickly bankrupt the gambling industry because with such large numbers it would be near impossible for bookkeepers to balance their expected payoffs. Thus, bookkeepers employ multiple methods in an effort to restrain the amount of money any individual can bet with them over a given period of time.

Many bookkeepers have maximum bets and also cut off gamblers who are too aggressive. My simulation had to keep track of these maximum bets and had to assure that I never exceeded the limits. These maximum bets significantly impair the ability of arbitrage gamblers to make significant amounts of money. I ran my simulation from January 1, 2009 to December 31, 2011 (36 months) with varying amounts of starting money per bookkeeper account. The results are summarized in Table 3 and in Figure 5. The second derivative of the curve that fits this data is negative, meaning that the rate of return on the money over the course of the simulation is decreasing. Although the simulation finishes with more money when starting with more money, the amount of profit is a less significant percentage than when starting with a smaller initial sum. This is shown in Figure 6, which maps the amount of money started to the percentage of money won over the course of the simulation. This value is calculated from the following equations:

$$A = P(1+x)^{36}$$

Start Money	End Money	Monthly Inc.	End Money (w/o Max)	Monthly Inc. (w/o Max)
10	$9,232.42	0.123	$9,232.42	0.123
20	$18,939.59	0.124	$18,464.84	0.123
40	$37,881.87	0.124	$36,929.68	0.123
80	$70,372.85	0.122	$73.859.36	0.123
160	$130,335.31	0.119	$147,718.72	.0123
320	$240,066.77	0.117	$270,545.79	.0121
640	$333,587.01	0.110	$418,180.79	0.113
1000	$474,773.91	0.103	$418,180.79	0.113
1280	$615,172.97	0.103	$681,845.82	0.106
2560	$972,452.26	0.096	$1,166,516.27	0.102
5120	$1,344,841.78	0.085	$1,933,238.24	0.096
10240	$1,777,629.83	0.072	$2,211,931.06	0.079

Table 3: This table shows the corresponding finishing account funds and average monthly percent increase for various sets of starting amounts both with and without a maximum bet.

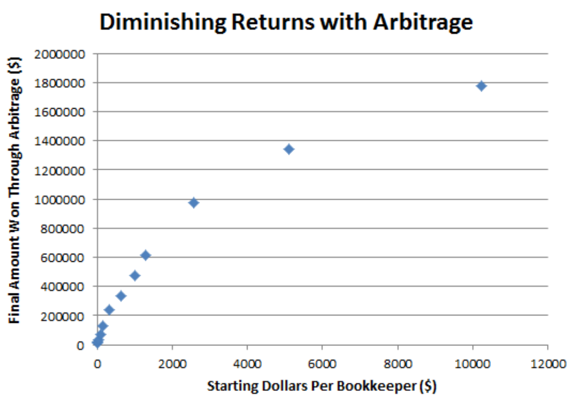


Figure 5: Shows the diminishing returns with the increasing starting dollars per bookkeeper.

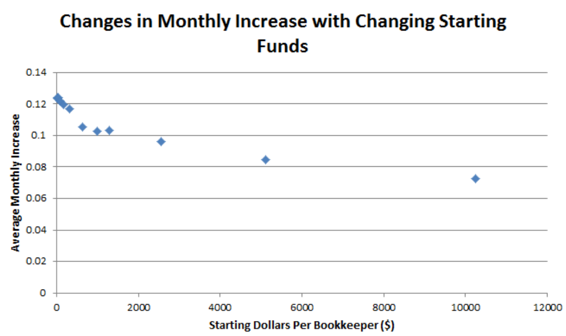


Figure 6: Shows the decreasing average monthly percent while increasing the starting dollars per bookkeeper.

$$\log_{36}\frac{A}{P} = 36\log_{36}(1+x)$$

$$x = \left(\frac{A}{P}\right)^{\frac{1}{36}} - 1$$

I ran the exact simulation without any maximum bets to see the effect that these maximum bets had on diminishing returns. Table 3 and Figures 7 and 8 show the results of the simulation. The diminishing returns are not as material and starting with more money per bookkeeper does not make as much of a difference in terms of the average percent monthly growth.

My program still results in diminishing returns without maximum bets because of the precautions taken by me to avoid being blacklisted by bookkeepers. In my simulation, I postpone betting on

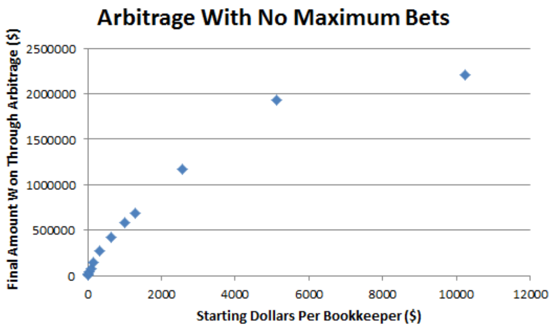


Figure 7: Shows the diminishing returns with the increasing starting dollars per bookkeeper without any maximum bets.

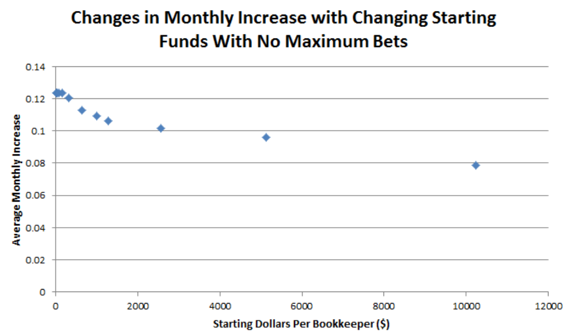


Figure 8: Shows the decreasing average month percent when increasing the starting dollars per bookkeeper without any maximum bets.

any bookkeepers after losing $10,000 with them in order to amortize the losses over a larger period of time. When I return betting with that bookkeeper, the losses will have accrued over a longer period of time (at least 6 months) reflecting an amortized loss that is more reasonable. For example, assuming that I can at worst lose $10,000 in 6 months, it will appear that I have lost less than $30 per day after the 6 month hiatus with the bookkeeper. Although it is impossible for me to determine when a given bookkeeper will blacklist an individual, I had to create some standard that I could use in my simulation. Each bookkeeper is running sophisticated algorithms to analyze the betting patterns of gamblers and some bookkeepers actively seek to ban arbitrage gamblers. Most of the bookkeepers my simulation uses have high tolerances for arbitrage gamblers. However, BetVictor and bet-at-home both actively try to prevent arbitrage gamblers. By eliminating the blacklist feature and assuming that no bookkeeper will blacklist me for losing too much money, there is a perfectly linear correlation between the amount of starting money and the final amount after three years of simulation (Figure 9). Regardless of the starting amount, the simulation averages 12% growth per month. Companies such as RebelBetting and Sports Arbitrage World claim to return 10% every month, but in reality that can only hold for so many months before maximum betting amounts and maximum loss limits established by bookkeepers would curtail continued growth at that rate. At that point, the compounding growth stops and there is a percent return (rather than growth) received each month. Sports World Arbitrage was founded in 1999, or at least 168 months ago[5]. If they really compounded 10 percent per month for 168 months, if one of the founders started by investing

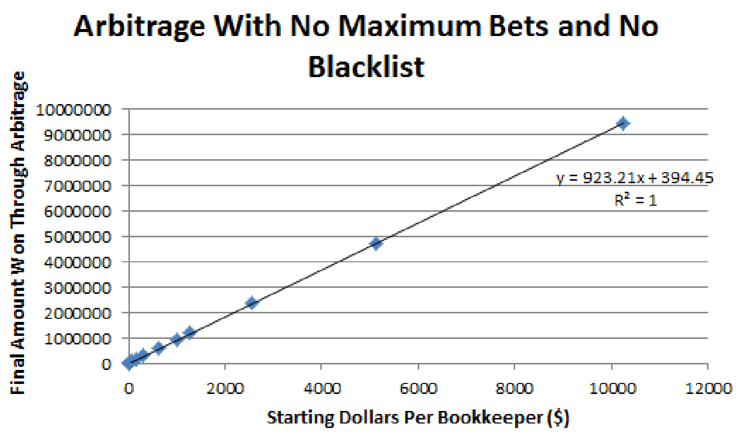


Figure 9: Shows that there is no diminishing returns if there are no maximum bets and no blacklisting.

$1,000 of his own money, he would now have $8,994,377,403.48, assuming he never withdrew any money from his various accounts. The entire world of online gambling had a market size of $35.80B in 2012, which includes all types of gambling and not just sports[2]. Given these numbers, it is almost certainly the case that the same diminishing returns affect RebelBetting and Sports Arbitrage World.

There are also some interesting correlations between average vigorish and usefulness in arbitrage opportunities. Consider the following scenario on April 4, 2009. The New York Mets hosted the Boston Red Sox in a preseason matchup. The Mets won the game 4 to 3. The identified bookmakers offered the following payoffs in decimal format in Table 4. Again, the probability of the New York Mets winning can be represented as a random variable and all of these bookkeepers are trying to estimate this random variable to the best of their ability. The average probability offered by these bookkeepers is .486 for the New York Mets winning and .514 for the Boston Red Sox winning. The standard deviations are both .028 (since all of the averages are related, summing to 1). Figure 10 is a diagram of the probabilities offered by the bookkeepers. The blue vertical line is the average probability of all the bookkeepers and the green vertical line represents standard deviations from the mean. The black rings represent the various probabilities that the bookkeepers have decided on, the gold rings are the optimal arbitrage opportunities, assuming a mini-max strategy.

A gambler exploiting the arbitrage opportunity would bet with Pinnacle Sports on the Mets

Bookkeeper	Home Payoff	Away Payoff	Vigorish	Home Prob.	Away Prob.
10Bet	1.77	2.15	2.9%	.452	.548
bet-at-home	1.80	1.90	7.6%	.486	.514
bet365	2.00	1.83	4.4%	.522	.478
BetVictor	1.80	1.91	7.3%	.485	.515
bwin	1.83	2.00	4.4%	.478	.522
Ladbrokes	2.00	1.83	4.4%	.522	.478
myBet	1.75	2.10	4.5%	.455	.545
Paddy Power	1.73	2.00	7.2%	.464	.536
Pinnacle Sports	2.02	1.91	1.8%	.514	.486

Table 4: Payoffs provided by various bookkeepers for the New York Mets, Boston Red Sox baseball game, as well as the corresponding probabilities.

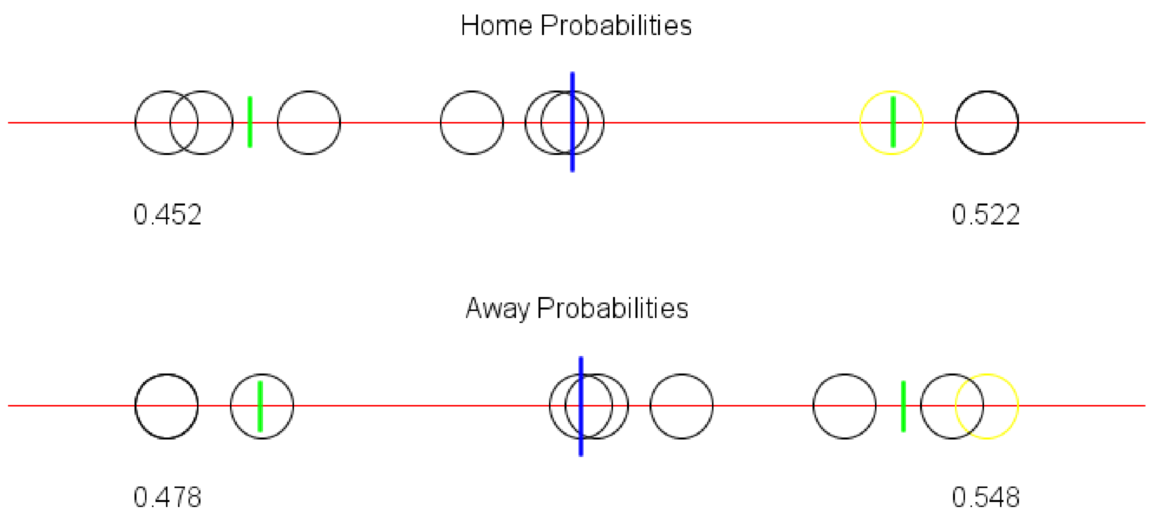


Figure 10: A graph showing the various probabilities different bookkeepers found for the New York Mets versus the Boston Red Sox baseball game.

winning and with 10Bet on the Red Sox winning. However, Pinnacle Sports did not believe that the Mets had the highest probability of winning. When accounting for vigorish, both bet365 and Ladbrokes have the Mets winning with probability 0.522, while Pinnacle Sports believes the Mets would win with probability .514. However, both bet365 and Ladbrokes had a 4.4 vigorish percentage while Pinnacle Sports had a 1.8 vigorish percentage. Thus, the better arbitrage opportunity came from betting on Pinnacle Sports even though Pinnacle Sport's algorithm gave the Mets a lower probability of winning. The following graph shows the average vigorish percentage taken per game by each bookkeeper versus the percentage of arbitrage opportunities associated with that bookkeeper over the years from 2006 to 2013. As you can see from Figure 11, there is a negative

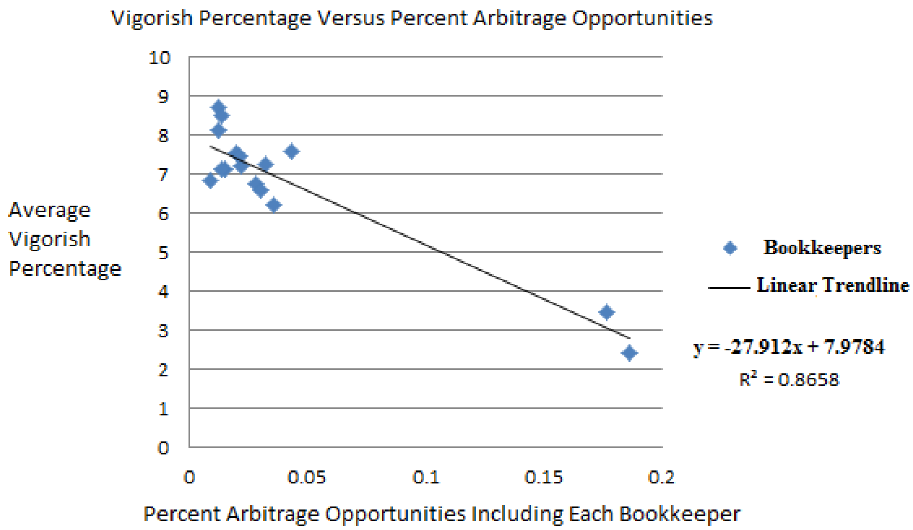


Figure 11: Accompanying graph showing the correlation between percent arbitrage opportunities and vigorish average for each bookkeeper.

linear correlation between the vigorish average for each bookkeeper and the number of arbitrage opportunities with that bookkeeper. Thus, bookkeepers that tend to charge less per game in vigorish can be used more often in arbitrage opportunities.

All of these bookkeepers have sophisticated computer models that attempt to predict the probabilities of various teams winning. The bookkeepers make the most money the closer their odds are to the actual probability, since otherwise gamblers would exploit the under predicted team and make a disproportionate amount of money. As time progresses, bookkeepers refine their algorithms based on their previous data to create a more accurate predictor of game probabilities. Theoretically, there is an actual set probability for each team winning, since we can consider each individual match as a random variable that sums up the probability of every possible sequence of events. Although bookkeepers do not consider each possible sequence of events, they are trying to find a good model to approximate all possible scenarios with a reliable probability. If one can assume that there is an actual probability associated with each outcome of a match, and that bookkeepers are trying to constantly improve their algorithms, then odds for all bookkeepers should converge, perhaps very

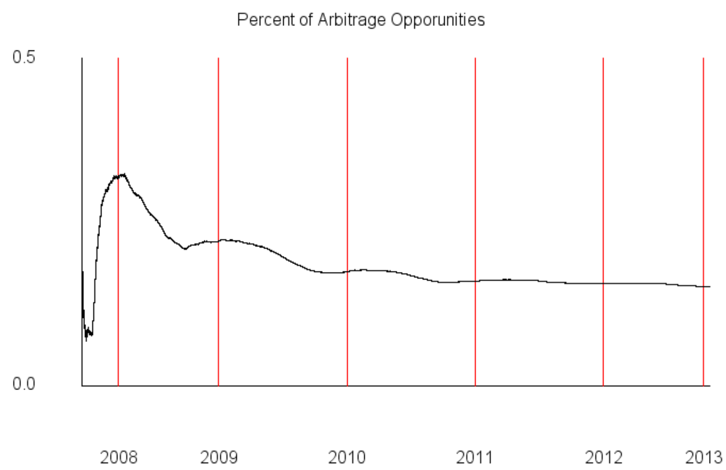


Figure 12: Decreasing number of arbitrage opportunities over the last six years.

slowly, to the probability of each team winning. Thus, bookkeepers' payoffs should move closer together, limiting arbitrage opportunities. As we can see in Figure 12, the number of arbitrage opportunities has decreased in the last six years.

9. Future Developments

My simulation only considered the closing odds right before a game started. Even still, there were many arbitrage opportunities (even though theoretically the market should have closed all arbitrage opportunities shortly after appearing). This means that the market with these particular bookkeepers does not act as quickly as previously hypothesized. It also means, however, that in the future I could look at the odds for games a day or more in advance and try to guess when the best payoff will occur. The trouble with betting on closing odds is that all of the information is known. However, a full day before the game there is a lot more variability in outcome and the payoffs tend to fluctuate more. The problem then becomes an online problem of randomly guessing on which odds to bet and when based on previous history, current situation, and payoff movements. There is some risk associated with arbitrage betting that follows this strategy. If, after betting on what appears to be an outlying odd, all of the bookkeepers push their odds to that value, there is going to be no way to counter the original bet to guarantee winning.

Alternatively, with changing odds over a span of days, if at any moment of time there appears an arbitrage opportunity a gambler could exploit the opportunity. Thus, there are more instances in time that arbitrage gamblers can consider when finding arbitrage opportunities. In a future extension to the project, the simulation could analyze many different games spanning a week's period and see if there are any arbitrage opportunities at a given time.

This entire project focused on the most simple gambling scheme - either the home team wins, draws, or losses. There are many other outcomes bookkeepers offer, such as who scores first, goal/score differential, and many other outcomes. These odds all offer more arbitrage opportunities that a future iteration of the project could exploit.

10. Conclusion

Sports gambling arbitrage is feasible, but there are diminishing returns. The returns claimed by Sports Arbitrage World and RebelBetting represent a short time span before the user reaches maximum bets. These services also charge for their services, and in the above analysis for the three year time period, a user would pay around 17% of his total profit to the arbitrage website if he began with $1400 in various accounts. There are also many issues with arbitrage gambling including avoiding detection by bookkeepers. The best solution for avoiding detection is to find the arbitrage opportunity and then bet a round amount (e.g. $30.00 instead of $29.39)[14]. For more profitable arbitrage opportunities, a gambler needs to constantly evaluate odds.

There are many risks associated with arbitrage gambling. Sports gambling is not legal in many countries. Also, some bookkeepers try to blacklist arbitrage betters. These bookkeepers fear that one bad odd that they produce will cost them thousands of dollars if arbitrage gambling is allowed on their website. These bookkeepers can blacklist gamblers whenever they want and do not need to accept bets from gamblers they have banned. Many bookkeepers operate out of third world countries, and some of these countries may not have a lot of legal communication with the United States. Although it is possible to mitigate some of the risks by selecting bookkeepers wisely, there are still issues.

It is possible to beat the current arbitrage methods by focusing on probabilities and redistributing the account balance of all the bookkeepers at frequent intervals. Using online accounting techniques and services, it is becoming cheaper and cheaper to shuffle money around. The market encourages bookkeepers to allow easy withdrawals, since their competitors do. Eventually, gamblers will be able to quickly withdraw all of the money from their accounts with minimal transaction fees.

There are long-term problems that come with arbitrage gambling. As bookkeepers improve their probability estimates and as more gamblers seek to gamble on arbitrage opportunities, the number of arbitrage opportunities and their duration decreases. If bookkeepers were to continually update their odds, there would be no arbitrage opportunities that would last more than a few seconds. Arbitrage gamblers bet down the extremely high payoff odds towards a final probability, and the more arbitrage gamblers, the quicker the odds will change because of the betting.

It is not clear whether arbitrage gambling is a viable long-term investment strategy. There are many risks associated with arbitrage gambling and as a gambler continues to bet it becomes more likely that he will be blacklisted and precluded from gambling with one or more bookkeepers. Thus, the risks of arbitrage gambling and the chance of blacklisting make it much harder to have a sustained investment in arbitrage. Arbitrage should be seen as a way to make a substantial amount of money over a shorter period of time, with the realization that in the future the arbitrage opportunities will vanish as the market progresses and as bookkeepers become more adapt at blacklisting.

This paper represents my own work in accordance with University regulations.

Brian Matejek - May 7, 2013

11. Appendix A

The following table shows the different bookkeepers that my simulation considered. These bookkeepers come from around the world and specialize in different sports.

Table 5: OddsPortal.com provided thousands of odds for different games from each of these bookkeepers.

Bookkeeper	Country
10Bet	United Kingdom
bet365	United Kingdom
Bet-at-Home	Malta
BetVictor	Gilbraltar
bwin	Gilbraltar
Dafabet	Philippines
Ladbrokes	United Kingdom
Marathon Bet	United Kingdom
myBet	Malta
NordicBet	Estonia
Paddy Power	Ireland
Pinnacle Sports	Curacao
SBOBET	United Kingdom
TonyBet	Estonia
Unibet	United Kingdom
William Hill	United Kingdom

12. Appendix B

Appendix B contains the list of all of the sports leagues that I scraped data from. There are over 1,000 teams from 13 different sports that make up these leagues. Some of the leagues are championship games and represent only a small number of teams).

Sport	Country	League
American Football	Canada	CFL
American Football	Germany	GFL
American Football	USA	NFL
American Football	USA	Ncaa I A
Aussie Rules	Australia	AFL
Aussie Rules	Australia	NAB Cup
Bandy	Finland	Bandylliga
Bandy	Norway	Eliteserien
Bandy	Russia	Super League
Bandy	Sweden	Allsvenskan Norra

Bandy	Sweden	Allsvenskan Sodra
Bandy	Sweden	Elitserien
Baseball	Italy	IBL
Baseball	Japan	NPB
Baseball	Mexico	LMB
Baseball	Netherlands	Hoofdklasse
Baseball	South Korea	KPB
Baseball	USA	IL
Baseball	USA	MLB
Baseball	USA	PCL
Baseball	World	World Baseball Classic
Basketball	Brazil	NBB
Basketball	England	BBL
Basketball	Europe	Euroleague
Basketball	Finland	Korisliiga
Basketball	France	LNB
Basketball	New Zealand	NBL
Basketball	Russia	PBL
Basketball	Spain	ACB
Basketball	Sweden	Ligan
Basketball	Turkey	TBL
Basketball	USA	NBA
Basketball	USA	NBA D League
Basketball	Ukraine	Superleague
Cricket	India	Premier League
Cricket	United Kingdom	County Championship One
Cricket	United Kingdom	County Championship Two
Handball	France	Division 1
Handball	Germany	Bundesliga
Handball	Russia	Superleague
Handball	Spain	Liga Asobal
Hockey	Czech Republic	Czech 1 Liga
Hockey	Czech Republic	Extraliga
Hockey	Europe	Karjala Cup
Hockey	Germany	DEL
Hockey	Poland	Polish Hockey League
Hockey	Russia	KHL
Hockey	Sweden	Elitserien
Hockey	USA	NHL
Hockey	United Kingdom	Elite League
Rugby League	Australia	NRL
Rugby League	Australia	NSW Cup
Rugby League	Australia	QLD Cup
Rugby League	England	Championship Cup
Rugby League	Europe	Challenge Cup

Rugby League	Europe	Super League
Rugby League	World	World Cup
Rugby Union	England	Aviva Premiership Rugby
Rugby Union	Europe	Challenge Cup
Rugby Union	Europe	Heineken Cup
Rugby Union	Europe	LV Cup
Rugby Union	Europe	Pro 12
Rugby Union	Europe	Six Nations
Rugby Union	France	Pro D2
Rugby Union	France	Top 14
Rugby Union	Italy	Super 12
Rugby Union	New Zealand	ITM Cup
Rugby Union	South Africa	Currie Cup
Rugby Union	World	Super Rugby
Soccer	Africa	Africa Cup Of Nations
Soccer	Australia	A League
Soccer	Brazil	Campeonato Brasileiro
Soccer	Canada	CSL
Soccer	Czech Republic	Gambrinus Liga
Soccer	England	Championship
Soccer	England	League One
Soccer	England	League Two
Soccer	England	Premier League
Soccer	France	Ligue 1
Soccer	France	Ligue 2
Soccer	Germany	2 Bundesliga
Soccer	Germany	Bundesliga
Soccer	Italy	Serie A
Soccer	Italy	Serie B
Soccer	Poland	Ekstraklasa
Soccer	Scotland	Division 1
Soccer	USA	Major League Soccer
Volleyball	Germany	1 Bundesliga
Water Polo	Hungary	OB I
Water Polo	Italy	A1

Table 6: All of the sports leagues that I scraped data for the seasons ranging from 2006 to 2013.

References

[1] Dafabet first deposit bonus (euro). [Online]. Available: http://www.dafamedia.com/dafabet/2012/03/dafabet-first-deposit-bonus-gbp/?trackingId=TRK247862736

[2] "Online gambling market - worldwide market volume 2003-2012 | statistic," *statista*. [Online]. Available: http://www.statista.com/statistics/168622/market-volume-of-online-gaming-worldwide-since-2003/

[3] (2010) Sports arbitrage pro - 10bet. [Online]. Available: http://www.sportsarbitragepro.com/reviews/10Bet

[4] (2013) Sports arbitrage - rebelbetting. [Online]. Available: http://rebelbetting.com/

[5] (2013) Sports arbitrage world | home. [Online]. Available: http://www.sportsarbitrageworld.com/

[6] BetVictor. (2013) Betvictor affiliates - terms & conditions page 1. [Online]. Available: http://www.victorsaffiliates.com/en-gb/terms-one.htm

[7] E. Franck, E. Verbeek, and S. Nüesch, "Inter-market arbitrage in sports betting," pp. 1–26, Oct. 2009.

[8] A. Hidayat. (2013) Phantomjs: Headless webkit with javascript api.

[9] T. H. Kjeldsen, "John von neumann's conception of the minimax theorem: A journey through different mathematical contexts," pp. 39–41, 2001.

[10] LiveSport s.r.o. 10 bet review. [Online]. Available: http://www.oddsportal.com/bookmaker/10bet/

[11] ——. mybet review, mybet bonus, free bet. [Online]. Available: http://www.oddsportal.com/bookmaker/myBet

[12] ——. Nordicbet review, nordicbet bonus, free bet. [Online]. Available: http://www.oddsportal.com/bookmaker/nordicbet/

[13] ——. Odds portal - betting odds monitoring service. [Online]. Available: http://www.oddsportal.com/

[14] ——. Odds portal news: Sure betting risks explained. [Online]. Available: http://www.oddsportal.com/news/sure-betting-risks-explained-16/

[15] ——. Sbobet review, sbobet bonus, free bet. [Online]. Available: http://www.oddsportal.com/bookmaker/sbobet/

[16] ——. Tonybet review. [Online]. Available: http://www.oddsportal.com/bookmaker/tonybet/

[17] ——. William hill review, william hill bonus, free bet. [Online]. Available: http://www.oddsportal.com/bookmaker/william-hill/

[18] D. Pennock, "The wisdom of the ProbabilitySports crowd," http://blog.oddhead.com/2007/01/04/the-wisdom-of-the-probabilitysports-crowd/, January 04 2007.

[19] PinnacleSports.com. Arbitrage friendly. [Online]. Available: http://www.pinnaclesports.com/betting-promotions/arbitrage-friendly

[20] R. Sedgewick and K. Wayne, "Stddraw.java," http://introcs.princeton.edu/java/stdlib/StdDraw.java.html.

[21] ——, "Stdin.java," http://introcs.princeton.edu/java/stdlib/StdIn.java.html.

[22] SureBetMonitor.com. Bet365 review - surebet monitor - sports arbitrage software, alert service, arbitrage bettingsurebet monitor – sports arbitrage software, alert service, arbitrage betting. [Online]. Available: http://www.surebetmonitor.com/reviews/bet365-review/

[23] ——. Betclic, expekt, bet-at-home (betclic everest group) review - surebet monitor - sports arbitrage software, alert service, arbitrage bettingsurebet monitor – sports arbitrage software, alert service, arbitrage betting. [Online]. Available: http://www.surebetmonitor.com/reviews/betclic-expekt-bet-at-home-betclic-everest-group-review/

[24] ——. Ladbrokes review - surebet monitor - sports arbitrage software, alert service, arbitrage bettingsurebet monitor – sports arbitrage software, alert service, arbitrage betting. [Online]. Available: http://www.surebetmonitor.com/reviews/ladbrokes-review/

[25] ——. Williamhill reveiew. [Online]. Available: http://www.surebetmonitor.com/reviews/williamhill-review/

[26] ——. (2012) Bwin loyalty points - added value to arbitrage betting « surebet monitor - sports arbitrage software, alert service, arbitrage bettingsurebet monitor – sports arbitrage software, alert service, arbitrage betting. [Online]. Available: http://www.surebetmonitor.com/2012/01/11/bwin-loyalty-points-added-value-to-arbitrage-betting/