Betting Market Efficiency:

An Examination of Australian Sport

Dario Dodig

A thesis presented for the degree of
Master of Research in Economics

MACQUARIE
University
SYDNEY · AUSTRALIA

Department of Economics
Macquarie University
Sydney, Australia
October 9, 2015

Statement of Candidate

I certify that the work in this thesis entitled "Betting Markets: An Examination of Australian Sport" has not been previously submitted for a degree nor has it been submitted as part of requirements for a degree to any other university or institution other than Macquarie University.

I also certify that the thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself have been properly acknowledged.

In addition, I certify that all information sources and literature used are indicated in the thesis.

Dodig Dario

Abstract

This study has two main purposes: (1) to investigate whether the sports betting markets of the Australian Football League, the National Rugby League and the A-League are efficient. (2) To examine whether differences in incentives or competitive balance can explain efficiency differences.

Efficiency is tested in three stages. Firstly, historical betting odds are analysed for arbitrage opportunities. Secondly, betting simulations are backtested to determine whether traditional biases such as the favourite-longshot bias exist. Thirdly, the extent of inefficiency is tested directly using a conditional logit model.

The data consist of historical odds and league statistics from 2010 to 2015. Between eleven to fifty-five bookmakers are surveyed, depending on the match. The results provide evidence of many arbitrage opportunities, lending support for inefficiency at a rudimentary level. Statistical testing reveals that average bookmaker odds are weakly efficient. The extent to which incentives and competitive balance can explain efficiency differences is varied.

Acknowledgements

Working on this thesis has provided a learning process and a written piece that have been made so much more thorough by the support of those around me. My supervisors, Dr. Jordi McKenzie and Dr. Christopher Heaton have provided me with invaluable input in all areas of my thesis. Without their help it would be a mere shell of the final product. Associate Professor Tony Bryant has created a journey that has been both enjoyable and a very mind-broadening experience. I have learnt more from sitting in a small room bouncing around ideas than I ever thought possible.

The support I have received from my friends and family has been incredible. Without their endless help and support in all aspects of life, this entire process would have been far more difficult. From accepting my lack of company in these past few months to proof-reading and suggestions; it is all greatly appreciated.

Finally, Macquarie University has provided generous financial support without which I could not have paid as much attention to my studies as I have.

Contents

List of Figures

List of Tables

1 Introduction

According to the Productivity Commission's (2010) report on gambling, roughly 70% of Australian adults participate in some form of gambling each year. It is an activity that has existed for thousands of years and is seen in the remains of many ancient societies (Preston et al., 1998). For example, an early version of the shell game is seen on the wall of an Egyptian burial vault from 2500 B.C.[1] In London, archaeological excavations uncovered a dice game that was played circa 2000 B.C. (Schwartz, 2013).

Likewise, many examples of betting on sporting events exist throughout history. The indigenous people of North America placed bets on ball games and foot races (Thompson, 2001). The ancient Romans had arenas scattered throughout the empire which were used primarily for chariot racing. These arenas drew in massive crowds of up to 260 000 people (Humphrey, 1986).

In modern society, with the advent of the Internet, people are able to place a bet on essentially any professional sporting event occurring almost anywhere in the world. The online betting markets where such bets are placed have developed into a thriving industry that has grown rapidly in popularity in the last decade (Direr, 2013). This thesis is concerned with the efficiency of these markets.

The concept of market efficiency was originally formulated for financial markets (Fama, 1970) but applies in a natural way to betting markets where asset prices are replaced by betting odds. A special feature of betting markets which distinguishes them from most other markets is that the payoff distribution of the assets (wagers) is observable with significant accuracy. This is extremely useful when it comes to estimating bettor preferences regarding odds pricing.

This thesis begins with a primer on gambling which is followed by a section on common nomenclature used in betting circles. Chapter 2 provides an overview of the literature on betting market efficiency and measuring competitive balance in sports. In Chapters 3 and 4 the analysis is presented. The efficiency of the betting markets of three Australian sporting codes is tested. These include the Australian Football League (AFL), the National Rugby League (NRL) and the A-League (the highest level of soccer[2] in Australia). Chapter 5 concludes with a brief summary of the findings and suggestions for future research.

[1] A common variant of the shell game involves placing a ball under one of three identical containers so that it can not be seen. These three containers are then shuffled and a player is invited to bet on which container holds the ball.

[2] "Football" and "soccer" are often used interchangeably. To avoid confusion, the latter term will be used throughout this thesis.

1.1 Gambling as Irrational Behaviour

Gambling is an activity that at first glance seems to contradict rationality as conventionally defined in economics. This is because individuals voluntarily use their own money to engage in an activity which will almost certainly lose them money in the long run. The observation that gambling has been a popular endeavour for millennia necessitates the examination of the theory if one is to arrive at a general description of bettor behaviour.

Economics has long been concerned with the rational utility-maximising agent (see von Neumann and Morgenstern, 1944). According to Stigler (1970), the three characteristics of a rational consumer are that their tastes are consistent, their cost calculations are correct, and that they make decisions which maximise utility. Various empirical observations provide evidence that refutes the utility maximisation hypothesis. Famous examples include the Allais paradox (Allais, 1953)[3] and the Ellsberg paradox (Ellsberg, 1961)[4], among others.

Findings such as these have greatly influenced the development of choice theory in economics. They have led to the creation of alternatives to the utility maximisation hypothesis, such as prospect theory, regret theory and many others (Machina, 1987, provides an overview of various alternatives). These theories have had a significant influence in explaining gambling behaviour.[5]

1.2 How do Betting Markets Compare to Other Financial Markets?

There are many similarities between betting markets and other financial markets, however, important differences exist. Betting markets generally have a negative expected outcome for bettors, whereas conventional markets have a positive expected outcome over time for investors. Information and skill are diffused through conventional markets but are concentrated on the supply side of betting markets. Regulation of conventional markets is generally stronger than regulation of betting markets (Coleman, 2013).

[3] The Allais paradox refers to the empirical observation that the modal choice pattern in certain decision games violates the independence axiom on preference ordering.

[4] The Ellsberg paradox refers to the empirical observation that people generally prefer to take on risk in situations where they know specific odds rather than an alternative risk scenario in which the odds are completely ambiguous. This holds even if the known probability of winning is low and the unknown probability of winning could be a guarantee of winning.

[5] Nyman et al. (2008) provide a useful summary of how various theories can be used to explain gambling behaviour before presenting their own model.

In some cases, direct comparisons have been made between certain financial assets and betting. For instance, Ruhm (2003) shows how certain positions in options trading are essentially simple bets, while Vecer et al. (2006) compare betting contracts with credit derivatives. Others have, however, suggested that the participants in betting markets may differ from the participants in other markets for various reasons (Coleman, 2013). If this is the case, then models of betting markets that assume concepts and behaviours identical to those in conventional markets have the potential to be misspecified and reach erroneous conclusions.

1.3 What Does Efficiency Refer to in Betting Markets?

Fama (1970) delineated three levels of information efficiency that vary in stringency. The first level, weak-form efficiency, stipulates that current prices embody all information contained in historical prices. In the context of betting markets, this implies that no betting strategy should be possible based on historical odds that can yield a positive expected return.

Ziemba (1995) further distinguishes between weak-form efficiency and strict weak-form efficiency. Ziemba states that betting at particular odds, or ranges of odds, should not yield different expected returns to betting at any other odds or range of odds. In betting markets, the most empirically cited violation of strict weak-form efficiency is what is known as the favourite-longshot bias (FLB), which is the empirical observation that the expected return from betting on horses/players/teams with short odds is greater than that from longshot bets, and expected return rises with the (implied) probability of a win.

In the betting literature, the FLB is often referred to as an inefficiency. A careful distinction must be made to note that there is a difference between statistical inefficiency and economic inefficiency. In betting markets, the latter is a sufficient condition for the former, but the converse is not true. Existence of the FLB indicates statistical inefficiency. It may or may not indicate economic inefficiency. This is discussed further when we review the literature in Chapter 2.

The second level outlined by Fama is semi-strong form efficiency, which requires that market prices reflect all publicly available information. In betting markets, this implies that profitable betting strategies based on publicly available information should not be possible. If market odds are known to over/underreact to news events then this information could be used to construct a profitable betting strategy.

The third and strictest notion of information efficiency is strong-form efficiency, which requires that market prices reflect all information whether publicly or privately held. Considering strong-form efficiency where private information is factored in, it should not be possible for any participant trading on superior information to make abnormal returns. Strong form efficiency in betting markets is problematic because of issues such as match fixing. In Australia and many other countries, professional sports players and people associated with the industry are barred from betting on their own sport.[6]

This thesis is concerned with weak-form efficiency. To reiterate, a distinction is made between statistical efficiency and economic efficiency. Statistical inefficiency refers to abnormal patterns in returns, whereas economic inefficiency refers to monetarily profitable opportunities.

[6] In Australia, match fixing is taken very seriously. For instance, the NRL has an Integrity Unit which was formed in 2013 and investigates players believed to be involved in betting on matches.

1.4 Terminology

This section provides a brief overview of some common terminology used in betting markets. Most of these terms are used by both scholars and regular bettors.

Odds[7]: The payout received from a bet. For example, a $1 winning bet placed with odds of 2.5 returns $2.50. This results in a profit of $1.50.

Implied Probability: The probability derived from the odds. This is calculated using 1/Odds. For example, odds of 2.5 have an implied probability of $1/2.5 = 0.4$.

Overround[8]: The amount charged by a bookmaker for their services. It is calculated by summing the implied probabilities of all possible outcomes and subtracting 1. For example, consider a soccer match between two teams, A and B. Assume team A has odds of 2.5 and team B also has odds of 2.5. Assume the odds on a draw are 3.5.
Overround $= (1/2.5) + (1/2.5) + (1/3.5) - 1 = 1.0857 - 1 = 0.0857$.
Converted to a percentage $= 8.57\%$.

Parimutuel markets: A form of wagering in which all bets are placed together in a pool. The overround is removed and the payoff odds are calculated by sharing the pool amongst all winning bets.

Fixed-odds markets: A form of wagering where the bettor knows the exact odds they will receive when they place a bet. The odds are fixed once the wager has been placed.

This list comprises some of the most common terms used in betting markets. However, it is not exhaustive and additional definitions are introduced as necessary. We now proceed to review the literature.

[7] Throughout this thesis, decimal (also known as European, digital or continental) odds are used.
[8] Overround has many names including bookmaker margin, vigorish, take, juice and cut. Throughout most of this thesis, overround is used.

2 Literature Review

This review of the literature on betting markets is split into two broad sections. The first section covers the bulk of the review and consists of the large volume of studies testing betting market efficiency. For the most part, a chronological approach is taken, whereby the literature covered begins with the early studies on horse-racing and is followed by developments in other sports markets.

The sports betting literature itself is further categorised. The evidence on arbitrage opportunities is discussed, followed by tests of efficiency on the two different market micro-structures; handicap markets and fixed-odds win markets. This section concludes with an examination of the evidence on Australian sports, which includes a mix of studies on both of these market micro-structures.

The final part of this chapter consists of an overview of measuring competitive balance in sports. It is not intended to be an exhaustive review of the literature on competitive balance. Rather, its purpose is to outline a gap in the current literature and motivate one of the research questions in this thesis.

2.1 Betting Market Efficiency: A Review of the Empirical Literature

2.1.1 The Early Studies

Historically, betting market research has focused on horse-racing. The vast majority of the literature on horse-racing has found evidence of deviation away from strict weak-form efficiency, in the form of the favourite-longshot bias (FLB). The FLB refers to the phenomenon that returns to bets on favourites are generally found to be higher than returns to bets on longshots.

Preston and Baratta (1948) were the first to find evidence of the existence of the FLB. However, this finding was not in a naturalistic setting. They conducted an experiment (with a card game specifically designed for the experiment) using undergraduate students and staff members of their university that aimed to determine whether people's perceived (psychological) probabilities matched actual (objective) probabilities.

Preston and Baratta (1948) found that subjects overvalued the likelihood of low probability events occurring and undervalued the likelihood of higher probability events occurring. The indifference point where their subjects' psychological probabilities matched mathematical probabilities was observed to be approximately 0.2.

The first naturalistic evidence of the FLB was from the psychologist Griffith (1949). He was inspired by the laboratory evidence of Preston and Baratta (1948) but wanted to test the results in a complex, non-laboratory environment. Analysing U.S. racetrack data, Griffith found that horses with low probabilities of winning were systematically overvalued, while horses with high probabilities of winning were systematically undervalued, a result consistent with that of Preston and Baratta.

McGlothlin (1956) replicated Griffith's (1949) study with a larger dataset. McGlothlin's data also suggested the existence of the FLB. After the initial findings, scholars and betting experts alike wondered if the FLB was a general phenomenon that applied to all racetracks or a mere coincidence. In the decades that followed these original studies, a significant body of evidence for the bias emerged in betting markets around the world.

In subsequent years, the FLB was tested for and uncovered in different parts of the United States (Ali, 1977; Asch et al., 1982; Thaler and Ziemba, 1988), the United Kingdom (Dowie, 1976; Williams and Paton, 1997), Australia (Tuckwell, 1983) and New Zealand (Gandar et al., 2001). Comprehensive surveys on the FLB in horse-racing markets are offered by Sauer (1998) and Williams (1999).

While the FLB was observed in most data sets on horse-racing, there were some exceptions. These included markets in Hong Kong and Japan (Busche and Hall, 1988; Busche, 1994), the market at one U.S. racetrack (Swidler and Shaw, 1995) and some markets in the UK (Smith et al., 2006).

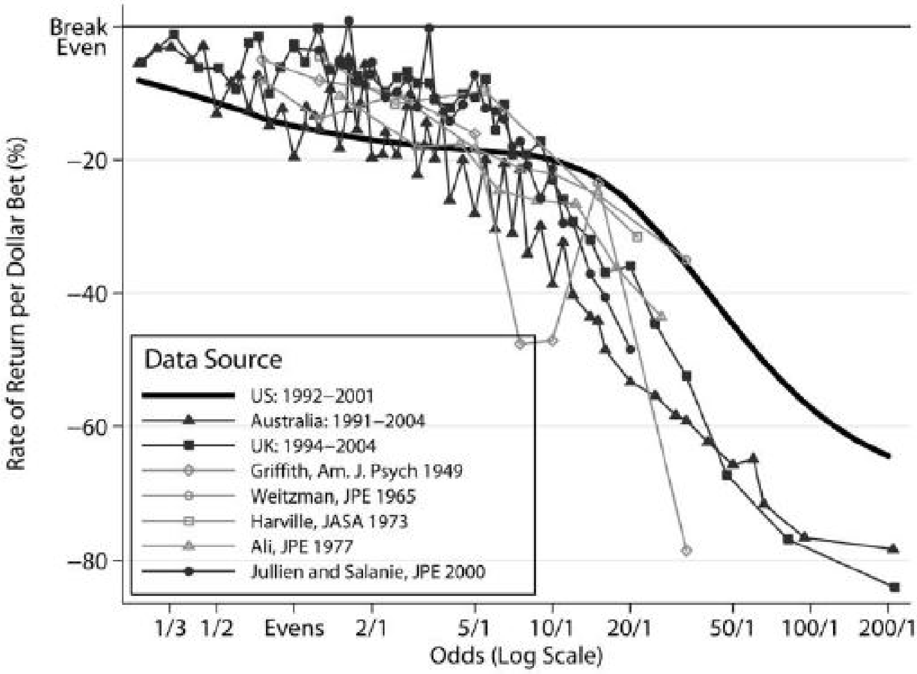
Although the FLB was widely observed, questions of robustness were raised given most studies examined relatively small samples of only a few thousand races. A large comprehensive study was still lacking until Snowberg and Wolfers (2010). Snowberg and Wolfers used a data set magnitudes larger than any previous study. It covered all 6.4 million horse-race starts in the U.S. between 1992 to 2001, as well as a further 2.7 million starts in Australia between 1991 to 2004 and 380,000 starts in the UK between 1994 and 2004. They concluded that the FLB exists in all three data sets.

What is interesting to note from Snowberg and Wolfers (2010) is that the FLB was found regardless of the structure of the betting market. In the U.S., almost all betting in their sample occurred in a parimutuel market. In the UK, almost exclusively fixed-odds bookmaker pricing is used. In Australia, it is a mixture of both; bookmakers competing with a parimutuel system.

Figure 1 appears in Snowberg and Wolfers (2010) and shows the FLB in horse-racing. It contains data from their own study as well as five previous studies. Returns to bets on favourite horses are on the left of the graph, while returns to bets on longshots are towards the right. There is a clear downwards trend in the rate of return as odds increase.[9] If there was no bias, then returns to bets would accurately reflect the odds implied probabilities on those bets. Essentially, the line on the graph would be horizontal at 1 minus the track take.

[9]Snowberg and Wolfers (2010) note that, in contrast with some findings such as Thaler and Ziemba (1988), there was no systematic profit opportunity on heavy favourites (horses with probability greater than 0.5).

Figure 1: The Favourite-Longshot Bias in Horse-Racing. Source: Snowberg and Wolfers (2010).



These results raise a number of questions. Why does the FLB exist? Why are returns to favourites systematically higher than returns to longshots? And why are there some exceptions to the rule? Answering these questions has proven very difficult. To this day, the issues are not completely resolved. However, the number of competing explanations has decreased over time. Some of the theoretical explanations for the existence of the FLB will now be discussed.

2.1.2 Explanations for the Existence of the Favourite-Longshot Bias

As the early studies by Preston and Baratta (1948) and Griffith (1949) were both conducted by psychologists, there was no economic theory behind the phenomenon. They gave vague explanations that attributed the FLB to the tendency of individuals to overestimate (underestimate) the occurrence of small (large) probability events.

Since then, many attempts have been made to explain why the FLB exists. Ottaviani and Sorenson (2007) provide an in-depth overview with mathematical derivations. Hurley and McDonough (2007) assert that the existing theories can be split up according to differences in one or more of the following three categories:

1. Bettor risk preferences.
2. Information (whether bettors have asymmetric or heterogeneous information).
3. Market micro-structure (whether the market is a parimutuel or bookmaker market).

The majority of the early literature concerning bettor risk preferences took the position that bettors were risk-loving, with a convex utility of wealth curve. This line of research would include the work of Weitzman (1965), Ali (1977), Quandt (1986) and Suvas (1992). The idea of risk preference has been referred to as an adrenaline factor from the increased excitement derived from betting on longshots (Bruce and Johnson, 1992), as well as bragging-rights for winning on longshots (Thaler and Ziemba, 1988). In either case, a convex utility of wealth function is derived.

The rest of the important contributions turn on information and/or market micro-structure. Ali (1977) studies a two-horse race where bettors have heterogeneous expectations and the betting mechanism is parimutuel. He shows that the FLB can arise under these assumptions. Blough (1994) extends Ali's model to the case of an n-horse race and finds that the bias will emerge even under the condition of symmetric heterogeneous expectations. Hurley and McDonough (2005) extend Blough's work to the case of sequential parimutuel betting with heterogeneous expectations. They show that both the usual and reverse biases are possible depending on the distributions of bettor beliefs.

Another set of models turns on asymmetric information. These models assume that there is a class of bettor which is more informed about the outcome of the contest and would include the work of Shin (1992), Hurley and McDonough (1995, 1996), Williams and Paton (1998), and Cain, Law and Peel (2003). Hurley and McDonough (1995, 1996) employ a parimutuel mechanism; the others study a bookmaker market. Hurley and

McDonough (2007) consider the case where bettors have heterogeneous beliefs and the bookmaker market is imperfectly competitive. They find that both the standard FLB as well as a reverse-FLB is theoretically possible in this scenario. Makropoulou and Markellos (2011) show how the FLB may occur due to information uncertainty alone.

The empirical observation that the FLB was not being found in every single market meant that it was necessary for any theoretical explanation to take this into consideration. Williams and Paton (1998) examine a model which has a demand side characterised by the presence of money from both 'uninformed' bettors (or noise traders) who cannot distinguish good bets, and 'informed' bettors who know the 'true' probabilities of each horse winning. They showed that a standard FLB, a reverse-FLB, or no bias was theoretically possible depending on the value of transaction costs (overround).

More recently, Snowberg and Wolfers (2010) examine two alternative theories, risk-preference and "probability misperception". They describe probability misperception by utilising prospect theory to explain the FLB. Snowberg and Wolfers proceed to compare a bettor with a concave utility function against a bettor who has a subjective utility function. This subjective utility function employs nonlinear probability weights that violate the reduction of compound lotteries in the form of "exotic" bets in horse-racing.[10] Snowberg and Wolfers conclude that the probability misperception model is better at explaining individual choices when evaluated using exotic bets.

It is interesting to see how economic theory has formalised what Preston and Baratta (1948) and Griffith (1949) found. However, it must be noted that Snowberg and Wolfers (2010) assume that bettors reduce the second choice in an exotic bet sequence in the same way as they do a single choice. While this is a reasonable assumption, their conclusion would not hold without it. As such, the debate on the theoretical reasons for the existence of the FLB is still not completely resolved. We now proceed to examine the sports betting literature, beginning with the evidence on arbitrage.

[10]The exotic bets Snowberg and Wolfers (2010) examine include exactas, quinellas and trifectas.

2.1.3 Arbitrage: The Evidence

Many authors have examined whether arbitrage opportunities exist in a variety of markets. In the context of sports betting, arbitrage would entail wagering on all possible outcomes and ensuring a risk-free profit. Arbitrage opportunities may exist if odds vary enough between different bookmakers such that a bettor could wager on different outcomes with different bookmakers and create an "underround" book.

Vergin and Scriabin (1978) were among the first to examine potential arbitrage opportunities in sports betting markets. They examined the National Football League (NFL) in the U.S. and found that bookmakers in different cities offered handicaps that differed enough to make arbitraging possible. Vergin and Scriabin noted that even though it would be difficult for an individual bettor to exploit, it would be possible for a syndicate to group together and split the winnings.

Pope and Peel (1989) study fixed-odds soccer markets in the UK and compare the odds between four different bookmakers. They find numerous examples where arbitrage was possible pre-tax, but only one match in the entire season that would have generated positive post-tax profit. Similarly, Dixon and Pope (2004) analyse the odds of three bookmakers over the three season period from 1993 to 1996 and find no arbitrage opportunities. Overall, Dixon and Pope found that the differences in odds between bookmakers was lower compared to those reported in Pope and Peel. They claimed that this is suggestive of more efficient forecasts for that season, or possibly the result of collusion between bookmakers (whether implicit or explicit).

Vlastakis et al. (2009) study five sets of bookmaker odds for 12,420 matches spanning 26 countries and events between 2002 to 2004. They find arbitrage opportunities in 63 matches, with an average return of 21.78%. Paton and Williams (2005) examine the English soccer betting market for booking-points.[11] They develop a "quasi-arbitrage" system designed to exploit bookmakers whose pricing differs significantly from the average spread. Paton and Williams use prices from five bookmakers and find that this quasi-arbitrage strategy generated positive returns in both the 'within' and 'reserved' samples of the 1999 to 2000 and 2000 to 2001 seasons.

These earlier studies examining arbitrage were conducted in a time where gambling was less accessible than it is today. It would have been far more difficult for bettors to uncover whether arbitrage opportunities exist, as they would have needed a physical

[11] Betting on booking points assigns point values to disciplinary actions. In soccer, a common structure is to assign 10 points for a yellow card and 15 points for a red card.

presence on site at the bookmaker's location.

With the development of the Internet, the betting industry has undergone a significant structural change (Smith and Williams, 2008). Due to the ease of comparing different bookmaker odds, arbitrage opportunities are less likely to exist as bettors would be able to easily take advantage of any differences in odds. Goddard (2011) examines whether this is the case for the UK football market over the 2009 to 2011 period. Goddard provides calculations of mean bookmaker overrounds from 10 prominent bookmakers. He notes that the "best odds" (the highest odds across all bookmakers for each outcome) overround was smaller than individual bookmaker overrounds, but that there was not a single match with an overround that was zero or negative.

The odds examined in all of these studies were the final odds given by each bookmaker immediately prior to the start of a match. The question remained as to whether arbitrage opportunities were possible at any point prior to the start of a match. Marshall (2009) examined this question and found that whilst arbitrage opportunities did often arise between different online bookmakers, they were removed within 15 minutes, on average, as bookmakers adjusted their odds. This provided evidence that betting markets were relatively efficient and similar to other markets in the way they removed arbitrage opportunities.[12]

In more recent studies, this result has not held and many arbitrage opportunities have been uncovered, even at final prices. Franck et al. (2013) examine arbitrage within the sports betting markets for the top five European soccer leagues and find that arbitrage opportunities existed for 19.2% of matches. They explain this apparent inefficiency by stating that bookmakers are behaving strategically with regards to long term profit. They claim that bookmakers are attempting to lure in new customers with favourable odds and are therefore prepared to make a short term loss on them.

Grant et al. (2015) also examine arbitrage opportunities in different European soccer leagues and find that arbitrage opportunities arise in 25.6% of matches (comparing 6 bookmakers). They conclude that mispricings are unlikely to be systematically exploitable as bookmakers may impose restrictions on the accounts of skilled bettors.

The explanation by Franck et al. (2013) seems more likely as there is little reason to suggest that individual bookmakers would not want arbitrageurs to use their service. Bookmakers would still be offering the same risk-reward opportunity as they do to any other customer and the fact that an individual bettor is hedging or arbitraging with another bookmaker should not be of concern.

[12]For example, Chordia et al. (2005) find that it takes between 5 and 50 minutes for order imbalances to be corrected by rational investors on the New York Stock Exchange.

As Levitt (2004) states, bookmakers act as profit maximising organisations and compete on price in the form of odds. Due to the extremely high level of competition, bookmakers may be willing to accept some losses (in the form of higher risk due to skewed odds) if it leads to bettor loyalty in the future. This effect could be similar to forms of below-cost pricing strategies that occur in other markets.[13] The possible existence of arbitrage opportunities in the markets under examination forms the initial analysis in this thesis.

[13] For example, the loss-leader strategy found commonly in retail markets, whereby a firm realises losses on one product in order to realise gains on other products (Hess and Gerstner, 1987).

2.1.4 Sports Handicap Markets: The Evidence

The early literature on sports betting was dominated by U.S. studies of handicap betting on the National Football League (NFL), the National Basketball Association (NBA), the National Hockey League (NHL) and Major League Baseball (MLB). During these earlier decades, betting markets in the U.S. generally offered handicap odds. Betting on an outright winner with fixed odds was rarely offered.[14]

These studies consisted of assessing whether there were biases in the way handicaps are set, for example, whether favourites were more or less likely to beat the handicap.[15] Roughly three quarters of these studies found deviations from market efficiency (Sobel and Ryan, 2008). Interestingly, in many of the cases, an opposite bias was found. That is, the underdog beat the handicap more often than the favourite.

To assess whether handicaps are set in an unbiased manner, generally an ordinary least squares (OLS) regression model is estimated. Gandar et al. (1988) and Sauer et al. (1988) set the standard test that many studies have emulated. The regression takes the form

$$Y_i = \alpha_0 + \beta_1 X_i + \varepsilon_i \tag{1}$$

where Y is the score difference between the home and the away team for match i, X is the betting handicap offered and ε is an error term. The joint hypothesis for efficiency is that $\alpha = 0$ and $\beta = 1$.

A reverse-FLB was found by Gandar et al. (1988) who showed that a simple strategy of betting on NFL underdogs in the handicap market would have produced a greater return than betting on favourites. Similarly, Kochman and Badarinathi (1992) found through a simple analysis of role (underdog or favourite), location of the game (home or away), and month in which the game was played, that a rate of success above break-even could be achieved for betting on NFL games between 1986 and 1990. Zuber et al. (1985) found an exploitable inefficiency in the NFL handicap betting market during the 1983 regular season, while Lacey (1990) finds some profitable opportunities in certain betting rules in NFL line markets from 1984 to 1986.

In the MLB, Woodland and Woodland (1994) find a reverse-FLB, which they confirm

[14] Fixed-odds win markets are referred to as "money-line" markets in U.S. nomenclature.

[15] Handicap markets require the bettor to pick whether Team A or Team B will beat its handicap, which is negative or positive depending on whether the team is the favourite or underdog, respectively. For example, Team A has a handicap of -10.5 and Team B a handicap of +10.5. If the bettor chooses Team B, then even if Team B loses by 5 points, the bettor still wins their bet because they didn't lose by more than 10.5. Payouts on handicap markets are generally slightly less than double the initial stake, for example, a \$10 wager will win \$9 plus the initial \$10 = \$19.

still exists after including 10 additional years of data in their updated study, Woodland and Woodland (2003). In their study of NHL betting markets, Woodland and Woodland (2001) find a strong opposite FLB, which is confirmed in an updated and revised version of this study by Gandar et al. (2004).

What is worthy of consideration is that incentives between the team and the individual bettor may be distorted when handicaps are considered. An individual bettor who wagers on the favourite wants them to win by more than the handicap, whilst the team playing may be content with simply winning.[16]

Examples of studies where efficiency was found include Sauer (1988), who tests the NFL over/under market, while Johnson and Pawlukiewicz (1992) find efficiency in over/under betting in the NBA.[17] To assess whether these markets were efficient, a similar procedure was employed as in Equation (1), where Y and X are replaced with total points and the over/under point estimate, respectively. These over/under markets are interesting because bettors are picking the total score rather than a winner. That a bias was not found in this category offers the possibility that bettors wagering in this market are more experienced and/or make more rational decisions.

A number of other studies also exist assessing whether simple strategies (analogous to trading strategies in financial markets) could have been implemented to make a positive rate of return. These include strategies such as the "hot-hand", which is similar to momentum trading strategies, and involves betting on a team who is on a winning streak of n games. There is also the alternative strategy, known as the "gambler's fallacy", which involves betting on a team who is on a losing streak of n games.

Camerer (1989) hypothesised that betting against teams on winning streaks could potentially be profitable. His hypothesis was that handicap lines set by bookmakers were influenced by winning streaks, but that game outcomes were not. Brown and Sauer (1993) tested Camerer's hypothesis and showed that both the handicap and the outcome of the game respond to streaks for teams in the NBA. Paul and Weinbach (2005) likewise tested the hot-hand in the NBA and concluded that betting against teams on winning streaks of two games or more to be profitable, but betting on teams on winning streaks of four games or more to be unprofitable.

Paul and Weinbach (2005) also tested the opposite strategy (the gambler's fallacy) and concluded that betting on teams experiencing losing streaks was not profitable. This

[16] This is not necessarily always the case, as teams who are tied on points at the end of a season are usually separated by a secondary measure such as goal difference. For that reason, teams may prefer to win by as large a margin as possible.

[17] The over/under market refers to betting on the total number of points or goals and whether the match will have more (over) or less (under) than the market predicts.

lends some support for inefficiency. However, the determination of the number of games, n, is quite arbitrary. It is possible that testing for the 'hot hand' or 'gambler's fallacy' for many different values of n will eventually come up with a profitable strategy for the data set under examination.

This covers some key findings in sports handicap betting markets. It is evident that while inefficiencies exist within these markets, the results are not as unanimous as they are in horse-racing. We now proceed to examine the evidence on fixed-odds sports markets.

2.1.5 Fixed-Odds Sports Markets: The Evidence

The literature in this section is the most relevant to this thesis, given fixed-odds on outright wins are used in the methodology. As such, more detail is provided in this section.

The earliest paper on statistical efficiency within a fixed-odds sports betting market is Pope and Peel (1989), who investigate a number of different approaches. Pope and Peel proposed a standard linear probability model, which has since been utilised in numerous studies of efficiency (for example, Schmanske, 2005 for golf; Schnytzer and Weinberg, 2007 for AFL; Lahvicka, 2014 for tennis). Pope and Peel begin with the assumption that a bookmaker will set odds reflecting his subjective probability estimate of each particular outcome and a margin on top.

To explain this, they assume that bookmakers are able to form unbiased and efficient estimates of the probabilities of outcomes so that

$$p_{ij} = p^*_{ij} + \varepsilon_{ij} \tag{2}$$

where p is the probability estimate, p^* is the true unobservable probability, ε is the estimation error and j refers to the outcome of match i. The efficiency condition is then $E(\varepsilon_{ij}) = 0$ and $E(\varepsilon_{ij} p_{ij}) = 0$.

Pope and Peel (1989) further assume that the posted odds ϕ are related to p in the manner

$$\phi_{ij} = p_{ij} + z_{ij} + t_{ij} \tag{3}$$

where t is transaction costs and z is the variable that reflects how new information affects the odds. The gross margin λ_{ij} is therefore equal to $z_{ij} + t_{ij}$. The probability estimates are therefore

$$\phi_{ij} = (p^*_{ij} + \varepsilon_{ij}) + \lambda_{ij}, \tag{4}$$

which can be expressed as the linear probability model

$$p^*_{ij} = \phi_{ij} - \lambda_{ij} - \varepsilon_{ij}. \tag{5}$$

Here, both λ_{ij} and ε_{ij} are unobservable, however p^*_{ij} can be estimated by retrieving the fitted values from

$$f_{ij} = \alpha_j + \beta_j \phi_{ij} + v_{ij} \tag{6}$$

where $f_{ij} = 1$ if outcome j of match i occurs and 0 if outcome k $(k \neq j)$ of match i occurs, α and β are regression coefficients, and v is a stochastic error term. The distribution of f_{ij} can be written

$$Pr(f_{ij} = 1) = p^*_{ij} \tag{7}$$

$$Pr(f_{ij} = 0) = (1 - p^*_{ij}) \tag{8}$$

Hence $E(f_{ij}) = p^*_{ij}$ and the model can be interpreted as generating estimates of the true probability of outcome j of match i conditional on the betting odds provided by the bookmaker. If $\beta = 1$ in Equation 6 then the bookmaker's odds are weakly efficient.

Pope and Peel (1989) acknowledge there are limitations with using ordinary least squares (OLS) in their work. Namely, that the estimated coefficient, β, would be unbiased and consistent but inefficient due to heteroskedasticity in the error term.

For this reason, Pope and Peel (1989) additionally use weighted least squares (WLS) and an approximation of the logistic regression. In their WLS model, the weights $(1/s(v_{ij}))$ are based on estimating the standard deviation of v_{ij} as $s(v_{ij}) = \hat{f_{ij}}(1 - \hat{f_{ij}})$ where $\hat{f_{ij}}$ is the fitted value obtained from the previous OLS estimate. Aitken (1936) showed that when a weighted sum of squared residuals is minimised, and each weight is equal to the reciprocal of the variance of the measurement, then the WLS estimator is the best linear unbiased estimator.

This solves the problem of heteroskedasticity but not the second problem of potentially implying probabilities which fall outside of the 0 to 1 range. To account for this, Pope and Peel (1989) employ a logit estimation of their linear probability model with the following conversions: $\alpha_{OLS} \approx 0.5 + 0.25\alpha_{LOG}$ and $\beta_{OLS} \approx 0.25\beta_{LOG}$, which are the approximate conversions between the logit and OLS models as given in Maddala (1983).

Pope and Peel (1989) examine the odds quoted by four national bookmakers for a total of 1,291 matches during the 1981 to 1982 season. They can not reject the null hypothesis of $\beta = 1$ and conclude that bookmakers set odds in a weakly efficient manner whereby implied probabilities reflect observed probabilities.

According to Forrest (2007), a significant limitation of the study by Pope and Peel (1989) is that their data only consisted of a single season. Forrest does note, however, that it is understandable given that it was a pioneering contribution, as they were the first to examine a fixed-odds sports betting market. Examining a single season is an issue because findings in sports may be unstable for various reasons. Testing over a longer period of time helps to increase robustness of conclusions by reducing the likelihood that the results occurred by pure chance.

A similar study was conducted by Kuypers (2000) who tested weak form efficiency in the English soccer betting market over the seasons 1993 to 1995 using a single bookmaker (Ladbrokes). Kuypers used a sample of 3,382 matches spanning four divisions and grouped into 24 categories based on the odds implied probability. Once again, an OLS regression was employed in the form of $Y_i = \alpha_0 + \beta_1 X_i + \varepsilon_i$, where Y is the implied probability of outcome i, X is the observed probability and ε is an error term. Kuypers was unable to reject the null hypothesis that $H_0 : \beta = 1$ at the 5% level and concluded that there was no systematic bias between implied probabilities and observed probabilities.

Many other sports have been tested for efficiency since the early studies on soccer. Cain et al. (2003) tested a range of different sports including baseball, boxing, cricket, snooker and tennis. They concluded that the FLB was evident to some extent in most of these sports, excluding baseball. However, they noted that it was unclear whether the bias is expected only at the very high and low ends of the probability spectrum, or whether returns from betting increase as a continuously monotonic function of winning probability.

The primary contribution of Cain et al. (2003) was to evaluate whether insider betting was present in these different markets. By evaluating their results based on the insider betting models by Shin (1991, 1992, 1993), Cain et al. concluded that there was probably a very small percentage of insider betting occurring in these markets; in the range of two to eight per cent.

A second wave of studies on betting market efficiency attempted to test whether fundamentals were appropriately taken into account when betting odds were priced. These involved determining whether information about the match could be used to obtain positive rates of return (for example, information on player injuries). These studies are essentially testing semi-strong form efficiency.

The potential for success in these types of studies depends on one of two things. Either the statistical model developed is superior at forecasting the result of matches, or it is able to exploit deliberate biases built into the odds. These biases may occur for strategic profit-maximising reasons to exploit cognitive biases among the general population of bettors or to account for preferences for betting on particular teams (Levitt, 2004). These studies have generally shown to yield only a limited potential for positive rates of return (Forrest, 2007).

The first attempts at 'fundamental analysis' modelled the scores of two teams in soccer as independent Poisson distributions. Dixon and Coles (1997) implement this procedure using means reflecting the goal-scoring record of the team and the goal-conceding record of the team's opponent. They generated *ex-ante* probabilities of match outcomes and compared these to bookmaker odds.

Dixon and Coles (1997) test for efficiency by placing a bet on a match outcome whenever the difference between their model's probability and the bookmaker odds implied probability varies by an arbitrary amount. While they were successful in some cases, the number of times that their model's probabilities deviated from bookmaker probabilities was very low and they conclude implementing the strategy on a practical level would prove difficult.[18]

Rue and Salvesen (2000) employ a Poisson model to assess whether they could have made positive returns from a strategy that placed bets such that expected profit was maximised subject to a constraint on its level of variance. The data spanned the top two English soccer leagues (the Premier League and the Football League Championship) during the 1997 to 1998 period and used the first half of the season to determine bets on the second half of the season. They showed this strategy to be very successful, yielding returns in the two leagues of 39.6% and 54%, with a total of 112 bets placed.

More modern and rigorous methods of fundamental analyses came about in the form of discrete choice, logit and probit models. These techniques model outcomes directly according to the probability distributions of the scores by each team. For instance, ordered logit and probit models are employed by Forrest and Simmons (2000) in a study of the efficiency of tipsters and by Koning (2000) in a paper on competitive balance; while Kuypers (2000), Dobson and Goddard (2001), Goddard and Asimakopolous (2004), and Graham and Stott (2008) develop ordered probit models to assess their efficacy as a source of profitable betting strategies.

Goddard (2005) demonstrates that the approach, of modelling results directly through such models, yields very similar results when compared with the older tradition of deriving probabilistic forecasts of outcome from estimates of probability distributions of goals scored by each team. Dobson and Goddard (2001) and Goddard and Asimakopoulos (2004) present an ordered probit forecasting model estimated from over 50,000 matches over a 10 year period. The variety of data captured by the regressors includes many variables such as league points won in the previous two seasons, the result of the most recent

[18]A similar issue was found in the horse-racing literature (Crafts, 1985).

home match, among many others.

Goddard and Asimakopoulos offer both "regression" and "economic" tests of betting market efficiency based on the results of their model estimated over 10 years and then applied to games in seasons 1999 to 2000 and 2000 to 2001. They estimate a WLS model similar to Pope and Peel (1989), adding a variable as follows:

$$P(homewin) = \alpha_0 + \beta_0 bookprobH + \delta_0(modelprobH - bookprobH) + \varepsilon_i \quad (9)$$

$$P(draw) = \alpha_1 + \beta_1 bookprobD + \delta_1(modelprobD - bookprobD) + \varepsilon_i \quad (10)$$

$$P(awaywin) = \alpha_2 + \beta_2 bookprobA + \delta_2(modelprobA - bookprobA) + \varepsilon_i \quad (11)$$

where *bookprob* refers to the bookmaker odds implied probabilities of a particular outcome (H for a home win, D for a draw, or A for an away win), *modelprob* is the probability according to the estimated forecasting model, and ε is an error term. Efficiency conditions are then $\alpha_0, \alpha_1, \alpha_2 = 0; \beta_0, \beta_1, \beta_2 = 1; \delta_0, \delta_1, \delta_2 = 0$. The coefficients δ_0, δ_1 and δ_2 will be zero if the bookmaker odds already take proper account of all the fundamental information exploited in the statistical model.

Excluding the apparent potential for employing a statistical model to secure positive returns late in the season, the literature reviewed so far has tended to find difficulty in establishing potential for using statistical modelling to secure positive as opposed to merely less negative returns. However, all these contributions relate to a period when transaction costs were higher than they are today. Transaction costs have fallen considerably over time with the rise of the Internet and increased competition (Flepp et al. 2014). This opens up the possibility for advanced statistical models to obtain positive returns.

Forrest et al. (2005) investigate whether odds were in fact set more efficiently as transaction costs fell. For the five seasons leading to 2003, they compare the forecasting performance of a richer version of the model used by Goddard and Asimakopoulos (2004) to that of a model where the only predictor variable is bookmaker odds. Results concluded that the fundamentals model is superior early in the period but inferior later on. This indicated that bookmakers responded to greater pressure to set efficient odds by adopting approaches that enabled them to do so; for example, by employing statistical modelling to supplement their intuition.

It has been found that parimutuel betting odds provide better forecasts of events than experts or statistical models for both sports and non-sports events (Wolfers and Zitzewitz, 2004). This is usually explained by the fact that closing odds reflect diverse information from large numbers of participants, induced by the prospect of financial gain to reveal their assessment of a future event (Forrest, 2007).

Fixed odds sports betting is different because odds do not change significantly as new money arrives in the market. The accuracy of odds depends solely on the expertise of odds-setters. Bookmakers have a sufficiently strong financial incentive to gather and process information accurately into the odds they post. This incentive has increased over time as transaction costs have fallen (Forrest, 2007).

Cain et al. (2000) consider just one season, 1991 to 1992, and point to superior (but still negative) returns to bets on teams that are strong favourites. However, since strong favourites are almost always playing at home, any bias they capture could in fact result from how bookmakers choose to price home advantage into the odds. By contrast, Dixon and Pope (1996) employ data from two bookmakers over 1992 to 1996 and find that backing longer odds teams yielded better results; they also report that backing home teams (or draws) yielded lower losses than betting on away teams.

Forrest (2007) argues that the FLB literature has benefited from authors using longer runs of data in more recent contributions. He asserts that even with lower transaction costs, there has been limited evidence of systematic bias that could be exploited to secure positive returns. Forrest made this assertion prior to the development and rapid expansion of online betting that has occurred in recent years. It may be the case that transaction costs have decreased enough such that positive returns can be made from simple strategies. Direr (2013) examines different European soccer leagues in the 2000 to 2011 period and finds profitable wagering opportunities by using a simple strategy of betting on heavy favourites with odds under 1.2. This suggests that either the FLB was extremely pronounced over this sample, or that transaction costs were low enough to take advantage of a standard FLB.

More recently, research efforts have attempted to explain why the FLB has been found in some sports and some leagues but not others. There have also been attempts to explain differences in the extent of the FLB. Lahvicka (2014) examined the FLB in tennis and controls for player rank, later round matches, and high-profile matches. The hypothesis was that the FLB would be greater in matches with higher importance, as better players would increase effort levels. This was tested using the model

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 \delta_1 + \beta_3 \delta_1 X_i + \beta_4 \delta_2 + \beta_5 \delta_2 X_i + \beta_6 \delta_3 + \beta_7 \delta_3 X_i + \varepsilon_i \qquad (12)$$

where Y refers to the result of match i, X refers to the odds implied probability of the outcome, δ_1 refers to whether a player was a lower rank, δ_2 refers to whether a match took place in a later round, δ_3 refers to whether the tournament was high profile (grand slam or similar), and ε is an error term.

A general critique of Lahvicka (2014) is that a linear probability model was used instead of a discrete choice model. Whilst heteroskedasticity-consistent standard errors were used, the problem of probabilities potentially lying outside of the $[0, 1]$ interval was not addressed.

Model issues notwithstanding, Lahvicka's (2014) results indicated that the extent of the FLB was stronger in later-round matches and in matches in high-profile tournaments (matches that are likely to attract higher betting volumes). On the other hand, the FLB was more pronounced in matches between lower-ranked players. Lahvicka asserts that these results are somewhat contradictory and concludes that some of the explanation may lie in the information asymmetry between bookmakers and bettors.

Lahvicka (2014) states that a possible direction for future research would be to examine if the FLB is more pronounced in higher profile matches in team sports. The claim being that the effect should be smaller in team sports because new information regarding individual members of a team (for example, sickness) have a smaller effect on the outcome of the match relative to tennis singles. This motivates one of the research questions examined in this thesis.

Another study that aims to discover what affects differences in the extent of the FLB within sports betting markets is Oikonomidis et al. (2015) who examine twenty-two soccer leagues across eleven European countries between 1999 and 2008. They find that the extent of the FLB varies greatly and they hypothesised that transaction costs and/or the extent of competitive balance could explain the differences in the extent of the FLB.

Oikonomidis et al. (2015) employ a conditional logit model with the dependent variable being binary and taking value 1 if the outcome occurs (for example, a home win) and 0 otherwise, and the lone independent variable being the (natural) logarithm of the odds implied probability. The conditional logit model has been shown in the literature to be the preferred way of detecting small deviation away from perfect calibration between implied and observed probabilities (Bacon-Shone et al., 1992; Johnson and Bruce, 2001; Sung et al., 2009). Its derivation in the context of betting markets appears in Johnson and Bruce (2001, pp.284-286) who apply it to British horse-racing. A detailed explanation is given in Section 3 when the methodology is discussed.

Oikonomidis et al. (2015) examine whether differences in the extent of the FLB can be attributed to transaction costs and/or competitiveness. They compare different specifications of their model and reject the hypothesis that transaction costs explain the extent of

FLB, as they find that it still exists when transaction costs are fully accounted for. They also compare each league's mean absolute goal difference (as a proxy for competitiveness) to the extent of the FLB and find that it is correlated. They conclude that competitiveness can explain the extent of the FLB.

Using mean-goal-difference is a relatively crude measure. It is difficult to translate across to other sports as there are differences in the way by which goals (or more generally, points) are scored. By using a standard measure of competitive balance from the sports economics literature, the conclusions of Oikonomidis et al. (2015) would hold more robustly. This serves to motivate one of the research questions in this thesis and is discussed in more detail in Section 2.2.

2.1.6 Australian Evidence

There have been only a handful of studies on betting market efficiency in Australian sports. Brailsford et al. (1995) were the first to examine the efficiency of betting in Australian sport. They looked at a novel data set consisting of parimutuel point spreads in the Australian Rugby League (ARL[19]), as well as parimutuel winning margins in the AFL. They employ an ordered probit model to test efficiency and find evidence of home bias and favourite-underdog bias in ARL. In AFL, they find profitability using a probit prediction model and betting on games where the win probability is in excess of 70% according to the model, but find no profitable simple strategies. Brailsford et al. are hesitant to conclude inefficiency in either case.

Schnytzer and Weinberg (2007) also test betting efficiency in the AFL. They examine the 2001 to 2004 period for both fixed-odds and handicap markets. They employ OLS but also use the nonparametric bootstrap method for robustness. Schnytzer and Weinberg conclude that there is no FLB in either of these markets once the possibility of a home ground advantage has been removed. Ryall and Bedford (2010) tested whether certain strategies could be used to make profit based on information at quarter-time, half-time or three-quarter time. Ryall and Bedford conclude that the team that is ahead at any time period is underbet to win, in the sense that greater returns can be made by betting on the team currently winning than the team currently losing.

Most recently, Norton et al. (2015) examined whether 'in-play' trading strategies can be profitable in One-Day International Cricket. This is arguably the strongest test of efficiency that has yet been conducted in the Australian betting literature. Norton et

[19] The NRL's inaugural season was in 1998. Prior to this it was known as the ARL.

al. construct a complex dynamic programming model to forecast the total score of the batting team during the first innings of the match as well as the probability that the chasing team will win the match. They then compare the results from this model to Betfair odds.[20] They show that the market's overreaction to the fall of early wickets provides profit opportunities greater than 20% even after commissions.

While some of these Australian studies have examined efficiency in the betting markets for AFL and NRL, there is no study that has yet used A-League data.[21] This thesis therefore provides insight into a previously untapped data source. It also updates the findings for AFL and NRL in the era of online betting. Furthermore, the Australian evidence has generally concentrated on determining whether profit opportunities in betting markets exist, rather than assessing the extent or causes of the FLB. It may be the case that certain aspects of the Australian sports being examined, such as the extent of competitive balance, can provide insight into the FLB phenomenon. This forms the final empirical analysis of this thesis. We now proceed to the final section of the literature review with an overview of measuring competitive balance in sport.

[20] Betfair is an online exchange where individual bettors are able to back or lay bets. There is no bookmaker that sets odds, however Betfair takes 5% commission from all winning bets.

[21] This is true to the best of the author's knowledge.

2.2 Measuring Competitive Balance in Sports: An Overview

This section is not directly related to the literature on betting market efficiency but offers a potential explanation of observed outcomes relating to the FLB. For an individual match, betting odds can be interpreted as a measure of match uncertainty *ex-ante*, as they assign implied probabilities to each team for a particular match. The literature on competitive balance within sports is often used to look at long term trends in competitiveness and whether structural changes (for example, regulation or institutional changes) in the sport have lead to changes in the level of balance (Fort and Maxcy, 2003).

According to Zimbalist (2002), "there are almost as many ways to measure competitive balance as there are to quantify the money supply" (p.112). This quote appropriately reflects the large number of ways competitiveness in sports can be measured. This section does not aim to cover every single measure. Rather, emphasis is placed on the most commonly used measures and some applications to the Australian sports being investigated in this thesis.

Dobson and Goddard (2011) offer a succinct overview of the various measures of competitive balance. These include the actual standard deviation (ASD); the relative standard deviation (RSD), which is the ratio of the actual and idealised standard deviations (Noll, 1988; Scully, 1989; Quirk and Fort, 1992; Fort and Quirk, 1995); the Gini coefficient (Fort and Quirk, 1995; Schmidt and Berri, 2001); the Herfindahl-Hirschman index (Depken, 1999); concentration ratios (Koning, 2000) and relative entropy (Horowitz, 1997). This list is not exhaustive but rather includes some of the most commonly used measures.

These various measures have different calculations with their own advantages and disadvantages. Fort and Lee (2007) assert that from these different measures, RSD is the most commonly used measure of within-season competitive balance in the sports economics literature. Its exact calculation will be presented in Section 3.4. Essentially, it takes a value of 1 in cases of "perfect balance" and higher values imply greater imbalance. Fort and Lee argue that it is the most useful because it controls for both season length and the number of teams.

The RSD measure has, however, been criticised in the literature for various reasons. Goossens (2006) criticises it because of the possibility that it can take values below 1. While this is rare, it implies that the league is more balanced than a perfectly balanced league; seemingly illogical. Owen and King (2015) examine the distributional properties of RSD via simulation and find that if there is imbalance in team strengths, then its distribution is sensitive to changes in season length and is therefore less useful when comparing between leagues that have different schedules.

Owen and King (2015) assert that cross-league comparisons are better suited to using a slightly modified version of RSD, which involves constructing a hypothetical "least-balanced" league. For this hypothetical league, RSD is calculated as RSD^{lb}. RSD^{*} is then calculated by taking RSD from the actual league and dividing it by RSD^{lb}. This serves to bound RSD^{*} between 0 and 1 with a value of 1 meaning that the league is as unbalanced as possible. It turns out that this figure can also be obtained by taking ASD and dividing it by ASD^{lb}, where ASD^{lb} is the ASD from the hypothetical "least-balanced" league. This enables slightly better cross-league comparison (Owen and King, 2015).

A number of studies have considered competitive balance in the context of Australian sports. Booth (2004) examines whether various labour market devices and revenue-sharing rules used in the AFL since its inception in 1897 have increased competitive balance. Booth (2005) examines the AFL, the NRL and the National Basketball League (NBL) to assess whether the introduction of player drafts and team salary caps, as well as changes in club location, have lead to an increase in competitive balance. Both of these studies examine changes in regulation over a long period of time and both use RSD to measure within-season competitive balance.[22]

Lenten (2011) examines the implications of an unbalanced schedule on various measures of competitiveness. He uses four different measures of competitive balance including RSD, the Herfindahl-Hirschman index, a concentration index of competitive balance and the Gini coefficient. Lenten proceeds to use a model to determine how these various measures are affected by changes in team scheduling.

All three of these Australian studies are assessing how the degree of competitive balance is being affected by some institutional change (or potential change). There are not many studies, Australian or otherwise, attempting to relate the level of competitive balance to the literature on betting markets.

Excluding Oikonomidis et al. (2015), who use mean-goal-difference to measure the extent of competitive balance, perhaps the closest link between these two different aspects of the economics literature is a study by Koning (2000). Koning uses an ordered probit model to forecast the probability of a team winning, using competitive balance as one of the variables in the model. The measures of competitive balance used by Koning include the standard deviation, the concentration ratio and a third measure that can be described as the sum of deviation away from average team quality. Koning's study, however, is more

[22] The distribution of premierships is used as a longer-term measure.

related to forecasting match outcomes rather than betting market efficiency, per se.

For the case of European soccer, Oikonomidis et al. (2015) find a positive relationship between the extent of the FLB and the degree of competitive balance within a league. However, using mean-goal-difference is not only a relatively crude measure, it is also difficult to translate across to other sports as there are differences in the way by which goals (or more generally, points) are scored. By using a standard measure of competitive balance, it may be possible to find evidence that supports or refutes the findings of Oikonomidis et al. Consequently this thesis considers whether the extent of competitive balance, as measured by RSD or RSD^*, can explain the extent of the FLB.

3 Methodology

This chapter outlines the various techniques used to estimate betting market efficiency. Firstly, historical bookmaker odds are compared and analysed for arbitrage opportunities. Secondly, betting simulations are backtested to determine whether traditional biases such as the favourite-longshot bias (FLB) exist. A selection of other common betting strategies are likewise backtested to determine whether positive returns based on simple betting rules would have been possible. Thirdly, the extent of inefficiency in each league is tested using a conditional logit model.

Various measures of competitiveness are calculated to determine the level of competitive balance within a league. The results from the conditional logit estimation are then compared to these measures of competitiveness to determine if a relationship exists. Each test of betting market efficiency is now outlined in detail along with corresponding hypotheses.

3.1 Arbitrage

The first and most obvious test for efficiency in betting markets is to determine whether arbitrage opportunities exist. In the context of sports betting, arbitrage would entail wagering on all possible outcomes and ensuring a risk-free profit. Arbitrage opportunities may exist if odds vary enough between different bookmakers such that a bettor could wager on different outcomes with different bookmakers and create an "underround" book. In the modern era of the Internet, numerous websites exist that serve to compare odds across a range of different online bookmakers.[23] This means that an individual bettor with multiple accounts across different bookmakers should be able to easily exploit such arbitrage opportunities if they exist.

Marshall (2009) shows that for an arbitrage opportunity to exist in betting markets, the inequality $\sum_{i=1}^{n} 1/H_i < 1$ must hold, where H refers to the highest betting odds available for outcome i and n refers to the total number of possible outcomes.

To profit from the existence of arbitrage, a bettor would wager proportion p on outcome i as follows

$$p_i = \frac{1/H_i}{1/H_1 + 1/H_2 + ... + 1/H_n}.$$

This can be broken down for the three sports covered in this thesis. For A-League football, where draws are not an uncommon occurrence, the necessary condition for arbitrage can be expressed as $(1/H_h) + (1/H_a) + (1/H_d) < 1$, where H_h refers to the highest odds available for a home win, H_a refers to the highest odds available for an away win and H_d refers to the highest odds available for a draw.

An arbitrageur wishing to profit the same amount regardless of match outcome would therefore stake $p_i = \dfrac{1/H_i}{(1/H_h) + (1/H_a) + (1/H_d)}$ where i refers to the outcome and takes the value h, a or d.

For AFL and NRL the necessary condition for arbitrage is the same without the draw component. However, it does depend on whether bookmakers follow what is known as the "dead-heat" rule. This rule states that in the rare event where a game ends in a draw and a draw outcome is not offered in the 'win' market, then wagers are paid at the face value of the ticket divided by the number of teams in the event.[24] For example, a \$1 wager placed at odds of \$1.60 on a match that results in a draw would be calculated as

[23] The website www.oddsportal.com is one such example.

[24] This is up to the discretion of the bookmaker. Contact was made with ten of the bookmakers in the data set and all confirmed this to be the case.

$\$1 \times \$1.60 \times 0.5 = \$0.80$. Schnytzer and Weinberg (2007) also describe this rule.[25]

Now, suppose the opposition in this match has odds of \$2.80. For the hypothetical arbitrageur betting \$100, we can calculate the wager on each outcome necessary to obtain a risk-free profit. We begin by calculating the bookmaker margin: $(1/1.60) + (1/2.80) = 0.9821$. By staking \$63.64 on the favourite and \$36.36 on the underdog the bettor gains $\$63.64 \times \$1.6 = \$101.82$ if the favourite wins and $\$36.36 \times \$2.8 = \$101.81$ if the underdog wins. This equates to \$1.82 and \$1.81 profit, respectively. In the event of a draw in this case, assuming the dead-heat rule applies, the bettor receives $(63.64 \times 1.6 \times 0.5) + (36.36 \times 2.8 \times 0.5) - 100 = \$1.82$. Regardless of the outcome, the bettor makes at least \$1.81 profit.

Thus, assuming the dead-heat rule applies, for AFL and NRL the arbitrageur does not need to wager on draws as well. As such, the necessary condition for arbitrage can be simplified to $(1/H_h) + (1/H_a) < 1$ and the arbitrageur stakes $p_i = (1/H_i)/[(1/H_h) + (1/H_a)]$.

As exploitable arbitrage opportunities would provide evidence of market inefficiency, we examine the data to find out if this was ever the case. This gives rise to the following hypothesis:

H1: No arbitrage opportunities exist.

Results related to this hypothesis are presented in Section 4.1.

[25]Footnote 18 of their paper.

3.2 Betting Simulation

We once again make the distinction between weak-form and strict weak-form efficiency. The former requires that no betting strategy is possible based on past price levels and movements that can yield a positive expected return. The latter requires that betting at particular odds, or ranges of odds, yields identical returns to betting at any other odds or ranges of odds.

To test strict weak-form efficiency, we simulate equal staking on each odds range. Firstly, the odds are categorised into two groups according to whether they corresponded to the favourite or the underdog for that particular match.[26] Then, each of these groups are split into deciles according to a grouping algorithm that results in slightly uneven groups but containing as much similarity as possible.[27] For strict weak-form efficiency to hold, every odds range must have the same return.

We also back-test simple strategies that have, at various times, been shown to yield positive rates of return in the literature. These strategies include:

(1) $Betting\ on\ home\ teams$,

(2) $Betting\ on\ away\ teams$,

(3) $Betting\ on\ favourites$,

(4) $Betting\ on\ underdogs$,

(5) $Betting\ on\ home\ underdogs$,

(6) $Betting\ on\ home\ favourites$,

(7) $Betting\ on\ away\ underdogs$,

(8) $Betting\ on\ away\ favourites$ and

(9) $Betting\ on\ draws$ (in A-League only).

For strict weak-form efficiency to hold, all of these strategies should yield the same return. For weak-form efficiency to hold, none of the strategies should yield a positive rate of return.

It is important to note that only a small number of simple strategies have been considered here. The total number of possible strategies that could be simulated is vast and dependent upon the information set used. These particular strategies are selected because the information set required to implement them is small, and some of them have shown

[26]This is done because sometimes odds for favourites and underdogs can not be distinguished by value alone. For example, there are matches where an underdog has odds of 1.87 and the corresponding favourite has odds of 1.85, and other times when the favourite has 1.87 and the underdog has 1.89. This discrepancy is caused by variance in bookmaker overround between matches.

[27]Using the quantile function from the stats package in R.

to yield positive returns in the literature.

We have chosen not to test for the “hot-hand” or the “gambler’s fallacy” as the selection of n (number of games on a winning or losing streak) is arbitrary and the number of strategies rapidly rise. Furthermore, as the number of strategies increase, the probability of one of them being profitable due to random chance also increase. This problem could also arise in the strategies that are being tested and we must therefore be careful in the interpretation of any strategy that does yield a positive return.

The final betting simulation involves isolating finals (playoff) matches to determine if the results for betting on favourites or underdogs differs when there is a greater incentive to win. Upsets may be less likely in a two-way contest if both teams are trying their best. This would extend the findings of Lahvicka (2014) who found that the FLB was relatively more pronounced in high-profile tennis matches.

Evidence that strategies could be implemented to obtain positive or non-constant returns would provide support for inefficiency. This gives rise to the following hypotheses:

> **H2a:** For strict weak-form efficiency, all strategies and odds ranges yield the same return.

> **H2b:** For weak-form efficiency, no strategy or odds range yields a positive rate of return.

> **H2c:** Returns to betting on favourites exceed returns to betting on underdogs by a greater amount in finals matches relative to regular season matches.

Results related to these hypotheses are presented in Section 4.2.

3.3 Statistical Test of Efficiency

3.3.1 Discussion of Potential Models

An alternate test of betting market efficiency is to employ a statistical model that examines whether betting odds are unbiased estimators of observed probabilities. As discussed in the literature review, there have been numerous approaches; some more appropriate than others. A few alternatives are briefly discussed here and then the final model is presented.

Most recent studies of betting market efficiency have recognised the weakness of Pope and Peel's (1989) OLS and WLS models when dealing with data aimed at providing probability estimates. As such, the literature has shifted towards the use of discrete choice models.

A question that arises immediately is which of these is best suited for the task? In the context of sports betting, one must look at the total number of possible outcomes. This brings us to the general case of three; the result is either a home-win, an away-win or a draw. One might assume that a multinomial model is the obvious choice. However, this is not the best option.

Due to the nature of draws being random, unpredictable events, it is better to model the outcomes individually (Forrest, 2005). The reason being is that one outcome might be perfectly priced by bookmakers whilst another outcome could be greatly mispriced. Deschamps and Gergaud (2007) test efficiency in the English Premier League and find that there is virtually no relationship between the draw odds and the probability of a draw outcome. They state that this anomaly is "unique to draw odds, since home and away odds are strongly correlated with the probability of home and away win" (p.67). For this reason, a binary model is preferred even when draw outcomes are relatively common, as in the A-League.

For AFL and NRL, draws are so rare and unpredictable that the draw outcome in a multinomial model would lack predictive power. Schnytzer and Weinberg (2007) test betting market efficiency in the AFL and state that draws are extremely rare, occurring less than 1% of the time and only available in the exotic bets section at very high odds. The same applies for NRL markets and in both cases the dead-heat rule applies regardless.

The question then becomes whether one should use logit or probit in the binary case. Cameron and Trivedi (2005) provide an overview of which should be preferred and why, which will be briefly summarised here.

The advantage of using a logit or probit model is that estimated probabilities are guaranteed to be between 0 and 1. This may not be the case in a linear probability model, especially for low and high probabilities. Furthermore, using maximum log-likelihood estimation to estimate the model eliminates the problem of heteroskedasticity. If we consider the dependent variable as binary, then

$$Y_i = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1-p. \end{cases}$$

As well as the lone independent variable, x_i, which in our context is the (natural) log of the betting-odds-implied-probability. The logit model in the form of the log odds ratio can be written as

$$log\frac{p}{1-p} = \beta_0 + \beta_1 x_i.$$

Converted to probability, is expressed as

$$p_i = Pr(Y_i = 1|x_i) = \frac{e^{(\beta_0+\beta_1 x_i)}}{1+e^{(\beta_0+\beta_1 x_i)}}.$$

It thus ensures that the probability estimates lie between 0 and 1. The implied marginal effect for the logit model is then

$$\frac{dp_i}{dx_i} = \frac{e^{(\beta_0+\beta_1 x_i)}}{1+e(\beta_0+\beta_1 x_i)}\beta_1.$$

The probit model on the other hand specifies

$$p_i = Pr(Y_i = 1|x_i) = \Phi(\beta_0 + \beta_1 x_i)$$

where $\Phi(.)$ is the cumulative distribution function for the standard normal density. The implied marginal effect is then

$$\frac{dp_i}{dx_i} = \phi(\beta_0 + \beta_1 x_i)\beta_1$$

where $\phi(.)$ is the density for the standard normal. In both the logit and the probit $dp_i/dx_i < 0$, however the coefficients cannot be compared directly.

Cameron and Trivedi (2005) assert that the correct choice of logit or probit depends on the underlying data generating process, but that empirically there is very little difference in the predicted probabilities between them. They state that for slope parameters, if $0.1 \leq p \leq 0.9$, then a good approximation is to use $\hat{\beta}_{Logit} \approx 1.6\hat{\beta}_{Probit}$. At either end of the probability range however, is where the departures are larger. They note that the logit model is usually preferred because it corresponds to the use of the canonical link for the binomial distribution (p.285). Furthermore, the interpretation of coefficients in terms of the log-odds ratio is quite intuitive.

Bacon-Shone et al. (1992) assess a number of different calibrations of the logit model and conclude that the conditional logit is the preferred method to determine the extent of the FLB in horse-racing.[28] It has been shown in the literature to be an ideal way of detecting small deviation away from perfect calibration between implied and observed probabilities (Bacon-Shone et al., 1992; Johnson and Bruce, 2001; Sung et al., 2009).[29] The derivation of the conditional logit model in the context of betting markets appears in Johnson and Bruce (2001, pp.284-286) who apply it to British horse-racing. A detailed explanation of the model in the context of the three sports being examined is given next.

[28] Bacon-Shone et al. (1992) refer to this as the "constant-β model" and find it has the highest log-likelihood ratio of the different calibrations of the logit model they consider (p.12).

[29] A few authors have utilised the standard binary logit model (for example, Koning, 2012 and Nyberg, 2014 for purposes of comparison only). However, it is less effective at detecting the extent of the FLB.

3.3.2 The Favourite-Longshot Bias in the Conditional Logit Model

The dependent variable, Y, is the match outcome and takes the value 1 for the event that occurred and 0 for the event(s) that did not occur. The probability that outcome j in match i occurs can be written as

$$P(Y_{ij} = 1) = \frac{e^{p_{ij}}}{\sum_{j=1}^{n} e^{p_{ij}}} \tag{13}$$

where p is the average-bookmaker-implied-probability (normalised for overround) and n is the number of outcomes. Bruce and Johnson (2001) state that it can be shown via McFadden (1974) that if the error terms obtained from estimating Equation 13[30] are assumed to be independent and identically distributed according to the double exponential (Laplace)[31] distribution, then the probability of outcome i occurring in match j is given as

$$p^*_{ij} = P(Y_{ij} = 1|p_{ij}) = \frac{e^{\beta p_{ij}}}{\sum_{i=1}^{n} e^{\beta p_{ij}}} \tag{14}$$

where p^* refers to the observed probability of the outcome, n refers to the number of outcomes and takes the value 2 in AFL and NRL, and 3 in A-League. Taking the (natural) logarithm of p_{ij} enables us to rewrite the model as

$$p^*_{ij} = P(Y_{ij} = 1|lnp_{ij}) = \frac{p^{\beta}_{ij}}{\sum_{i=1}^{n} (p_{ij})^{\beta}}. \tag{15}$$

The parameter β is determined by maximising the joint probability of observing the results of all the matches in the sample. That is, we find β which maximises the likelihood function L, such that

$$L = \prod_{j=1}^{m} p^*_{ij} \tag{16}$$

where m is the total number of matches. It can be interpreted as follows: if $\beta = 1$, then for the match outcome under investigation, the log of the betting-odds-implied-probability is the same as the log of the odds-ratio of success for that outcome.

If $\beta < 1$, then the results are indicating a reverse-FLB, meaning that implied-probabilities

[30] Notation for Equation 13 was obtained and modified from Grant et al. (2015).

[31] This assumption implies that if the estimate is taken from a large sample, then β is approximately normally distributed with a mean of β and a standard deviation which can be consistently estimated by the standard errors obtained from maximum likelihood estimation (Johnson and Bruce, 2001, p.286).

on longshots are underestimated and/or implied-probabilities on favourites are overestimated. If $\beta > 1$, then the results are indicating the standard FLB, meaning that implied probabilities on longshots are overestimated and/or implied-probabilities on favourites are underestimated.

This can be demonstrated with a simple example. Consider the case of a moderate favourite with betting odds of 1.5. Assume for simplicity that overround = 0. The implied probability of success = 1/1.5 = 0.67. In the case of $\beta = 1$ we have $\frac{0.67^1}{0.67^1 + 0.33^1} = 0.67$. That is, the observed probability of success is the same as the implied probability of success. If $\beta = 0.8$ we have $\frac{0.67^{0.8}}{0.67^{0.8} + 0.33^{0.8}} = 0.638$. This means that the observed probability of success is lower than the implied probability of success. Likewise, if $\beta = 1.2$ we have $\frac{0.67^{1.2}}{0.67^{1.2} + 0.33^{1.2}} = 0.701$. This means that the observed probability of success is higher than the implied probability of success.

This example can be generalised as follows:

$$\text{if } \beta > 1 \text{ then } \frac{p_i^\beta}{p_i^\beta + (1 - p_i^\beta)} > p^*, \qquad \text{(standard FLB)}$$

$$\text{if } \beta = 1 \text{ then } \frac{p_i^\beta}{p_i^\beta + (1 - p_i^\beta)} = p^* \text{ and} \qquad \text{(no FLB)}$$

$$\text{if } \beta < 1 \text{ then } \frac{p_i^\beta}{p_i^\beta + (1 - p_i^\beta)} < p^* \qquad \text{(reverse-FLB)}$$

where p refers to the bookmaker implied probability for the favourite of outcome i and p^* refers to the observed probability. Finding statistical evidence that bookmaker odds do not reflect observed probabilities would provide strong evidence that the markets are inefficient. Once again, we isolate finals (playoff) matches to determine if the results differ when incentives are greater. This gives rise to the following hypotheses:

H3a: For efficiency, $\beta = 1$ for all sports and all outcomes. This implies no FLB.

H3b: The FLB is more pronounced in finals matches.

Results related to these hypotheses are presented in Section 4.3.

3.4 Relative Standard Deviation

We now outline the measure of competitiveness and the calculations before providing a hypothesis of the expected relationship between competitiveness and the extent of the favourite-longshot bias (FLB).

The level of league competitiveness is measured for each season using the Relative Standard Deviation (RSD) measure, which is calculated by dividing the Actual Standard Deviation (ASD) over the Idealised Standard Deviation (ISD). ASD can be calculated as

$$ASD = \sqrt{\sum_{i=1}^{N} [p_i - \bar{p}]^2/(N-1)}$$

where N is the number of teams in the league and $p_i = P_i/T_i$ which is defined as the team's points ratio. P_i and T_i are, respectively, the actual number of points accumulated and the maximum possible points that could be obtained if the team won every match, $\bar{p} = \sum_{i=1}^{N} p_i/N$ and refers to the league's mean points ratio. If draws are not possible or are worth half as much as a win (as in AFL and NRL), then $\bar{p} = 0.5$.

ISD depends on the structure of the league. Owen (2012) discusses the calculation of ISD when draws are possible. Owen argues that the standard calculation[32] is inappropriate and offers alternative calculations that are based on the (historical) probability of a draw and the structure of the points system (whether wins are 2 points or 3 points).[33] The modified calculations are used in this thesis.

For AFL and NRL, where draws are worth half a win, $ISD = \sqrt{(1-d)/4t}$, where d is the implied probability of a draw based on historical outcomes and t is the number of rounds.[34] For A-League, where draws are worth a third of a win, $ISD = \sqrt{(1-d)(d+9)/4t}$.

A significant problem arises when trying to compare RSD values across different leagues or different sports (Owen, 2010). This is due to the differences in league structure, for instance, differences in number of teams or number of matches that teams play. As such, Owen (2010) recommends the use of an alternative measure, RSD^*, which involves calculating a hypothetical "least-balanced" league. For this hypothetical league, RSD is calculated as RSD^{lb}.

RSD^* is then calculated by taking RSD from the actual league and dividing it by

[32] $ISD = \sqrt{1/4t}$ where t is the number of rounds in the season.

[33] In AFL, wins are worth 4 points and ties are worth 2 points. Owen (2012) shows that the effect on RSD is the same, as long as draws are worth half a win.

[34] When d is extremely low (as in AFL and NRL) then modified ISD and regular ISD are almost identical.

RSD^{lb}. This serves to bound RSD^* between 0 and 1 with a value of 1 meaning that the league is as unbalanced as possible.

To assess whether a relationship exists between competitiveness and the direction or extent of the FLB, the results from the conditional logit model will be compared to the results from the measures of competitiveness. This gives rise to the following hypothesis:

> **H4:** Leagues with lower competitiveness are associated with a greater degree of FLB.

Results related to this hypothesis are presented in Section 4.4.

3.5 Summary of Methodology

To reiterate, first bookmaker odds are compared and assessed for arbitrage opportunities. For efficiency, we expect no arbitrage opportunities to exist.

Secondly, betting simulation based on historical odds is undertaken to assess the performance of various strategies. For weak-form efficiency, we expect no strategy or odds range to yield a positive rate of return. For strict weak-form efficiency, we expect all strategies and odds ranges to yield the same rate of return.

Thirdly, conditional logit estimation is employed to assess the degree of the favourite-longshot bias (FLB) in each sport. For efficiency, we expect that there is no FLB, meaning that bookmaker odds reflect actual observed probabilities.

Fourthly, the results for finals matches are isolated to determine if the apparent incentive difference changes the results. Betting simulation and conditional logit estimation is undertaken on these matches. The hypothesis is that the FLB is more pronounced in finals matches as upsets are less likely.

Finally, measures of league competitiveness are calculated for each sport. These are then compared to the extent of the FLB from the conditional logit estimation. The hypothesis is that leagues with lower competitiveness are associated with a higher degree of FLB. All of the hypotheses being tested are listed in Table 1.

Table 1: Summary of Hypotheses

Section	Hypothesis	Expected Result
Arbitrage	H1	No arbitrage opportunities.
Betting Simulation	H2a	All strategies and odds ranges yield the same return.
Betting Simulation	H2b	No strategy or odds range yields a positive return.
Betting Simulation	H2c	Returns to bets on favourites are greater in finals matches.
Statistical Testing	H3a	$\beta = 1$ for all outcomes. This implies no FLB.
Statistical Testing	H3b	The FLB is more pronounced in finals matches.
Competitive Balance	H4	Lower competitiveness is associated with higher FLB.

3.6 A Description of the Data

The data being analysed can be categorised into two sections, betting odds and sport statistics. Historical betting odds for all AFL, NRL and A-League matches between between 2010 to 2014 (2010/11 to 2014/15 for A-League) were gathered from the bookmaker comparison website *www.oddsportal.com.* The full range of betting odds available from the source is used (excluding the 2009 season as it was incomplete). Between 11 and 55 bookmakers are available, depending on the match. The website is free to use, however, an account is required to obtain information from all bookmakers in the database. Without an account, the user is restricted in the number of bookmakers they can compare. A partial screenshot of the website with the list of bookmakers and their respective odds is provided in Appendix A.1.

League tables used to calculate the different measures of competitiveness were obtained from various websites depending on the sport.[35] Actual league tables were used to calculate RSD. The hypothetical "least-balanced" league tables were constructed to calculate the highest possible value for RSD, RSD^{lb}. The construction of these hypothetical tables varies depending on the match scheduling system. For four of the five A-League seasons, there are ten teams and each team plays every other team three times. This results in twenty-seven total games per team. Team 1 has twenty-seven wins, Team 2 has twenty-four, Team 3 has twenty-one, and so forth. The fifth A-League season differs slightly as there are eleven teams.

For AFL and NRL, the calculation differs due to unbalanced schedules. That is, some teams play each other once and some teams play each other twice. For this reason, teams around the middle of the hypothetical table may have their ordering changed based on who they play and how many times. An example of one of the hypothetical league tables used to construct RSD^{lb} is provided in Appendix A.2. We now proceed to examine the results of the analysis.

[35]http://afltables.com/afl/afl_index.html for AFL;
http://afltables.com/rl/rl_index.html for NRL;
http://www.ultimatealeague.com/standings.php for A-League.

4 Results

This chapter provides results for the different tests of betting market efficiency discussed in Chapter 3. All data analysis and estimation was undertaken using R version 3.2.0.

4.1 Arbitrage

This section deals with arbitrage opportunities between bookmakers. All analysis is based on the "best odds" available across all bookmakers. Many arbitrage opportunties were uncovered for each sport. Arbitrage opportunities were evident in 30.67%, 59.90% and 59.94% of matches for AFL, NRL and A-League, respectively.

The extent of arbitrage can be visualised by plotting bookmaker overround for each match. Figure 2 consists of scatter plots showing the extent of overround for each sport. These are accompanied by descriptive statistics of the overround in Table 2.

Table 2: Bookmaker Overround — Descriptive Statistics

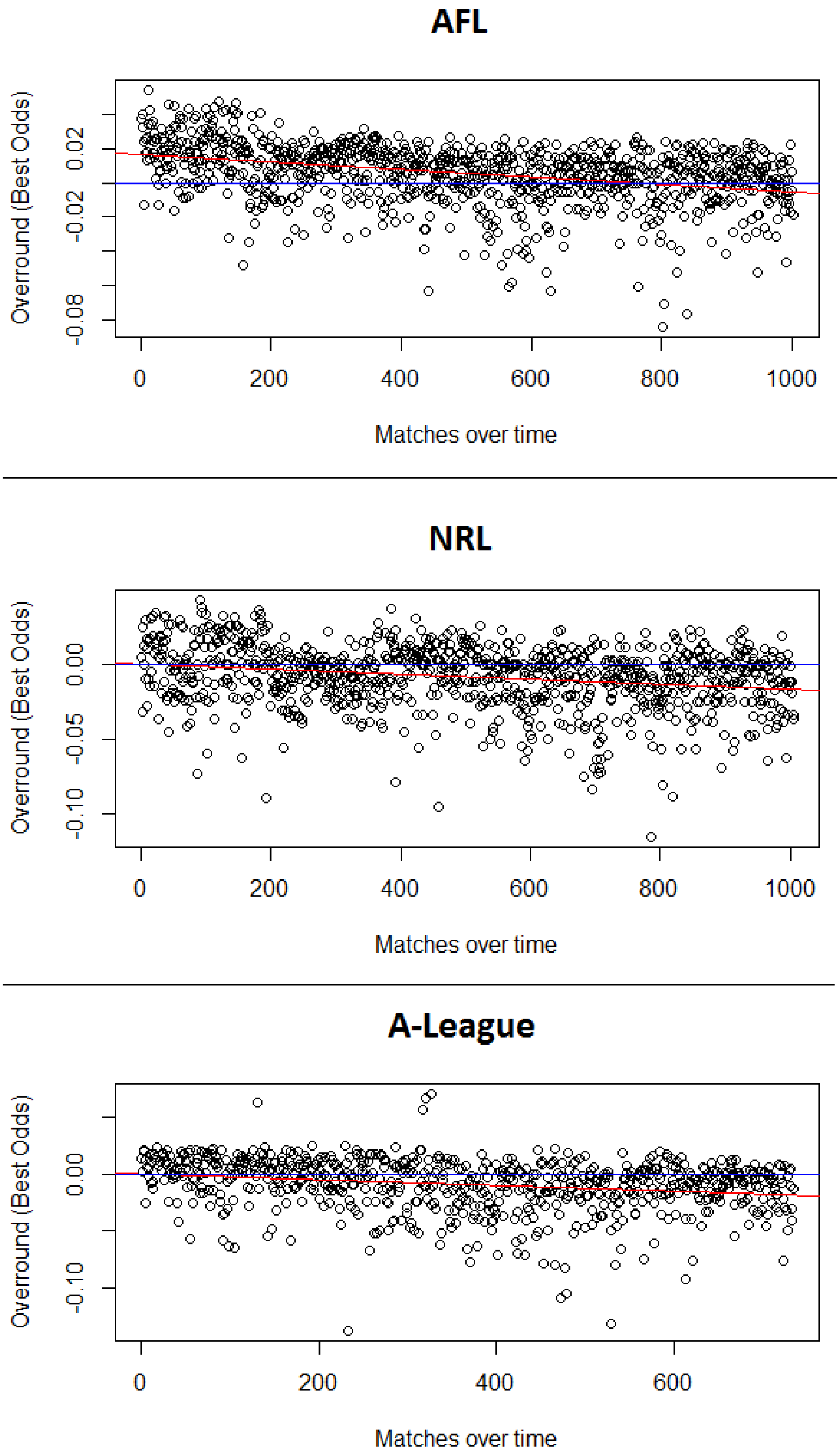
	Sport	Matches (No.)	Mean	Median	Min	Max
All Matches	AFL	1001	0.005	0.008	-0.084	0.054
	NRL	1005	-0.008	-0.005	-0.115	0.043
	A-League	734	-0.009	-0.005	-0.138	0.071
Matches With Arbitrage	AFL	307	-0.015	-0.010	-0.084	-0.000
	NRL	602	-0.020	-0.015	-0.115	-0.000
	A-League	440	-0.022	-0.016	-0.138	-0.000

By examining Table 2 and looking at 'all matches', we can see that the mean and median overround is very close to zero for all three sports. For AFL, both the mean and median are slightly positive, while for NRL and A-League they are both slightly negative. By looking at matches where arbitrage opportunities were identified, the median values between one and two per cent suggest frequent and large pricing differences between bookmakers.

The minimum values refer to matches which would have generated the highest rates of return from arbitrage. We can see that these values are -0.084, -0.115 and -0.138 for AFL, NRL and A-League, respectively, suggesting risk-free returns in excess of 10%.[36]

[36]In the initial data set there were a few (approximately 10) outliers where the overround took very low values (between -20 and -40), suggesting extremely large arbitrage opportunities of up to 40%. This is indicative of either a severe mispricing for that match or a problem with the data itself. An examination

Figure 2: Scatter Plots of Potential Arbitrage Opportunities



All points under the blue lines (fixed at 0) are matches with arbitrage opportunity. The red lines are trend lines and show that overround for best odds has slightly decreased over time. However, this apparent trend must be interpreted carefully. Whilst it may appear that the market has become more inefficient over time, the number of bookmakers available in the data set increased over time. The average number of bookmakers per match in 2010 was just 13 whilst the average for 2014 was 35. This could instead mean that a higher number of bookmakers results in more arbitrage opportunities.

The apparent existence of many arbitrage opportunities with a non-trivial rate of return is intriguing. If markets are informationally efficient, then systematic mispricings of this size should not occur. In the recent literature on arbitrage within betting markets, two main reasons are given for this apparent mispricing.

Franck et al. (2013) examine arbitrage within the sports betting markets for the top five European soccer leagues (the English Premier League, the French Ligue 1, the German Bundesliga, the Italian Serie A and the Spanish Primera Division) and similarly find that arbitrage opportunities existed for 19.2% of matches. They explain this apparent inefficiency by stating that bookmakers are behaving strategically with regards to long term profit. They claim that this is because they are attempting to bring in new customers with favourable odds and are therefore prepared to make a short term loss on them.

Grant et al. (2015) likewise look into arbitrage opportunities in different European soccer leagues and find that arbitrage opportunities arise in 25.6% of matches (comparing 6 bookmakers). They conclude that mispricings are unlikely to be systematically exploitable as bookmakers may impose restrictions on the accounts of skilled bettors.

Whilst apparent arbitrage opportunities existed for AFL, NRL and A-League, it is difficult to determine whether they would have been systematically exploitable. If they were exploitable, it is likely that this is due to strategic bookmaker behaviour as Franck et al. (2013) and Levitt (2004) suggest.

of the outliers suggested that in all of these cases it was almost certainly a data issue, whereby the odds were reversed between teams for one bookmaker. These matches were removed from the analysis.

4.2 Betting Simulation

Betting simulation was undertaken to determine whether systematic mispricings were occurring that could have been exploited by a simple strategy. The simple strategies considered are listed in Table 3.

Table 3: Returns for Various Betting Strategies

Sport	Strategy	Bets (No.)	Win (%)	Return/Wager (Best odds)	Return/Wager (Avg odds)
AFL	Home Team	1001	56.73	1.09	1.00
	Away Team	1001	42.37	0.98	0.88
	Favourite	1001	70.19	0.97	0.93
	Underdog	1001	28.91	1.14	0.98
	Home Favourite	588	73.30	0.98	0.95
	Home Underdog	413	32.93	1.23	1.07
	Away Favourite	413	67.07	0.95	0.90
	Away Underdog	588	26.70	1.07	0.91
NRL	Home Team	1005	57.91	1.06	0.99
	Away Team	1005	41.69	1.03	0.93
	Favourite	1005	63.08	0.98	0.94
	Underdog	1005	36.52	1.13	1.01
	Home Favourite	653	66.46	1.00	0.96
	Home Underdog	349	42.12	1.19	1.07
	Away Favourite	349	57.88	0.95	0.90
	Away Underdog	653	33.54	1.09	0.97
A-League	Home Team	734	46.73	1.06	0.98
	Away Team	734	28.61	1.04	0.92
	Favourite	734	50.41	1.07	0.97
	Underdog	734	25.34	1.05	0.93
	Home Favourite	566	51.06	1.07	0.97
	Home Underdog	168	32.14	1.09	0.99
	Away Favourite	168	44.64	1.10	0.98
	Away Underdog	566	24.20	1.03	0.91
	Draw	734	24.66	0.92	0.84

Strategies that would have yielded a positive return with average odds are shaded in grey. These strategies would have yielded positive returns even without shopping around for better odds. Betting on the home-underdog in AFL and NRL would have been the most successful strategy yielding a return of 7.2%. Betting on every home team in AFL or every underdog in NRL would have likewise resulted in a positive return, however these are only slightly above 1.00 in both cases. This suggests that for AFL and NRL over this time period, home underdogs were more successful than expected based on the odds implied probability.

To examine whether systematic mispricings occurred based solely on the value of the odds, odds were divided into favourite and underdog categories and then further divided into deciles based on the implied probability of success. These results are reported in Tables 4, 5 and 6, and are depicted graphically in figures 3, 4 and 5.

Table 4: Returns to Bets at Different Odds — AFL

Odds Range	Bets (No.)	Win (%)	Return/Wager (Best odds)	Return/Wager (Avg odds)
1.01 - 1.04	109	92.66	0.97	0.95
1.04 - 1.11	95	88.42	0.98	0.96
1.11 - 1.18	113	79.65	0.95	0.92
1.18 - 1.25	92	76.09	0.96	0.92
1.25 - 1.32	102	64.71	0.87	0.83
1.32 - 1.40	92	63.04	0.90	0.86
1.40 - 1.48	102	66.67	1.01	0.96
1.48 - 1.60	102	62.75	1.01	0.97
1.60 - 1.72	95	55.79	0.97	0.92
1.72 - 1.88	101	56.44	1.07	1.01
1.87 - 2.07	103	48.54	1.01	0.95
2.07 - 2.28	108	43.52	1.03	0.95
2.28 - 2.55	91	36.26	0.97	0.87
2.55 - 2.85	100	36.00	1.08	0.99
2.85 - 3.26	100	34.00	1.13	1.03
3.26 - 3.90	103	33.98	1.36	1.22
3.90 - 4.68	98	21.43	1.02	0.89
4.68 - 6.01	101	17.82	1.07	0.92
6.01 - 9.69	101	12.87	1.17	0.94
9.69 - 24.7	98	8.16	1.52	1.03
5.61 - 10.4	70	15.71	1.46	1.18

Figure 3: Returns to Bets at Different Odds — AFL

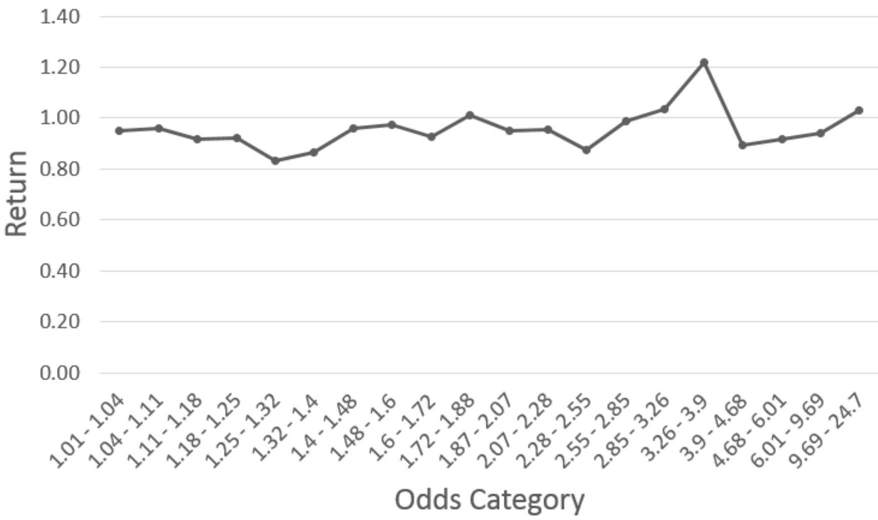


Table 5: Returns to Bets at Different Odds — NRL

Odds Range	Bets (No.)	Win (%)	Return/Wager (Best odds)	Return/Wager (Avg odds)
1.02 - 1.22	101	78.22	0.93	0.90
1.22 - 1.32	101	68.32	0.91	0.87
1.32 - 1.39	124	67.74	0.96	0.92
1.39 - 1.44	95	68.42	1.02	0.98
1.44 - 1.48	92	63.04	0.97	0.93
1.48 - 1.55	106	58.49	0.93	0.89
1.55 - 1.61	95	57.89	0.97	0.92
1.61 - 1.68	91	54.95	0.96	0.91
1.68 - 1.77	100	64.00	1.17	1.10
1.77 - 1.91	100	52.00	1.02	0.95
1.90 - 2.06	123	49.59	1.05	0.99
2.06 - 2.18	78	44.87	1.02	0.96
2.18 - 2.31	121	39.67	0.97	0.90
2.31 - 2.45	80	37.50	0.97	0.90
2.45 - 2.63	103	41.75	1.15	1.06
2.63 - 2.76	108	36.11	1.06	0.98
2.76 - 2.99	93	37.63	1.21	1.08
2.99 - 3.36	98	27.55	0.99	0.87
3.36 - 4.13	101	29.70	1.25	1.09
4.13 - 14.4	100	23.00	1.61	1.26

Figure 4: Returns to Bets at Different Odds — NRL

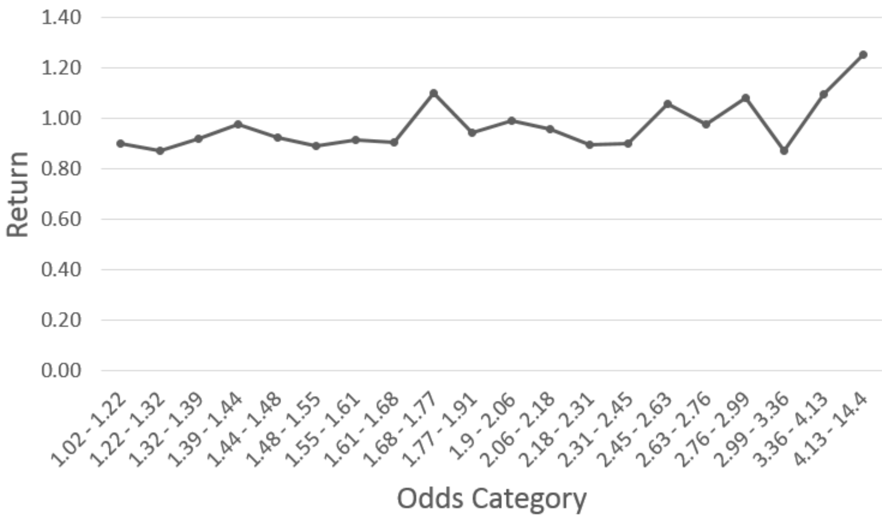
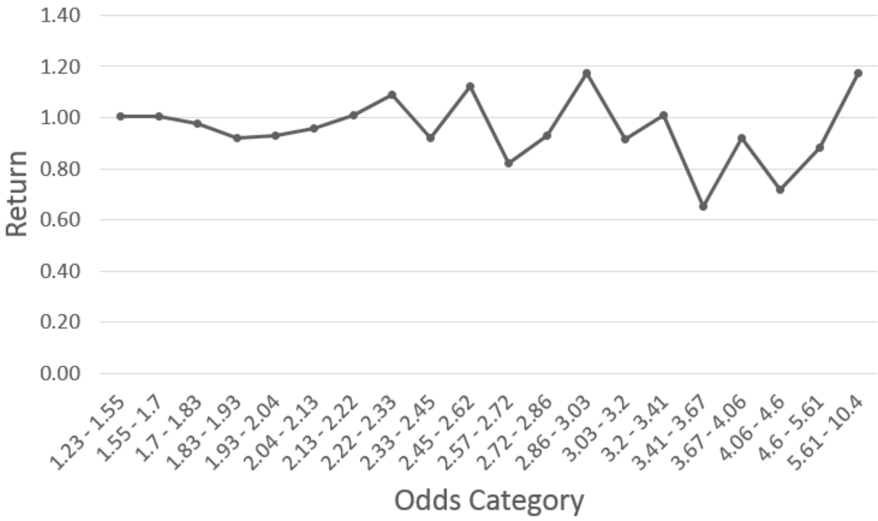


Table 6: Returns to Bets at Different Odds — A-League

Odds Range	Bets (No.)	Win (%)	Return/Wager (Best odds)	Return/Wager (Avg odds)
1.23 - 1.55	76	68.42	1.06	1.00
1.55 - 1.70	73	61.64	1.07	1.00
1.70 - 1.83	76	55.26	1.05	0.98
1.83 - 1.93	83	48.19	0.98	0.92
1.93 - 2.04	67	46.27	1.01	0.93
2.04 - 2.13	69	46.38	1.05	0.96
2.13 - 2.22	71	46.48	1.10	1.01
2.22 - 2.33	76	47.37	1.18	1.09
2.33 - 2.45	73	38.36	1.01	0.92
2.45 - 2.62	70	44.29	1.25	1.12
2.57 - 2.72	85	30.59	0.89	0.82
2.72 - 2.86	70	32.86	1.03	0.93
2.86 - 3.03	66	39.39	1.31	1.17
3.03 - 3.20	78	29.49	1.04	0.91
3.20 - 3.41	76	30.26	1.14	1.01
3.41 - 3.67	71	18.31	0.75	0.65
3.67 - 4.06	68	23.53	1.05	0.92
4.06 - 4.60	73	16.44	0.83	0.72
4.60 - 5.61	73	17.81	1.07	0.88
5.61 - 10.4	70	15.71	1.46	1.18

Figure 5: Returns to Bets at Different Odds — A-League



The data suggest that most odds categories had a negative return. For perfect efficiency, we would expect an almost horizontal line equal to 1 minus the bookmaker overround.[37] This would mean that all bets had a return of (1-overround) $\times$ 100%. Recall from the literature review the figure sourced from Snowberg and Wolfers (2010). Their graph indicated a monotonically decreasing rate of return as odds increased. The above graphs do not appear as such. The variance of returns appears slightly larger as odds increase, particularly in the A-league. It is, however, difficult to spot a clear pattern that would indicate an odds bias, with the exception of perhaps very high odds having positive returns, possibly indicative of a reverse-FLB.

This is in contrast to most of the previous literature on the FLB, which has generally found that favourites have higher returns than longshots, however it is not the first evidence of a reverse result. For example, Woodland and Woodland, 1994, find a reverse-FLB in baseball. Testing the extent of this potential bias is undertaken statistically in the next section.

All three sports under examination follow a relatively similar finals series system. At the end of the regular season, a certain number of the top teams enter a playoff competition to determine the overall winner. One hypothesis is that upsets are less likely to occur because of the increased incentive to win. This is because both teams are expected to play their absolute best as there is more to be gained from winning. Results for finals matches are provided in Table 7.

In finals matches, for all three sports, returns to bets on favourites greatly exceeded returns to bets on underdogs. For AFL, betting on the favourite yielded a positive return with average odds, but a negative return for underdogs even with best odds. For NRL, returns were negative for both favourites and underdogs even with best odds, but were less negative for favourites.

For A-League finals matches, betting on favourites yielded a positive return only with best odds, while betting on draw or underdog yielded the worst return of any strategy by a large margin. For every dollar wagered on an underdog in the A-League for a finals match, a bettor would have received \$0.22 with average odds or \$0.26 with best odds, resulting in a loss of \$0.77 (or \$0.74 with best odds) per bet. This suggests that over this time period, underdogs won finals matches in A-League less than the odds implied, however the sample size is quite small and this result could potentially change over a longer time period.

[37]The line would not be perfectly horizontal because overround is not constant.

Table 7: Betting Simulation — Finals Matches

Sport	Strategy	Bets (No.)	Win (%)	Return/Wager (Best odds)	Return/Wager (Avg odds)
AFL	Favourite	46	69.57	1.09	1.05
	Underdog	46	30.43	0.85	0.81
NRL	Favourite	54	72.22	0.96	0.91
	Underdog	54	27.78	0.73	0.66
A-League	Favourite	30	64.29	1.02	0.95
	Underdog	30	11.90	0.26	0.22
	Draw	30	23.81	0.53	0.49

4.3 Estimation Results

Conditional logistic regressions were employed to test the direction and extent of the FLB. The dependent (binary response) variable takes the value 1 if the outcome occurred and 0 if it did not. The sole independent variable is the log of the betting-odds-implied probability. Average bookmaker odds normalised for overround were used for all estimation. Each sport was tested using the full sample of five seasons and then individually for each season. Both home and away outcomes were tested for all three sports, with the additional draw outcome tested for A-League. The results for each sport's full sample is provided in Table 8.

To reiterate, β can be interpreted as follows: if $\beta = 1$, then for the match outcome under investigation, the log of the betting-odds-implied-probability is the same as the log of the odds-ratio of success for that outcome. If $\beta < 1$, then the results are indicating a reverse-FLB, meaning that implied-probabilities on longshots are underestimated and/or implied-probabilities on favourites are overestimated. If $\beta > 1$, then the results are indicating the standard FLB, meaning that implied probabilities on longshots are overestimated and/or implied-probabilities on favourites are underestimated.

Table 8: Conditional Logit Estimation — Full Sample

	β Coef.	Std. Error	z-score	p-value
AFL Home	0.88	0.10	9.04	0.000
AFL Away	1.00	0.10	10.36	0.000
NRL Home	0.76	0.15	5.22	0.000
NRL Away	0.83	0.14	6.07	0.000
A-League Home	0.96	0.22	4.38	0.000
A-League Away	1.07	0.21	5.02	0.000
A-League Draw	1.72	1.01	1.70	0.089

The coefficient for the home outcome is less than 1 for all three sports, indicating a slight reverse-FLB. This effect is greatest for NRL with the lowest coefficient at 0.76 and smallest for A-League with a coefficient of 0.96. This implies that bookmaker odds are slightly undervaluing the probability that home-longshots have of winning, and/or overvaluing the probability of away-favourites winning. This is in concurrence with the betting simulation results, which showed positive returns to bets on home-longshots even with average odds.

Whilst these results seem indicative of a weak reverse-FLB, we can not reject the null hypothesis that $\beta_1 = 1$ in any of the three sports for the home outcome. Confidence intervals constructed at the 95% and even 90% levels all contain 1. This means that even though the coefficient is negative and a slight reverse-FLB is evident, conclusions of overall inefficiency are unfounded.

The coefficient for the away outcome is negative only for NRL, slightly positive for AFL and moderately positive for A-League. This is indicative of a reverse-FLB for NRL, a standard-FLB for A-League, and almost no bias at all for AFL. Of all three sports, the away outcome for AFL appears to be closest to 1, meaning that over the course of the past five seasons, odds on this outcome are priced at a rate that almost perfectly reflects the observed probability of the outcome. Once again, in all cases, confidence intervals constructed at 90% all contain 1, meaning we do not reject the null hypothesis that bookmaker odds implied probabilities accurately reflect observed probabilities.

In all cases except for the draw outcome for A-League, the coefficient, β, is statistically significant at the 0.1% level. The coefficient for the draw outcome in A-League is only statistically significant at the 5% level. In the context of betting markets and over a sample as large as the one used here, this is a reasonable discrepancy. It means that the bookmaker implied-probabilities for draws in A-League have slightly less predictive power than other outcomes. This is consistent with the findings of Deschamps and Gergaud (2007) who also noted lower predictive power for draw outcomes. As we will see below, this occurs for other outcomes when results are confined to a smaller sample, however only in A-League draws when looking at results for the entire sample of five seasons.

Table 9 provides the conditional logit estimation results for each season. For AFL, for both home and away outcomes, the coefficients are less than 1 for 2010, 2011 and 2014, but greater than 1 for the other two seasons. For NRL, the coefficients are less than 1 in all cases except the away outcome in 2012, indicating that a slight reverse-FLB existed almost every season. For A-League, four of the five home outcome coefficients are less than 1, but three of the five away outcomes are greater than 1. Only one of the five coefficients for A-League draw outcomes is statistically significant which further suggests that the draw odds are not reflecting the observed probabilities with which they occur. Furthermore, coefficients on draw outcomes vary greatly. The 2012-13 season draw outcome coefficient has a value of 2.04, while for the 2010-11 season it is 3.99.

What is important to note here is the discrepancy in the statistical significancy of results between seasons. Forrest's (2007) largest critique of Pope and Peel (1989) was that they only used a single season in their data set. From the results reported above, it is evident that the robustness of any conclusions made for individual seasons would be quite weak. By looking at the standard errors, it is clear that confidence intervals would be extremely wide. This makes rejecting the null hypothesis that $\beta = 1$ not possible in any case.

Table 10 reports the results of conditional logit estimation on finals matches. These results are quite erratic and do not affirm what was found in the betting simulation. In all cases standard errors are extremely large. It appears that the away outcome in A-League was statistically significant, however the coefficient is 5.32 which makes this result seem quite dubious and may have occurred due to random chance. The likely explanation for this is the low sample of finals matches. Only 46, 54 and 30 finals matches are available

Table 9: Conditional Logit Estimation — Yearly

Sport	Season	β Coef.	Std. Err.	z-score	p-value
AFL	2014 Home	0.92	0.22	4.27	0.000
	2014 Away	0.98	0.22	4.53	0.000
	2013 Home	1.04	0.22	4.69	0.000
	2013 Away	1.17	0.21	5.49	0.000
	2012 Home	1.01	0.21	4.78	0.000
	2012 Away	1.14	0.21	5.48	0.000
	2011 Home	0.78	0.21	3.69	0.000
	2011 Away	0.82	0.20	4.08	0.000
	2010 Home	0.53	0.24	2.22	0.027
	2010 Away	0.78	0.26	3.01	0.003
NRL	2014 Home	0.72	0.31	2.34	0.019
	2014 Away	0.80	0.28	2.87	0.004
	2013 Home	0.94	0.29	3.23	0.001
	2013 Away	0.96	0.26	3.65	0.000
	2012 Home	0.88	0.34	2.56	0.011
	2012 Away	1.13	0.34	3.37	0.000
	2011 Home	0.75	0.35	2.17	0.030
	2011 Away	0.74	0.35	2.12	0.034
	2010 Home	0.38	0.39	0.99	0.322
	2010 Away	0.42	0.34	1.22	0.222
A-League	2014-15 Home	0.95	0.51	1.85	0.065
	2014-15 Away	0.97	0.42	2.29	0.022
	2014-15 Draw	3.95	2.72	1.46	0.145
	2013-14 Home	0.82	0.46	1.77	0.076
	2013-14 Away	0.71	0.46	1.52	0.128
	2013-14 Draw	2.29	2.26	1.01	0.312
	2012-13 Home	0.84	0.47	1.79	0.073
	2012-13 Away	1.02	0.49	2.07	0.039
	2012-13 Draw	2.04	2.75	0.74	0.457
	2011-12 Home	0.67	0.52	1.28	0.200
	2011-12 Away	1.17	0.52	2.27	0.023
	2011-12 Draw	3.32	1.71	-1.94	0.053
	2010-11 Home	1.48	0.49	3.01	0.003
	2010-11 Away	1.50	0.49	3.05	0.002
	2010-11 Draw	3.99	2.47	1.62	0.106

for AFL, NRL and A-League, respectively, making it difficult for the model to properly converge. This problem could be solved by using a larger data set spanning a longer time period.

Table 10: Conditional Logit Estimation — Finals Matches

Sport	β Coef.	Std. Err.	z-score	p-value
AFL Home	0.95	1.05	0.91	0.37
AFL Away	0.78	0.91	0.85	0.39
NRL Home	0.78	1.11	0.71	0.48
NRL Away	0.65	1.02	0.64	0.52
A-League Home	1.67	1.49	1.12	0.26
A-League Away	5.32	2.08	2.56	0.01
A-League Draw	-0.96	7.14	0.14	0.89

4.4 Competitive Balance and the Favourite-Longshot Bias

The extent of competitiveness for each sport and each season was calculated in a number of different ways and is reported in Table 11. RSD is the measure of competitive balance most commonly reported in the literature and is very useful when looking at the same league over time. However, it is less effective when making comparisons across leagues with different scheduling structures. As such, RSD^* was calculated and serves as a better measure when attempting to compare across different leagues/sports. RSD^* is bounded between 0 and 1, with higher values indicating a less balanced season. If $RSD^* = 1$ then RSD is the highest value it can possibly be, meaning that the season was as unbalanced as possible.

Table 11: Competitive Balance Statistics

	Season	ASD	ISD	RSD	RSD^{lb}	RSD^*
AFL	2014	0.198	0.106	1.867	3.048	0.612
	2013	0.231	0.106	2.181	2.943	0.741
	2012	0.230	0.106	2.167	2.934	0.738
	2011	0.244	0.106	2.302	2.922	0.788
	2010	0.180	0.106	1.696	3.004	0.565
NRL	2014	0.128	0.102	1.259	2.568	0.490
	2013	0.159	0.102	1.560	2.568	0.608
	2012	0.150	0.102	1.470	2.568	0.572
	2011	0.182	0.102	1.786	2.568	0.695
	2010	0.135	0.102	1.328	2.568	0.517
A-League	2014/15	0.171	0.083	2.051	4.037	0.508
	2013/14	0.099	0.083	1.189	4.037	0.295
	2012/13	0.132	0.083	1.588	4.037	0.393
	2011/12	0.120	0.083	1.443	4.037	0.357
	2010/11	0.152	0.079	1.922	4.195	0.458

For AFL, RSD^* takes its lowest value in 2010 at 0.57, its highest value in 2011 at 0.79, and a mean value of 0.69. For NRL, RSD^* takes its lowest value in 2014 at 0.49, its highest value in 2011 at 0.70, and a mean value of 0.58. For A-League, RSD^* takes its lowest value in 2013/14 at 0.30, its highest value in 2014/15 at 0.51, and a mean value of 0.40. From this, we can say that according to the RSD^* measure, A-League has been the most competitive of the three sports and AFL the least competitive.

The hypothesis was that greater competitiveness leads to a reduction in the extent of the traditional FLB. This would extend the findings of Oikonomidis et al. (2015) who concluded that the extent of competitive balance within European soccer can explain the extent of the FLB. This hypothesis was tested by examining whether a relationship exists between RSD^* values and coefficients from the conditional logit estimation.

Table 12: Competitive Balance and the Favourite-Longshot Bias

Sport	Season	RSD	RSD^*	Home Coef.	Away Coef.	Avg Coef.
AFL	2014	1.87	0.61	0.92	0.98	0.95
	2013	2.18	0.74	1.04	1.17	1.11
	2012	2.17	0.74	1.04	1.17	1.11
	2011	2.30	0.79	0.78	0.82	0.80
	2010	1.70	0.57	0.53	0.78	0.66
NRL	2014	1.26	0.49	0.72	0.80	0.76
	2013	1.56	0.61	0.94	0.96	0.95
	2012	1.47	0.57	0.88	1.13	1.00
	2011	1.79	0.70	0.75	0.74	0.75
	2010	1.33	0.52	0.38	0.42	0.40
A-League	2014-15	2.05	0.51	0.95	0.97	0.96
	2013-14	1.19	0.30	0.82	0.71	0.77
	2012-13	1.59	0.39	0.84	1.02	0.71
	2011-12	1.44	0.36	0.67	1.17	0.92
	2010-11	1.92	0.46	1.48	1.50	1.49

From Table 12 and Figures 6 and 7, it is difficult to spot any evidence of a relationship. Seasons which had the lowest level of competitiveness (highest RSD and RSD^*) did not, in any sport, have the highest average β coefficient. In AFL, the season with the lowest RSD (2010) had the lowest average coefficient, however, this is not true in NRL nor A-League. These results do not support the findings of Oikonomidis et al. (2015) that competitiveness can explain the extent of the FLB. However, they do not refute their findings either. Oikonomidis et al.'s conclusion was that a greater (standard) FLB was associated with greater imbalance. The results reported here are not showing the standard FLB. It may be the case that the relationship is only clearly observable when $\beta > 1$. When $\beta < 1$ or $\beta = 1$, the same conclusion may not hold.

Alternatively, it may be the case that a relationship does exist but there are not enough data points to find evidence of it. Scatter plots for individual sports are provided in Appendix A.3. The issue with examining individual sports across time is that there are only five seasons, making meaningful inference very difficult. The graphs are possibly upwards trending, however, five data points is far too little to make any reasonable conclusion in that regard. More seasons across a longer time period would increase the sample size and help to support or dispute these results.

This concludes the analysis. We now proceed to summarise the results and conclude, with directions on potential avenues for future research.

Figure 6: Plot of RSD^* Against β — All Sports: Home Outcome.

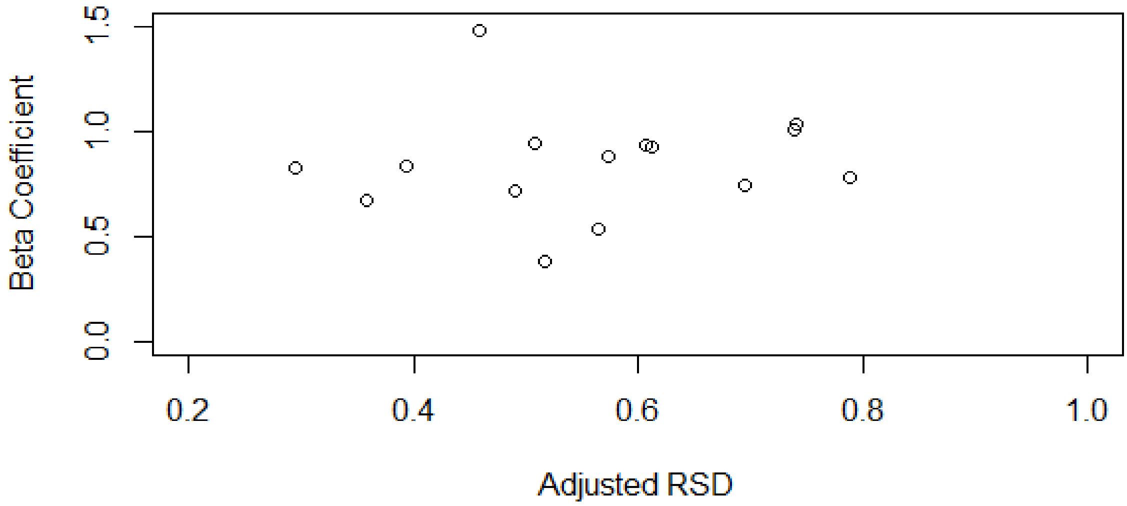
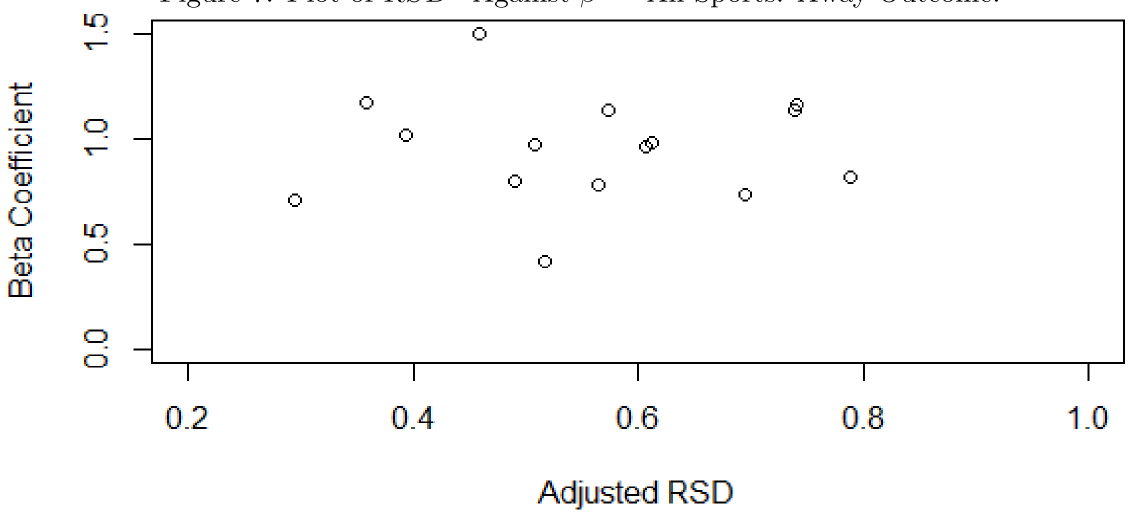


Figure 7: Plot of RSD^* Against β — All Sports: Away Outcome.



4.5 Summary and Discussion of Results

Arbitrage opportunities were uncovered in roughly half of all matches, which seems indicative of extreme market inefficiency. However, these arbitrage opportunities can be explained by two alternative reasons. Firstly, the arbitrage opportunities may arise due to strategic bookmaker odds-setting in an attempt to attract new customers and gain their loyalty, as Franck et al. (2013) suggest. Secondly, the ability to exploit these arbitrage opportunities systematically could be severely limited as bookmakers may restrict the accounts of so-called "informed bettors", as Grant et al. (2015) suggest.

The former explanation seems more likely as there is little reason to suggest that bookmakers would lose revenue from arbitrageurs. Bookmakers would still be offering the same risk-reward opportunity as they are to any other customer, and the fact that an individual is hedging or arbitraging with another bookmaker should be of little concern. Either way, a conclusion of inefficiency due to the mere existence of arbitrage opportunities is too strong. Bookmakers act as profit maximising organisations (Levitt, 2004) and compete on price in the form of odds. Due to the high level of firm competition, bookmakers appear to be willing to accept some implied losses (in the form of higher risk due to skewed odds) if it leads to customer loyalty in the future.

Betting simulation was conducted to discover whether simple strategies could have been implemented in order to secure positive returns. The results here suggest that there were some strategies that could have been implemented in both AFL and NRL that would have resulted in positive returns using average odds. Weak evidence of the home-underdog bias appeared in both AFL and NRL, with a return of 7.2% in both cases. For A-League, betting on home underdogs would have yielded negative returns, however it is still the most profitable strategy with a return of -0.006%, the next best A-League strategy yielding -2.5%. This suggests that the home ground advantage for underdogs may be higher than is priced into bookmaker odds.

From the betting simulation results it was evident that greater returns would have been made from betting on favourites rather than underdogs in finals matches, which was not true when looking at the full sample. This result held across all three sports, suggesting that incentive or match importance may play some role in the direction or extent of the FLB. It is necessary to note that these simulation results are only indicative of returns that would have been achieved in the sample period under examination and have not been held to statistical robustness checks. The same results may not necessarily hold in the future or over a longer historical sample period.

The direction and extent of the FLB was examined using a conditional logit model. The results indicate that bookmaker odds in these markets are weakly efficient. β coefficients were in most cases statistically significant at 0.05% or lower and no coefficient had a 90% confidence interval that did not contain 1. This result holds across all sample sizes and all outcomes for all three sports. Therefore, the hypothesis that bookmaker odds accurately reflect observed probabilities is not rejected. Controlling for finals matches to determine if the extent of the FLB was greater when incentive was higher was unsuccessful. The sample size was too small for the model to properly converge which resulted in erratic estimation results.

The observed relationship between RSD^* and the extent of the FLB is weak to non-existent. Nevertheless, some interesting observations can be made. AFL, which has the most imbalance according to the RSD^* measure, exhibits the lowest variation between bookmaker-implied probability and observed probability. A-League, which is the most balanced according to the RSD^* measure, exhibits the greatest variation between bookmaker-implied probability and observed probability. This latter result however, can mostly be attributed to the mispricing of draw outcomes as the other two outcomes for A-League (home and away) are priced reasonably close to their observed-probabilities.

It is difficult to make a meaningful conclusion in this regard given that the data set spans over only five seasons. By using RSD^* instead of RSD, an attempt was made to merge the sports together, but even this creates a sample of only 15 seasons. As such, a relationship between competitiveness and the extent of FLB can not be ruled out. It could be the case that the results are affected more by league or sport-specific variables, not included in the analysis. Evaluating this relationship in individual sports over a longer time period could provide more insight into this phenomenon.

5 Conclusion

This thesis had a number of aims. First and foremost, the research question being asked was whether the sports betting markets of the AFL, the NRL, and the A-League are efficient. In this regard, the answer is "for the most part".

Arbitrage opportunities appear to exist in these markets. The extent to which they were or are exploitable in real time remains a question for future research. Betting simulation showed that implementing some simple strategies in the AFL and the NRL betting markets would have yielded positive rates of return. No simple strategies were found to be profitable in the A-League betting market. This result may be coincidental over this time period and robustness checks are necessary. Statistical testing of efficiency showed that bookmaker odds approximately reflect observed probabilities of match outcomes. While small differences in the extent of deviation away from perfect calibration between implied and observed probabilities exist, these differences are not robust enough to conclude that a clear bias exists.

The evidence on the relationship between incentive and the extent of the favourite-longshot bias is mixed. Betting simulation showed that differences in returns to betting on favourites compared to betting on underdogs was greater in finals matches for all three sports. Statistical testing did not support nor reject this finding, as the sample size for finals matches was too small for the model to properly converge. In this regard, the author posits two avenues for potential future research.

Firstly, increasing the time period under consideration, which would serve to increase the robustness of any conclusions. Secondly, a better measure of match incentive could be used. Rather than simply controlling for finals matches, using a measure of incentive that assigns a value of importance to each team for a particular match could be incorporated. The sports economics literature consists of various ways to measure match importance which have not been discussed in this thesis. Perhaps one of these measures could be applied to better identify how incentive affects the extent of the favourite-longshot bias.

The evidence on the relationship between competitive balance and the extent of the favourite-longshot bias is weak. In this regard, we recommend increasing the sample period and examining individual sports independently over time. A sample period of 20 years, possibly more, may be enough to make a more robust conclusion than possible with the current data.

A Appendix

A.1 Example of Bookmaker Odds

Figure 8: Partial Screenshot of the Oddsportal Website

Essendon Bombers - Adelaide Crows

Saturday, 15 Aug 2015, 04:45

Final result 8.11 (59)-27.9 (171) (4.3 (27)-3.3 (21), 0.3 (3)-6.3 (39), 2.4 (16)-7.2 (44), 2.1 (13)-11.1 (67))

1X2 | Home/Away | AH | O/U | DNB | HT/FT | O/E

FT including OT | Full Time | 1st Half | 2nd Half | 1Q | 2Q | 3Q | 4Q

Bookmakers	1	2	Payout
10Bet	↓ 3.50	↑ 1.30	94.8%
138	↓ 3.50	↑ 1.30	94.8%
5Dimes	↓ 3.65	↑ 1.31	96.4%
5plusbet	↓ 3.40	↑ 1.27	92.5%
bet-at-home	↑ 3.50	↓ 1.26	92.6%
bet365	↑ 3.70	↓ 1.30	96.2%
Betclic	4.00	1.20	92.3%
Betfred	↑ 3.50	↓ 1.29	94.3%
Betsafe	↓ 3.75	↑ 1.27	94.9%
Betsson	↓ 3.75	↑ 1.27	94.9%
Betway	↑ 3.50	1.30	94.8%
Boylesports	4.00	1.25	95.2%
bwin	↓ 3.80	↑ 1.25	94.1%
Coral	↓ 3.70	↑ 1.25	93.4%
DOXXbet	↓ 3.62	↑ 1.27	94.0%
Expekt	4.00	1.20	92.3%
Interwetten	3.60	1.20	90.0%
Island Casino	↓ 3.65	↑ 1.31	96.4%
Luxbet	↑ 3.75	↓ 1.29	96.0%
MarathonBet	↓ 3.58	↑ 1.34	97.5%

A.2 Example of a Hypothetical League Table Used to Construct RSD^{lb}

Table 13: Hypothetical Least-Balanced League Table — NRL 2014

Pos.	Team Name	Played	Win	Draw	Lose	Points
1	Sydney Roosters	24	24	0	0	48
2	Manly-Warringah Sea Eagles	24	22	0	2	44
3	South Sydney Rabbitohs	24	20	0	4	40
4	Penrith Panthers	24	21	0	3	42
5	North Queensland Cowboys	24	18	0	6	36
6	Melbourne Storm	24	16	0	8	32
7	Canterbury-Bankstown Bulldogs	24	13	0	11	26
8	Brisbane Broncos	24	12	0	12	24
9	New Zealand Warriors	24	13	0	11	26
10	Parramatta Eels	24	9	0	15	18
11	St. George Illawarra Dragons	24	8	0	16	16
12	Newcastle Knights	24	6	0	18	12
13	Wests Tigers	24	6	0	18	12
14	Gold Coast Titans	24	3	0	21	6
15	Canberra Raiders	24	1	0	23	2
16	Cronulla-Sutherland Sharks	24	0	0	24	0

In Table 13, teams shaded grey would have had their ordering swapped based on their points. Points awarded for byes were not included in hypothetical nor actual league tables.

A.3 Plots of RSD^* Against β for Individual Sports

Figure 9: Plot of RSD^* Against β — AFL Home Outcome

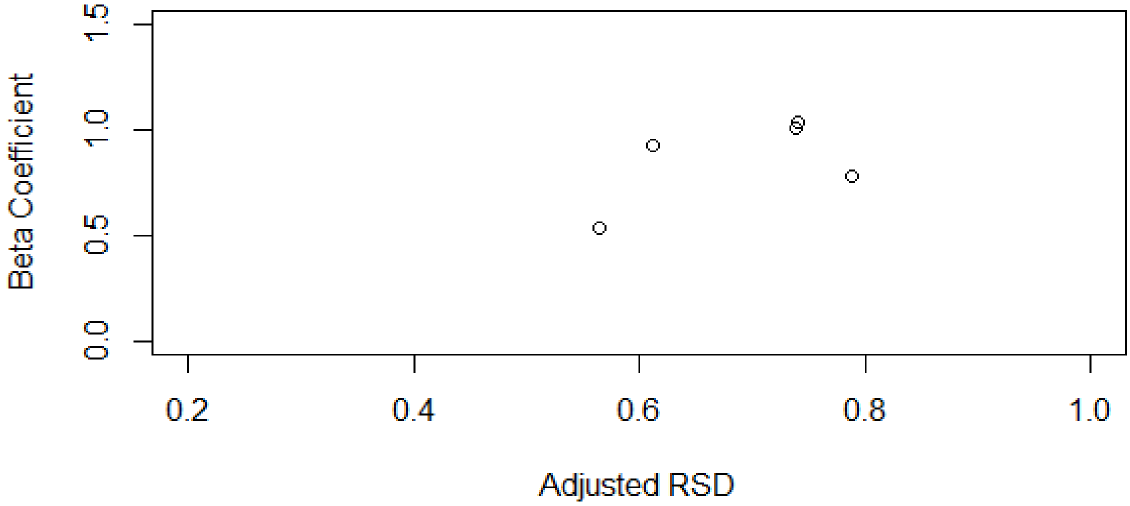


Figure 10: Plot of RSD^* Against β — AFL Away Outcome

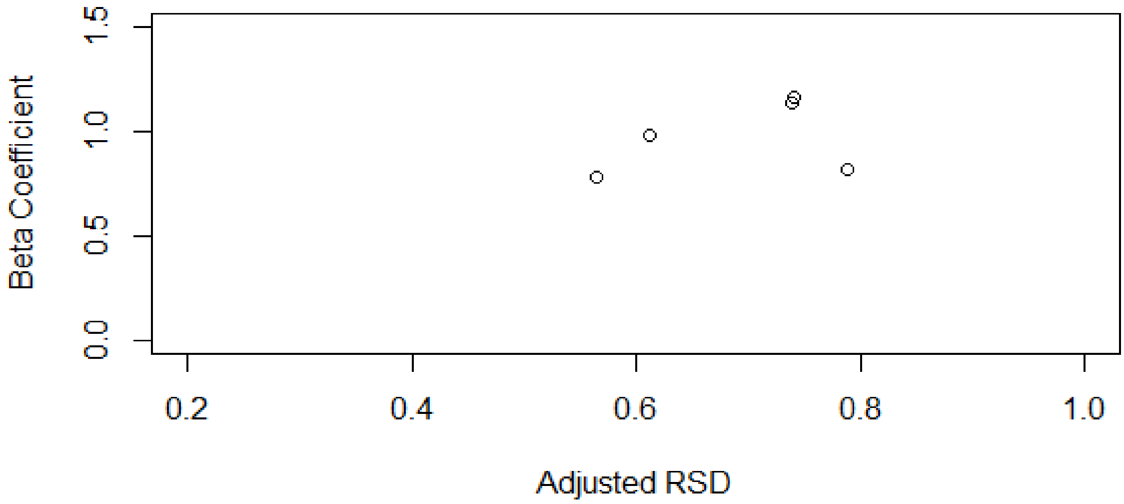


Figure 11: Plot of RSD^* Against β — NRL Home Outcome

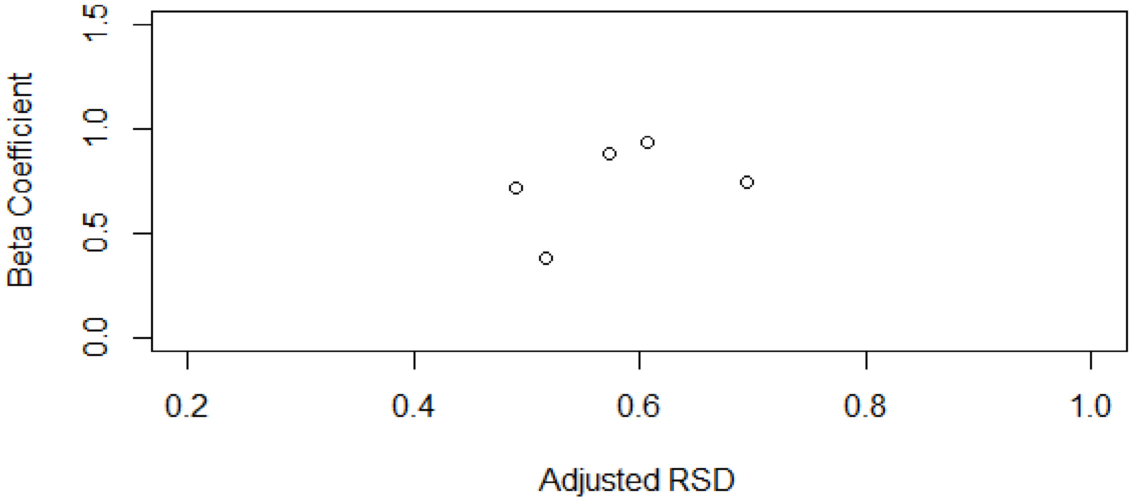


Figure 12: Plot of RSD^* Against β — NRL Away Outcome

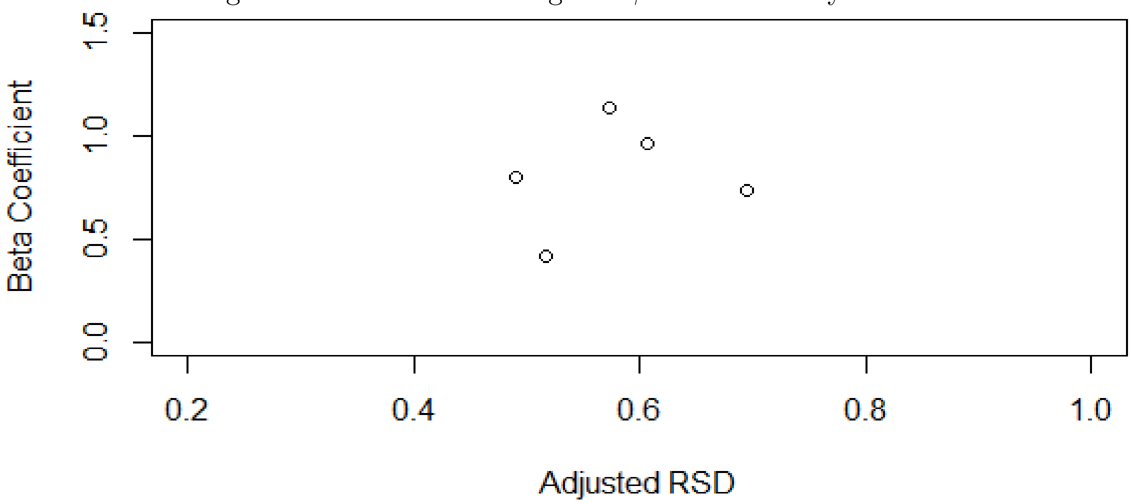


Figure 13: Plot of RSD^* Against β — A-League Home Outcome

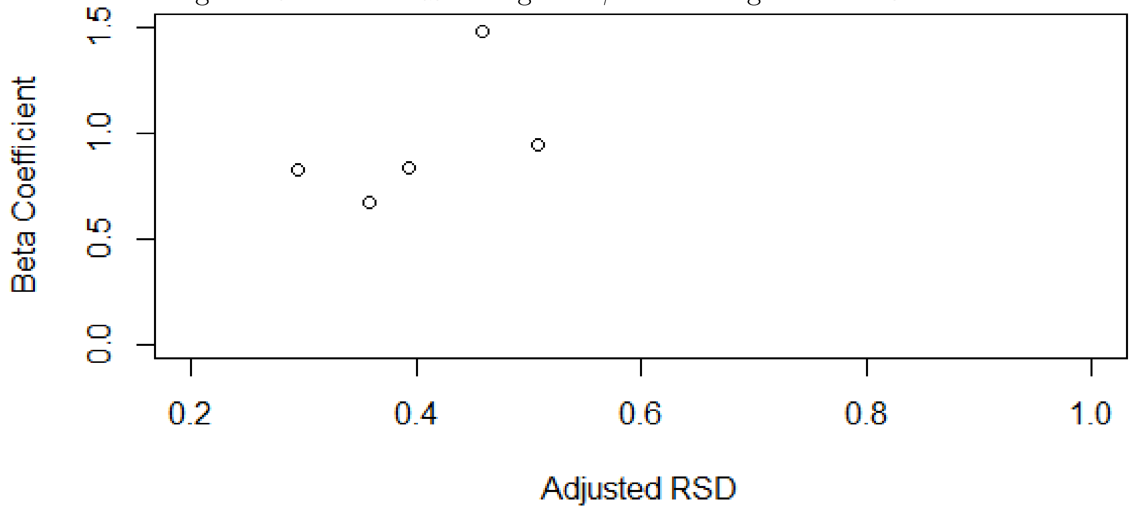
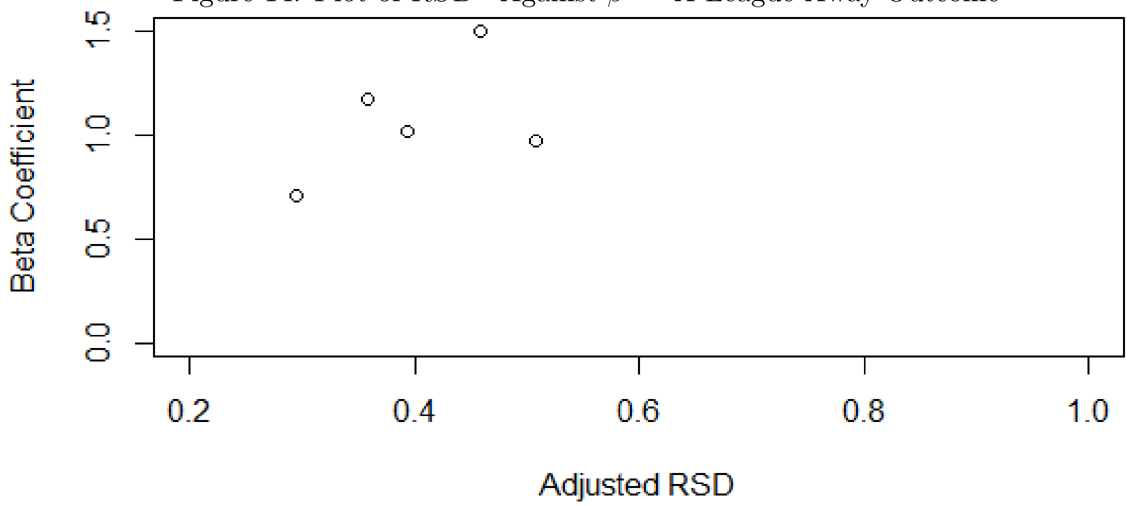


Figure 14: Plot of RSD^* Against β — A-League Away Outcome



References

Aitken, A. (1936). IV. On Least Squares and Linear Combination of Observations. *Proceedings of the Royal Society of Edinburgh*, *55*, 42-48.

Ali, M. (1977). Probability and Utility Estimates for Racetrack Bettors. *Journal of Political Economy*, *85*(4), 803.

Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica*, *21*(4), 503.

Asch, P., Malkiel, B., & Quandt, R. (1982). Racetrack Betting and Informed Behavior. *Journal of Financial Economics*, *10*(2), 187-194.

Bacon-Shone, J., Lo, V., & Busche, K. (1992). Modelling Winning Probability. *University of Hong Kong: Research Report, 10*.

Bar-Eli, M., Avugos, S., & Raab, M. (2006). Twenty Years of "Hot Hand" Research: Review and Critique. *Psychology of Sport and Exercise*, *7*(6), 525-553.

Booth, R. (2004). The Economics of Achieving Competitive Balance in the Australian Football League, 1897-2004. *Economic Papers: A Journal of Applied Economics and Policy*, *23*(4), 325-344.

Booth, R. (2005). Comparing Competitive Balance in Australian Sports Leagues: Does a Salary Cap and Player Draft Measure Up? *Sport Management Review*, *8*(2), 119-143.

Brailsford, T., Gray, P., Easton, S., & Gray, S. (1995). The Efficiency of Australian Football Betting Markets. *Australian Journal of Management*, *20*(2), 167-195.

Brown, W., & Sauer, R. (1993). Does the Basketball Market Believe in the Hot Hand? Comment. *The American Economic Review*, 1377-1386.

Bruce, A., & Johnson, J. (1992). Toward an Explanation of Betting as a Leisure Pursuit. *Leisure Studies*, *11*(3), 201-218.

Busche, K., & Hall, C. (1988). An Exception to the Risk Preference Anomaly. *The Journal of*

Business, *61*(3), 337.

Cain, M., Law, D., & Peel, D. (2000). The Favourite-Longshot Bias and Market Efficiency in UK Football betting. *Scottish Journal of Political Economy*, *47*(1), 25-36.

Cain, M., Law, D., & Peel, D. (2003). The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets. *Bull Economic Research*, *55*(3), 263-273.

Camerer, C. (1989). Does the Basketball Market Believe in the Hot Hand? *The American Economic Review*, 1257-1261.

Cameron, A., & Trivedi, P. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Crafts, N. (1985). Some Evidence of Insider Knowledge in Horse Race Betting in Britain. *Economica*, 295-304.

Depken II, C. (1999). Free-agency and the Competitiveness of Major League Baseball. *Review of Industrial Organization*, *14*(3), 205-217.

Deschamps, B., & Gergaud, O. (2012). Efficiency in Betting Markets: Evidence from English Football. *The Journal of Prediction Markets*, *1*(1), 61-73.

Direr, A. (2013). Are Betting Markets Efficient? Evidence from European football Championships. *Applied Economics*, *45*(3), 343-356.

Dixon, M., & Coles, S. (1997). Modelling Association Football Scores and Inefficiencies in the Football Betting Market. *Journal of the Royal Statistical Society*, *46*(2), 265-280.

Dixon, M., & Pope, P. (2004). The Value of Statistical Forecasts in the UK Association Football Betting Market. *International Journal of Forecasting*, *20*(4), 697-711.

Dobson, S., & Goddard, J. (2001). *The Economics of Football*. New York: Cambridge University Press.

Dobson, S., Goddard, J., & Wilson, J. (2001). League Structure and Match Attendances in English Rugby League. *International Review of Applied Economics*, *15*(3), 335-351.

Dowie, J. (1976). On the Efficiency and Equity of Betting Markets. *Economica*, *43*(170), 139.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, *75*(4), 643.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work*. *The Journal of Finance*, *25*(2), 383-417.

Forrest, D. (2008). Soccer betting in Britain. *Handbook of Sports and Lottery Markets*, 421-446.

Forrest, D., & Mchale, I. (2007). Anyone for Tennis (Betting)? *The European Journal of Finance*, *13*(8), 751-768.

Forrest, D., & Simmons, R. (2000). Forecasting sport: the behaviour and performance of football tipsters. *International Journal of Forecasting*, *16*(3), 317-331.

Forrest, D., Gulley, O., & Simmons, R. (2010). The Relationship Between Betting and Lottery Play. *Economic Inquiry*, *48*(1), 26-38.

Fort, R., & Lee, Y. (2007). Structural Change, Competitive Balance, and the Rest of the Major Leagues. *Economic Inquiry*, *45*(3), 519-532.

Fort, R., & Quirk, J. (1995). Cross-subsidization, Incentives, and Outcomes in Professional Team Sports Leagues. *Journal of Economic Literature*, 1265-1299.

Franck, E., Verbeek, E., & Nuesch, S. (2013). Inter-market Arbitrage in Betting. *Economica, 80*(318), 300-325.

Gandar, J., & Zuber, R. (2004). An Evaluation of the Debate Over "Testing Market Efficiency in the Major League Baseball Over-Under Betting Market". *Journal of Sports Economics, 5*(1), 100-105.

Gandar, J., Zuber, R., & Johnson, R. (2001). Searching for the Favourite-Longshot Bias Down Under: an Examination of the New Zealand Pari-mutuel Betting Market. *Applied Economics*, *33*(13), 1621-1629.

Gandar, J., Zuber, R., & Lamb, R. (2001). The Home Field Advantage Revisited: a Search for the Bias in other Sports Betting Markets. *Journal of Economics and Business*, *53*(4), 439-453.

Gandar, J., Zuber, R., O'Brien, T., & Russo, B. (1988). Testing Rationality in the Point Spread Betting Market. *The Journal of Finance*, *43*(4), 995.

Goddard, J. (2005). Regression Models for Forecasting Goals and Match Results in Association Football. *International Journal of Forecasting*, *21*(2), 331-340.

Goddard, J., & Asimakopoulos, I. (2004). Forecasting Football Results and the Efficiency of Fixed-odds Betting. *Journal of Forecasting*, *23*(1), 51-66.

Goossens, K. (2006). Competitive Balance in European Football: Comparison by Adapting Measures: National Measure of Seasonal Imbalance and Top 3. *Revista Di Dirrito Ed Economica dello Sport*, *2*, 77-122.

Graham, I., & Stott, H. (2008). Predicting Bookmaker Odds and Efficiency for UK Football. *Applied Economics*, *40*(1), 99-109.

Grant, A., Johnson, J.E.V. and Oikonomidis, A. (2015) Arb-mirage: Exploring the Extent to which Apparent Inefficiency in Betting Markets is an Illusion. In, *22nd Annual Conference of the Multinational Finance Society, Halkidiki, GR, 28 Jun - 01 Jul 2015.*

Griffith, R. (1949). Odds Adjustments by American Horse-Race Bettors. *The American Journal of Psychology*, *62*(2), 290.

Hausch, D., & Ziemba, W. (1995). Efficiency of Sports and Lottery Betting Markets. *Handbooks in Operations Research and Management Science*, *9*, 545-580.

Horowitz, I. (1997). The Increasing Competitive Balance in Major League Baseball. *Review of Industrial Organization*, *12*(3), 373-387.

Humphrey, J. (1986). *Roman Circuses: Arenas for Chariot Racing*. University of California Press.

Hurley, W., & McDonough, L. (2005). 10 On the Improbability of Informationally Efficient Parimutuel Betting Markets in the Presence of Heterogeneous Beliefs. *Information*

Efficiency in Financial and Betting Markets, 247.

Hurley, W., & McDonough, L. (2007). Imperfect Market-maker Competition, Heterogeneous Expectations, and the Favourite-Longshot Bias in Wagering Markets. *The Journal of Gambling Business and Economics*, *1*(1), 3-12.

Isaacs, R. (1953). Optimal Horse Race Bets. *The American Mathematical Monthly*, *60*(5), 310.

Johnson, J., & Bruce, A. (2001). Calibration of Subjective Probability Judgments in a Naturalistic Setting. *Organizational Behavior and Human Decision Processes*, *85*(2), 265-290.

Johnson, R., & Pawlukiewicz, J. (1992). The Efficiency of the Over-under Betting Market for NBA Basketball Games. *Journal of Economics*, *18*, 97-100.

Kochman, L., & Badarinathi, R. (1992). Pricing Inefficiencies in the Football-betting Market. *Atlantic Economic Journal*, *20*(3), 104-104.

Koning, R. (2000). Balance in Competition in Dutch Soccer. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *49*(3), 419-431.

Kuypers, T. (2000). Information and Efficiency: An Empirical Study of a Fixed Odds Betting Market. *Applied Economics*, *32*(11), 1353-1363.

Lacey, N. (1990). An Estimation of Market Efficiency in the NFL Point Spread Betting Market. *Applied Economics*, *22*(1), 117-129.

Lahvivcka, J. (2014). What Causes the Favourite-Longshot Bias? Further Evidence from Tennis. *Applied Economics Letters*, *21*(2), 90-92.

Lenten, L. (2011). The Extent to which Unbalanced Schedules Cause Distortions in Sports League Tables. *Economic Modelling*, *28*(1-2), 451-458.

Levitt, S. (2004). Why are Gambling Markets Organised so Differently from Financial Markets? *Economic Journal*, *114*(495), 223-246.

Machina, M. (1987). Decision-Making in the Presence of Risk. *Science*, *236*(4801), 537-543.

Maddala, G. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge [Cambridgeshire]: Cambridge University Press.

Marshall, B. (2009). How Quickly is Temporary Market Inefficiency Removed? *The Quarterly Review of Economics and Finance*, *49*(3), 917-930.

McGlothlin, W. (1956). Stability of Choices among Uncertain Alternatives. *The American Journal of Psychology*, *69*(4), 604.

Norton, H., Gray, S., & Faff, R. (2015). Yes, One-day International Cricket "In-Play" Trading Strategies can be Profitable! *Journal of Banking & Finance*.

Nyman, J., Welte, J., & Dowd, B. (2008). Something for Nothing: A Model of Gambling Behavior. *The Journal of Socio-Economics*, *37*(6), 2492-2504.

Oikonomidis, A., Bruce, A., & Johnson, J. (2015). Does Transparency Imply Efficiency? The Case of the European Soccer Betting Market. *Economics Letters*, *128*, 59-61.

Owen, P. (2010). Limitations of the Relative Standard Deviation of Win Percentages for Measuring Competitive Balance in Sports Leagues. *Economics Letters*, *109*(1), 38-41.

Owen, P. (2012). Measuring Parity in Sports Leagues with Draws: Further Comments. *Journal of Sports Economics*, 13(1), 85-95.

Owen, P., & King, N. (2015). Competitive Balance Measures in Sports Leagues: the Effects of Variation in Season Length. *Economic Inquiry*, *53*(1), 731-744.

Paton, D., & Williams, L. (2005). Forecasting Outcomes in Spread Betting Markets: Can Bettors Use "Quarbs" to Beat the Book? *Journal of Forecasting*, *24*(2), 139-154.

Pope, P., & Peel, D. (1989). Information, Prices and Efficiency in a Fixed-Odds Betting Market. *Economica*, *56*(223), 323.

Preston, F., Bernhard, B., Hunter, R., & Bybee, S. (1998). Gambling as Stigmatized Behavior: Regional Relabeling and the Law. *The Annals of the American Academy of Political and Social Science*, 186-196.

Preston, M., & Baratta, P. (1948). An Experimental Study of the Auction-value of an Uncertain Outcome. *The American Journal of Psychology*, 183-193.

Productivity Commission, (2010). *Gambling Inquiry Report*. Canberra.

Quirk, J., & Fort, R. (1997). *Pay Dirt: The business of Professional Team Sports*. Princeton University Press.

Rue, H., & Salvesen, O. (2000). Prediction and Retrospective Analysis of Soccer Matches in a League. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *49*(3), 399-418.

Ruhm, C. (2003). Good Times Make You Sick. *Journal of Health Economics*, *22*(4), 637-658.

Ryall, R., & Bedford, A. (2010). An Optimized Ratings-based Model for Forecasting Australian Rules Football. *International Journal of Forecasting*, *26*(3), 511-517.

Sauer, R. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 2021-2064.

Sauer, R., Brajer, V., Ferris, S., & Marr, M. (1988). Hold Your Bets: Another Look at the Efficiency of the Gambling Market for National Football League Games. *Journal of Political Economy*, *96*(1), 206.

Schilling, M. (1994). The Importance of a Game. *Mathematics Magazine*, *67*(4), 282.

Schmidt, M., & Berri, D. (2001). Competitive Balance and Attendance: The Case of Major League Baseball. *Journal of Sports Economics*, *2*(2), 145-167.

Schnytzer, A., & Weinberg, G. (2007). Testing for Home Team and Favorite Biases in the Australian Rules Football Fixed-Odds and Point Spread Betting Markets. *Journal of Sports Economics*, *9*(2), 173-190.

Schwartz, D. (2013). Roll the bones: The history of gambling.

Scully, G. (1900). The Business of Major League Baseball. *University of Chicago Press Economics Books*.

Shin, H. (1991). Optimal Betting Odds Against Insider Traders. *The Economic Journal*, *101*(408), 1179.

Shin, H. (1992). Prices of State Contingent Claims with Insider Traders, and the Favourite-Longshot Bias. *The Economic Journal*, *102*(411), 426.

Shin, H. (1993). Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims. *The Economic Journal*, *103*(420), 1141.

Shmanske, S. (2005). Odds-setting Efficiency in Gambling Markets: Evidence from the PGA Tour. *Journal of Economics and Finance*, *29*(3), 391-402.

Smith, M., & Vaughan Williams, L. (2008). Betting Exchanges: a Technological Revolution in Sports Betting. *Handbook of Sports and Lottery Markets, New York: North Holland*.

Smith, M., Paton, D., & Williams, L. (2006). Market Efficiency in Person-to-Person Betting. *Economica*, *73*(292), 673-689.

Snowberg, E., & Wolfers, J. (2010). Explaining the Favorite-Longshot Bias: Is it Risk-Love or Misperceptions? *The Journal of Political Economy*, *118*(4), 723-746.

Stigler, G. (1970). The theory of price.

Sung, M., Johnson, J., & Dror, I. (2009). Complexity as a Guide to Understanding Decision Bias: A Contribution to the Favorite-Longshot Bias Debate. *Journal of Behavioral Decision Making*, *22*(3), 318-337.

Swidler, S., & Shaw, R. (1995). Racetrack wagering and the "Uninformed" Bettor: A study of Market Efficiency. *The Quarterly Review of Economics and Finance*, *35*(3), 305-314.

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

Thompson, W. (2001). *Gambling in America: An Encyclopedia of History, Issues, and Society*. Abc-Clio.

Tuckwell, R. (1983). The Thoroughbread Gambling Market: Efficiency, Equity and Related

Issues. *Australian Economic Papers, 22*(40), 106-118.

Vaughan Williams, L., & Paton, D. (1997). Why is There a Favourite-Longshot Bias in British Racetrack Betting Markets? *Economic Journal*, *107*(440), 150-158.

Vecer, J., Ichiba, T., & Laudanovic, M. (2006). Parallels Between Betting Contracts and Credit Derivatives: Lessons Learned from FIFA World Cup 2006 Betting Markets. *Science*, 1-15.

Vergin, R., & Scriabin, M. (1978). Winning Strategies for Wagering on National Football League Games. *Management Science, 24*(8), 809-818.

Vlastakis, N., Dotsis, G., & Markellos, R. (2009). How Efficient is the European Football Betting Market? Evidence from Arbitrage and Trading Strategies. *Journal of Forecasting*, *28*(5), 426-444.

Von Neumann, J. (1953). Morgenstern, O.(1944). Theory of Games and Economic Behavior. *Princeton: Princeton University Press*.

Weitzman, M. (1965). Utility Analysis and Group Behavior: An Empirical Study. *Journal of Political Economy*, *73*(1), 18.

Williams, L. (1999). Information Efficiency in Betting Markets: a Survey. *Bulletin of Economic Research*, *51*(1), 1-39.

Woodland, L., & Woodland, B. (1994). Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market. *The Journal of Finance*, *49*(1), 269.

Woodland, L., & Woodland, B. (2001). Market Efficiency and Profitable Wagering in the National Hockey League: Can Bettors Score on Longshots? *Southern Economic Journal*, *67*(4), 983.

Woodland, L., & Woodland, B. (2003). The Reverse Favourite-Longshot Bias and Market Efficiency in Major League Baseball: An Update. *Bull Economic Research*, *55*(2), 113-123.

Zuber, R., Gandar, J., & Bowers, B. (1985). Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games. *Journal of Political Economy*, *93*(4), 800.

Websites:

www.oddsportal.com. (2015). Odds Portal: Odds Comparison, Sports Betting Odds. Retrieved July 7, 2015, from http://www.oddsportal.com

www.ultimatealeague.com,. (2015). Standings : 2015-16 : Ultimate A-League. Retrieved July 8, 2015, from http://www.ultimatealeague.com/standings.php

www.afltables.com,. (2015). AFL Tables. Retrieved July 8, 2015, from http://afltables.com/afl/afl_index.html

www.afltables.com,. (2015). Rugby League Tables & Statistics. Retrieved July 8, 2015, from http://afltables.com/rl/rl_index.html

R Packages

R Core Team (2015). R: A language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria, http://www.R-project.org/

Therneau, M. (2015). A Package for Survival Analysis in S. Version 2.3.8. http://CRAN.R-project.org/package=survival