TESTING FOR EFFICIENCY IN THE NEW ZEALAND HORSE RACETRACK BETTING MARKET

A Thesis in partial fulfilment
of the requirements for the degree of
Master of Applied Economics at
Massey University

Malcolm McKee

Massey University
1998

Abstract

Using a large sample of New Zealand pari-mutuel horse race betting data, this study tests for market efficiency. This involves testing for weak form efficiency and an anomaly known as the favourite longshot bias. Additionally, a test developed by Henery (1985) is used to examine the extent to which bettors discount their losses. Also, two utility estimations are calculated using the first three moments surrounding the distribution. Each test is performed twice, firstly with the odds at the close of the tote and secondly with the odds quoted 30 minutes before the tote closes.

A number of previous studies are reviewed. The data set is discussed along with its limitations. An extensive description of the research methodology is presented, followed by the results, interpretations and discussion.

Many former studies have found that racetrack betting is not weak form efficient, but instead there exists a negative risk-return trade-off in the market. This study, exhibiting the negative risk-return trade-off and the favourite longshot bias, is consistent with previous studies. Using opening odds, there is much evidence to suggest that the favourite longshot bias exists 30 minutes before the tote closes but is essentially eliminated in the final 30 minutes of betting. The estimation of Henery's test is consistent with his results that bettors discount approximately 2% of there losses as 'not typical'. The implications of this are also discussed. The estimation of bettor's utility functions shows that bettors are risk lovers and, contrary to one study, the inclusion of the third moment is insufficient to prove bettors are in fact risk averse.

Acknowledgements

I am greatly indebted to my research supervisor, Professor John Gandar, Department of Economics, University of North Carolina at Charlotte, USA, for his extensive guidance, time and constructive criticism. Without his support, this thesis would not have been possible. The initial computer programming in this study must also be credited to John. I also wish to thank my second supervisor, Mr. Gary Buurman, Department of Applied and International Economics, Massey University, for his extensive proof reading and his unique ability to make the unreadable readable. I also owe gratitude to Dr. Claudio Michelini, Department of Applied and International Economics, Massey University, for his help with the econometric techniques adopted in this thesis. His advice, criticism and humour was greatly appreciated.

Thanks must also go to the Totalisator Agency Board (T.A.B.) for supplying me with the data set. Also, a special thanks to Ron McKenzie, System Support Manager at the T.A.B., who was always pleased to answer any queries I had throughout the year and for proof reading the section on the New Zealand Racing Industry.

I am also appreciative of my partner, Andrea Thomas, who has helped me proof read sections and has offered advice when asked. She can also be credited with helping me stay sane at times.

Also, a big thank you to my parents, Hugh & Jean, and my two brothers, Kelvin & Hayden, for their support, both financial and personal, throughout my days at Massey. Finally, a special thanks to my Mother for supplying me with my brains, as she so often reminds me.

This thesis is dedicated to my late Grandmother, Gladys McKee, who after gaining her university entrance was not able to fulfil her dream of completing a degree.

This one's for you Nana.

Table of Contents

List of Figures

List of Tables

1. Introduction

For centuries, societies have been intrigued with racing. They have been enthralled by racing everything; cars, bikes, boats, dogs, snails, frogs and of course, horses. More importantly, humans have been interested in trying to pick the winners of these races. People have a fascination with trying to predict the outcome of an uncertain event. Bettors believe that they can overcome this uncertainty and pick winners. It is from this aspect of human nature that betting was born. A bettor may think they can pick the winner of a race, and therefore they may be willing to put money forward as an indication of how much they believe in their choice. If bettors are able to find another person who is willing to back an opposing horse, we have a potential betting system, as systems require bettors with differing opinions. Bookmakers and organisations see this opportunity and hence, facilitate bets, as they sense there is money to be made by providing this service and bringing bettors together. In effect, they receive all the bets from bettors and pay back a proportion to those who correctly predict the race outcome.

1.1 New Zealand Racing Industry

The facilitator of bets in New Zealand is the Totalisator Agency Board, or the T.A.B for short. This board is closely monitored by Parliament and it must abide by the guidelines detailed in the Racing Act (1971) and its various amendments in 1986, 1989, 1992 and 1995. This Act prohibits any other person or organisation running their own bookmaking system, hence the T.A.B. has a monopoly in New Zealand.

The T.A.B. runs what is called a pari-mutuel betting system, meaning they take bets from the public, remove a percentage to cover costs, profits, the Governments' 6% tax on gambling, and return the remainder to the bettors who correctly predict the outcome. To illustrate this, consider bets placed on the winner of a race. The T.A.B. will receive money from thousands of bettors who place bets at odds offered by the T.A.B. Once the T.A.B. receives this money, they place it in what is called a win pool. When the race is finished and the winner is determined, they return the money less the track take to the bettors that correctly predict the winner. The T.A.B. cannot lose. Irrespective of the betting volume, the T.A.B. always retain a set percentage. In New Zealand, the track take is currently 15.5% for win and place bets, 21% for quinellas, 25% for trifectas and 26% for pick 6 and six pack. In the horse racing market, there are two main players; the bettors and the T.A.B. Since the T.A.B. always retains the track take, bettors as a whole MUST lose, even though it is possible for individual bettors to make money. Imagine a bettor that only makes one bet in their lifetime. If that bet wins, then this bettor has made money. In this situation, you will hear bettors say that they have made money out of the T.A.B., yet the T.A.B. have still collected their set percentage. The winnings are at the expense of fellow bettors.

Currently there are four ways you can bet; on course, off course, by the telephone and over the internet. On course betting allows bettors at the race meeting itself to use the numerous betting windows operated by the T.A.B., who are contracted by the local racing club. The T.A.B. also operates what they term agencies and sub-agencies. The agencies are T.A.B. retail outlets scattered throughout the country which are operated by an agent on a commission basis. Sub-agencies are unrelated businesses offering T.A.B. services, most commonly being public bars, bookshops and service stations. Additionally, the T.A.B. operates a telephone betting system where bettors can put money into a T.A.B. account. When they ring up and quote a PIN, they can make bets with the money in their account. Winnings are credited to their account and losses debited. Internet betting works identically to telephone betting.

The most popular bets are on the win and place pools. When bettors bet in the win pool, they are attempting to predict the winner of the race. A place bet wins when the horse finishes 1^{st}, 2^{nd} or 3^{rd}. Bettors can also bet on what are termed 'exotic bets', like quinellas, trifectas, doubles, trebles and pick 6. A quinella involves correctly picking the two horses that finish 1^{st} and 2^{nd} in any order. A trifecta requires accurately predicting the horses that finish 1^{st} 2^{nd} and 3^{rd} in the correct order[1]. Picking the winner of two races at one race meeting is called a double, while a treble involves picking the winner of three races at one race meeting. The T.A.B. has also designed a betting system called 'Pick 6'. A bettor can attempt to pick the winners of six races themselves or they can get the T.A.B. computer to pick six horses in a lucky dip[2] situation. The computer chooses six horses with a weighting towards the favourites by assigning sets of numbers to all horses in the six races. As the level of favouritism falls, the horses have less numbers assigned to them. The computer then randomly chooses one number from

[1] A variation on a trifecta is to 'box' the horses so that they can finish in any order. You can also box three or more horses for a quinella bet. These options obviously cost more.
[2] Note that 'lucky dips' are available on all bet types.

each of the six races. You are more likely to get the favourite horse as more numbers are assigned to it, however, it is still possible to get allocated six longshots. This thesis studies data in the win pool only.

The question of when bettors can bet depends upon the size of the race and when it is being run. For a typical weekday race meeting, the tote 'on course' and 'off course' will open approximately 1 hour before the first race. At this time, you can bet on all the races up until the horses 'jump' at the start of the race. At the end of the raceday, 'on course' betting windows stay open for an hour, allowing bettors to collect their winnings on the last race and to bet on other races around New Zealand and Australia that have not yet finished. For a typical Saturday meeting, the 'off course' tote will open on the Friday. The 'on course' tote will not open until the morning of the races.

To comprehend the size of the New Zealand racing industry, in 1997 the T.A.B. took approximately 1 billion dollars worth of bets, equating to approximately $275 per capita. A typical Saturday race meeting will gross approximately $1,500,000. The biggest race meeting of the year is the Melbourne Cup, where the T.A.B. can gross approximately $9,500,000 for all 10 races, with most of this from the cup race. The biggest race within New Zealand is the Auckland Cup, which grosses approximately $4,700,000, followed by the Wellington Cup, which grosses $3,000,000[3].

There are three main horseracing publications in New Zealand; 'Turf Digest', 'Best Bets' and the 'Friday Flash'. All publications contain similar information about gallops. A vast amount of information about races is contained in these publications;

- Horse name
- Weight to be carried
- Barrier draw
- Trainers / owners and stable
- Age, sex, sire, dam
- Details of last three starts, including;
  - when, where, weight, race distance, details of run down the home straight, placing, margin, time, jockey
- Placing in last six starts
- Records on different tracks (NZR, F, E, S, H, C, D, FR) where NZR= NZ races, F = firm, E = easy, S = soft, H = heavy, C = current track, D = record over current distance, FR = racing fresh (after a spell of 42 days or longer)
- Prize money won

[3] 1997 Figures

The New Zealand T.A.B. also covers some of the big race meetings in Australia. For these often major races, the fixed odds market may open a week before the actual race. In this situation, the T.A.B. runs what is termed fixed odds betting up until the day of the race, where the minimum bet is $5. On the actual raceday, only pari-mutuel odds will be offered. The next section will outline how pari-mutuel and fixed odds systems work.

1.2 Pari-mutuel and Fixed Odds Betting

The pari-mutuel betting system in New Zealand works as follows. Imagine there are N horses running in a race. Each of these horses has an amount bet on them to win, which is denoted as $b_1, b_2, \ldots, b_N$. Let the total win pool for the j[th] race be denoted by B_j, such that

$$B_j = b_1 + b_2 + \cdots + b_N$$

Letting t denote the track take, supposing horse *i* wins the race, the payoff per dollar bet on horse *i* will be

$$OW_i = \frac{B_j(1-t)}{b_i}$$

where OW_i represents the odds quoted at T.A.B. agencies[4]. However, the OW_i quoted at the time the bettor places their bet may not be the same odds that get paid out if the horse wins. The odds that are to be paid out are calculated after the tote closes for the race. After a bet has been placed, the odds can change depending on where other bettors place their money.

The T.A.B. also runs a fixed odds betting system. This system is used for all sports betting and on important horse races as explained above. Once you make a bet under fixed odds, the odds quoted at the time are paid out if the bet wins, hence the payout is not dependent upon other bets. When the fixed odds market opens before a race like the Melbourne Cup, a large amount of information is made public between the time the fixed odds market opens and the race. If the T.A.B. only offered pari-mutuel betting prior to the race, this information could change the odds dramatically within that time. Compare this to pari-mutuel betting on the day of the race. By the raceday, most information affecting the odds has been made public. Bettors are able to predict the closing odds more accurately a day rather than a week before the race. If fixed odds are

[4] Odds are rounded down to the nearest 5 cents. This is known as breakage.

offered, bettors are more likely to bet earlier in the week than with a pari-mutuel betting system.

1.3 Thesis Overview

There are four different areas of concern in this thesis, all related to each other. Firstly, there is the issue of efficiency. In Chapter 2, efficiency is formally defined and linked to the racing industry. This allows the hypothesis, that the New Zealand racing industry is inefficient since not all information is reflected in the odds, to be examined. The expected returns from adopting different strategies are also calculated. These strategies will be simple, for example, bet on 1st favourite to win, bet on 2nd favourite to win, bet on 3rd favourite to win and so on. Alternatively, what would the bettor's return be if they adopted a strategy of only betting on horses within a predetermined subjective probability interval? A test to see if the continued support of certain odd categories can make a positive profit will be conducted to test the hypothesis that, for every strategy, the returns will equal one minus track take. From here, some inference can be made as to which strategy is best.

The New Zealand racing data is then used to perform a well-documented test called the favourite longshot bias. This test will be performed on differing levels. For example, this thesis attempts to answer questions such as; 'is the favourite longshot bias more pronounced 30 minutes before a race than at the start of the race?', 'does the favourite longshot bias become more prolific throughout a raceday?', 'is there any vast difference in the favourite longshot bias when we define objective probabilities in different ways?'.

Henery (1985) discloses a theory that attempts to explain the occurrence of bettors participating in a "negative sum game". His idea is that bettors underestimate their losses. Henery's test will be conducted using the New Zealand data.

Another theory of why bettors participate in a "negative sum game" will then be examined. To explain this phenomenon, the New Zealand data will be tested for evidence of skewness using a test conducted by Golec and Tamarkin (1998). A power utility function developed by Ali (1977) will also be applied to the New Zealand data.

1.4 Importance of Research

Each of the tests performed in this thesis have been developed and tested by others using different countries data sets, however very little has been done using the New Zealand racing data. Although this study may not have a direct contribution to knowledge, the repetition of an idea with a new data set is not immaterial. It is possible that a new discovery could be found in the New Zealand racing industry that did not occur when the tests were performed in other countries. The results that are found will either confirm or reject conclusions of previous studies.

Two groups could benefit from this research. Firstly, because the T.A.B. do not conduct these sorts of tests on their data, they could use the results in numerous ways to make the betting market more efficient. Inefficiencies arise from three aspects of information. These are; a lack of information, a lack of information flow (where information is present in a market, but not everyone has access to it), and where a bettor has access to information but does not consider it when making their bets. This thesis could then make the market more efficient since it introduces new information.

By no means is this thesis an attempt to explore the efficiencies of a racing market in its entirety. New ideas and avenues are continually explored in an attempt to fully understand the intricacies of such a market. It is my hope that this thesis could be used in conjunction with further research to explore future ideas. Only through the continuation of research could we possibly understand the workings of such a market.

Finally, this thesis may induce some bettors to reassess where they should be investing their money.

1.5 Thesis Outline

Chapter 2 contains an extensive literature review followed by a chapter on research methodology, which defines exactly how each test will be conducted. Chapter 4 is devoted to disclosing the data; its size, origins and limitations. Chapter 5 then looks at the results of each test. The discussion of all the results and conclusions will come in Chapter 6, followed by a section on general conclusions in Chapter 7. Chapter 8 will discuss further research ideas.

2. Literature Review

This chapter reviews a cross-section of literature on market efficiency in the horse racing industry, the favourite longshot bias, the three moment model and Henery's (1985) test. It covers the relevant theories and in some instances, the empirical work. Each of these will be tested in this thesis using New Zealand data.

2.1 Efficiency

This thesis is predominately concerned with testing whether the New Zealand horse betting market is efficient. This requires a formal definition of efficiency. Fama (1970, p.383) defines efficiency as "a market in which prices 'fully reflect' available information". To extend this idea, Fama (1970) devised three efficiency categories; weak form efficiency, semi strong form efficiency and strong form efficiency. Firstly, a market is said to be weak form efficient if historical information is fully reflected in the price. In a racing context, this historical information is the knowledge of past winners, track conditions, jockeys, winning margins, weight carried, training runs etc. This kind of information is contained in racing form guides such as 'Best Bets' and the 'Turf Digest'. Fama (1970) states that if bettors include all this information in their decision making, then the odds will reflect this information and the market will be weak form efficient.

A semi strong form efficient market occurs where current information is reflected in the price. From a racing point of view, current information includes information that is made available throughout the betting process, such as scratchings of horses, their agitation levels, or even sudden dramatic changes in the odds. If bettors include this information in their decision making, it will be reflected in the odds and the market will be semi strong efficient. Very little research has been done in this area. Semi strong efficiency will not be covered in this thesis.

Fama (1970) concludes that the market is strong form efficient if all participants have equal information, that is, no group of investors have superior information. A group that possess superior information are referred to as insider traders. In a racing market, insider traders are those who actively participate in the market on a full-time basis. These people could be horse trainers, jockeys, owners, gambling addicts or simply punters that absorb exceptional amounts of racing information through an obsessive fixation with the market. If a market does not have a group like this, then according to Fama (1970), it is strong form efficient.

Note that this thesis looks at the racing market in the context that it is an information market. There will obviously be several efficiency issues relevant to the racing market as a service industry. Some examples of efficiency criteria in terms of a service industry could be, 'are there an efficient number of racetracks in New Zealand?' or 'are there too many betting agencies?'. These efficiency issues are beyond the confines of this thesis.

The weak form, semi strong form and strong form efficiency tests, collectively known as the efficient markets model, were developed by Fama (1970) to study a stock market. This model is adapted in this thesis to fit the betting market. Whether Fama's (1970) model can be adapted lies in the similarities between the two markets. If the markets are similar, using Fama's definitions should pose no problems when applied to a betting market. If however, there are crucial differences, this adaptation may be inappropriate. Upon investigation, a stock market is very similar to a betting market. Firstly, and most importantly, both markets have uncertain outcomes. The most competent financial advisor could never predict the outcome of a stock with 100% accuracy. He/she may investigate the information regarding a stock to better predict the outcome, but irrespective of the amount of information gathered, unforeseen external shocks will always result in a stock having an associated probability distribution. Similarly, a bettor will never know with certainty what the outcome of a race will be. Like the financial advisor, the bettor could not eliminate all uncertainty surrounding a horse race.

Secondly, in both markets there are only three possible outcomes. These are:

- Make a positive profit.
- Break even.
- Make a negative profit.

In deciding whether to invest money in each market, investors will assign subjective probabilities to each of these outcomes. The subjective probability is the best estimate the bettor can make as to the outcome. Also attached to each of these outcomes is an objective probability, which is the probability distribution when all information is known. Hence, it reflects a horses true chance of winning. Ideally, investors would like to know the objective probabilities, as then there would be no uncertainty surrounding the investment. The investors decision criteria would then be simple; invest if the expected return is greater than my discount rate, otherwise steer clear. Hence, when investors assign subjective probabilities to their outcomes, it is their intention to correctly predict the objective probability. This goal is not always achieved, but there are steps an investor can take to improve the accuracy of predicting the objective probability. The objective probability could be termed the 'full information' probability, so any movement toward full information must make the prediction of the objective probability easier. This is why bettors absorb information in both the stock market and the betting market. An investor will almost certainly seek advice to better understand the risk involved. This could be achieved by reading the latest financial report, or seeking the opinion of their financial advisor.

It is impossible for a bettor to place money on a horse without the accumulation of knowledge. At the very least, the bettor will need to know the number of the horse. Like the investor, bettors tend to feast on information. Publications such as the "Turf Digest", "Best Bets" and the "Friday Flash" contain an exhaustive amount of information for bettors to absorb. Bettors also source information from newspaper articles, friends and family and by examining the changing odds. In some ways, this latter source of information is the most important as it reflects the opinions of all bettors. An individual with no knowledge of horses could, at least through examining the odds, roughly decipher which horses have a higher chance of winning than others.

The most notable way in which the two markets differ is in their expected returns. For the most part, the expected return for a stock is positive, while the expected return for a bet is always negative. The reasons why a better would invest in a 'negative sum game' will be investigated later.

Thaler and Ziemba (1988) explain that betting markets are better suited for testing market efficiency than a stock market, because a betting market has a well defined termination point when the value of the bet becomes certain. A stock's termination point is never reached, hence its value is dependent upon its future earnings and is therefore never certain. Thaler and Ziemba (1988) believe that because a betting market has a quick, repeated feedback mechanism, this should facilitate learning. They believe that this constant learning from historical data occurrence could mean that a betting market is more efficient than a stock market.

2.2 Testing Fama's (1970) Weak Form Efficiency

Now that the similarities and differences between the stock and betting markets have been discussed, Fama's (1970) definition of weak form efficiency must be adapted to test the betting market. Recall that weak form efficiency occurs when historical information is fully reflected in the odds. The weak form efficiency test, most commonly used by researchers, is to determine the expected returns for different betting strategies. This is how Snyder (1978) approached weak form efficiency when he studied 1730 races during the 1972-74 season. In conjunction with his data set, he used five previous studies by Fabricant (1965), Griffith (1949), McGlothlin (1956), Seligman (1975) and Weitzman (1965)[5] to test for evidence of weak form efficiency. The test investigates whether knowledge about the subjective odds assigned by bettors through their dollars placed can be used to earn an above average return, where the subjective odds are simply the closing odds quoted by the tote. Essentially, Snyder tested whether adopting different strategies would yield the same return. Adapting Fama's (1970) definition, a market is weak form efficient if, irrespective of the strategy you adopt, your return on one dollar is one minus the track take. If this holds, "horse racing would then be said to fully reflect the publics use of all available information and consequently the market could be called efficient" (Snyder, 1978, p.1110). To conduct his test, Snyder (1978) grouped horses into eight odds categories and calculated each groups rate of return using the formula

$$RR = \frac{W(O+1)-N}{N}$$

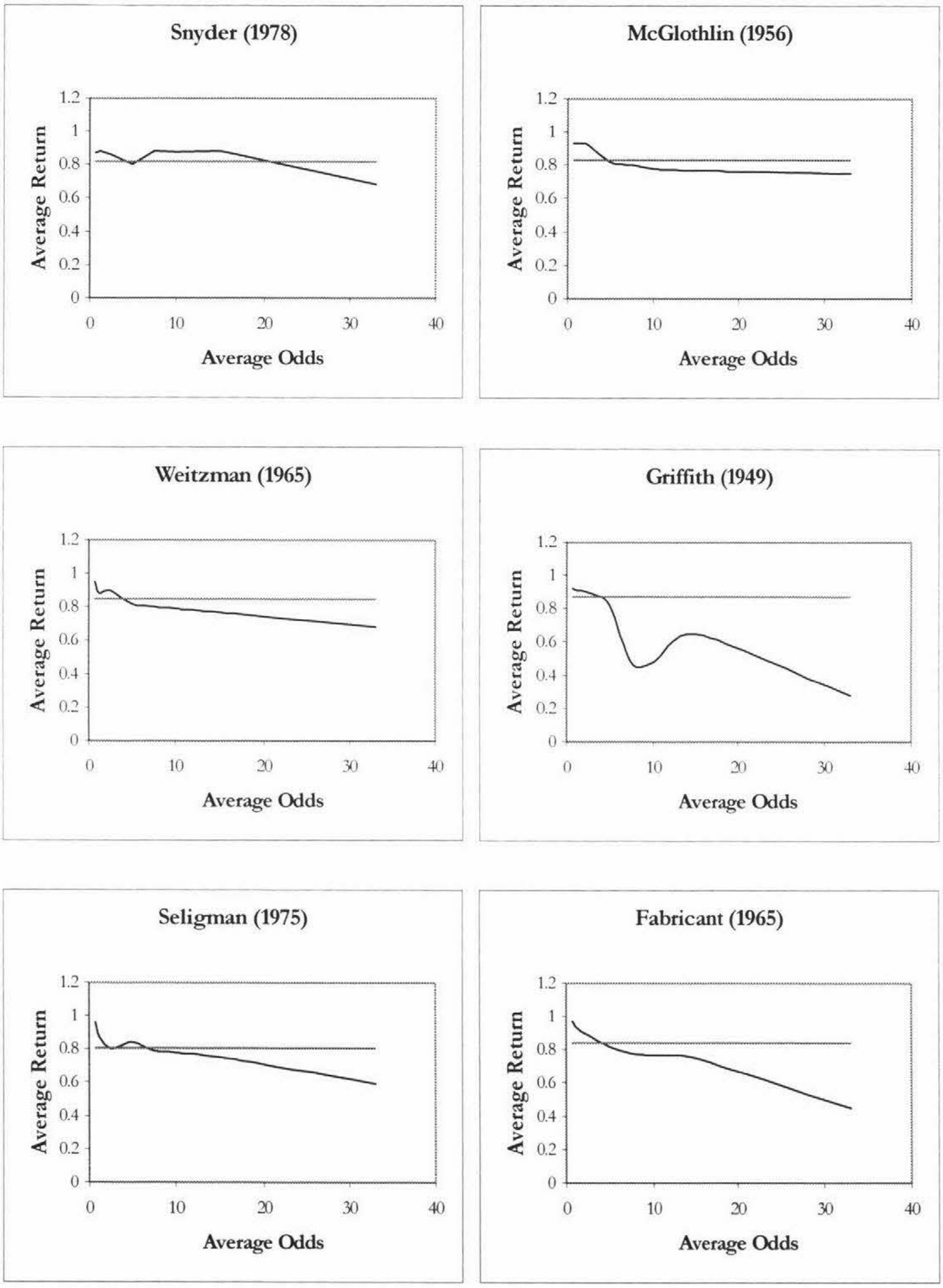
where W is the expost number of winning horses in each group, O are the odds[6] and N is the total number of horses in each group. Figure 2.1 shows Snyder's results as well as the results of the five previous studies mentioned above. Weak form efficiency requires a straight line through one minus the track take[7]. The graphs indicate that none of these six studies show evidence of weak form efficiency. In fact, in Fabricant and Griffith, the average returns across the odds categories varied wildly. In Fabricant's study of 1000 horses from 1955-62, the continued backing of a horse with odds of roughly $1 would have lost a bettor around 4 cents in every dollar, while the bettor would have lost approximately 55 cents in the dollar if they continued to back horses with odds of about

[5] Weitzman did not publish his data, but it was used again and published by Rosett (1965)

[6] The plus one is necessary because the odds stated in the U.S.A. do not include the initial $1 wagered.

[7] The track takes in these studies are approximately Fabricant 16%, Griffith 13%, McGlothin 17%, Seligman 19%, Weitzman 15% and Synder 18%.

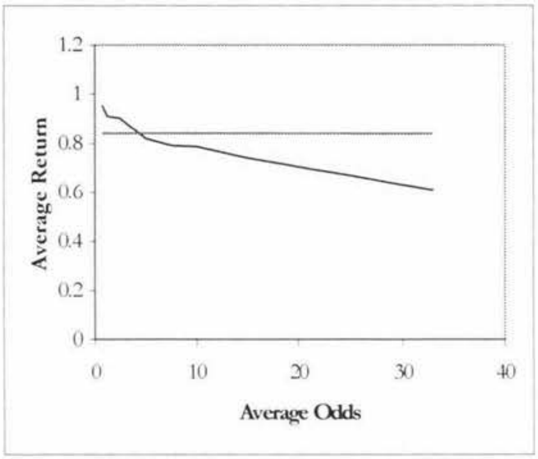
Figure 2.1: Weak Form Efficiency Results of Snyder (1978), McGlothin (1956), Weitzman (1965), Griffith (1949), Seligman (1975) and Fabricant (1965)



Note: Estimated figures only as actual observation data was not given in Snyder (1978). The horizontal line represents the track take and is therefore the weak form efficiency line.

$40. As for Griffith, the spread is even more pronounced, ranging from a loss of approximately 8 cents for every dollar (at odds of approximately $1) to 78 cents in the dollar (at odds of approximately $33). Overall, each study exhibits a negative relationship between the average return and the average odds categories. This outcome is even more evident when all studies are combined, as Snyder (1978) did. Figure 2.2 presents the results of this amalgamation that now covers 35,285 races consisting of over 300,000 horses between 1947 and 1975. The conclusion can be made that there is little evidence of weak form efficiency.

Figure 2.2: Combined Weak Form Efficiency Results of Snyder (1978), McGlothin (1956), Weitzman (1965), Griffith (1949), Seligman (1975) and Fabricant (1965)



A paper by Henery (1985) also explored the existence of weak form efficiency using two data sets, the first set for the 1973 flat season covering 22,109 horses, and the second set covering 883 races in the 1979-80 flat season. Like Snyder (1978), Henery divided his data set into groups according to their odds and determined the expected returns for each. He broke his first data set into 19 categories and the second into 12. The results from each set are presented in Figure 2.3. Because of the negative relationship exhibited in the graphs, we can conclude that there is no evidence of weak form efficiency.

Swindler and Shaw (1995) presented a study that most closely fitted Fama's criteria of weak form efficiency. They divided 288 races at the Trinity Meadows Raceway (Weatherford, Texas) during 1991 into 8 odds categories and found expected returns using the same equation as Snyder (1978). Their results are depicted in Figure 2.4 below. The track take in Swindler and Shaw's (1995) study was 18.8%.

Figure 2.3: Henery's (1985) Weak Form Efficiency Results

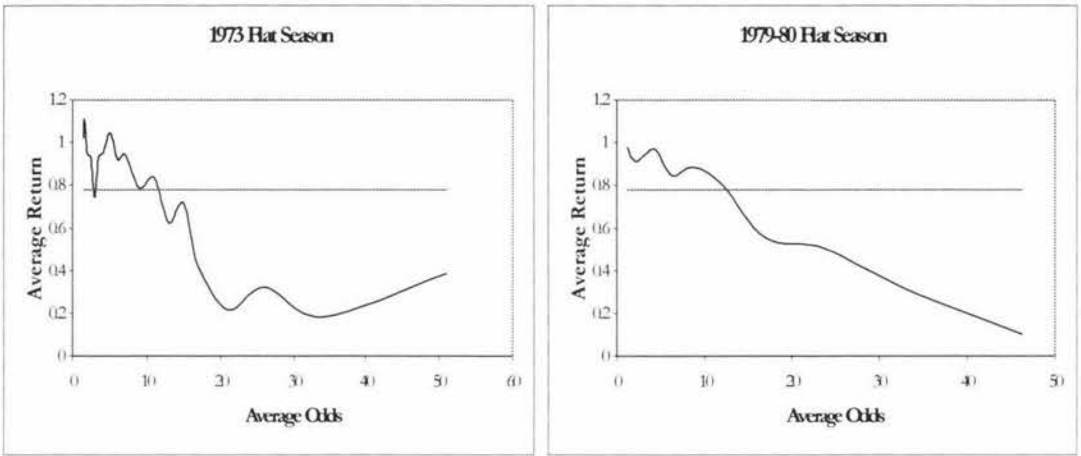
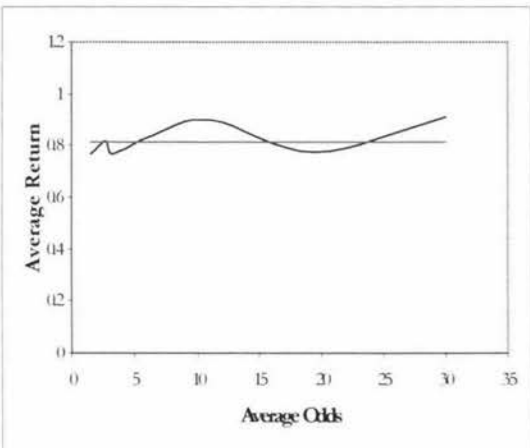


Figure 2.4: Swindler and Shaw's (1995) Weak Form Efficiency Results



Sharpe (1970) commented that the risk-return trade-off existing in a betting market is opposite to that existing in a stock market. In a stock market, an investor requires a greater rate of return for assuming greater risk. However, in a betting market, the previous studies, except Swindler and Shaw (1995), have shown that if you assume more risk (back a horse with higher odds), then you can expect a lower return.

There are differing 'schools of thought' as to why this bettor bias exists. Psychologists Griffith (1949, p.293) and McGlothin (1956, p.614) explained that it is a "... reconfirmation of a general tendency for all bettors to prefer low probability-high prize combinations over high probability-low prize combinations, findings first reported in laboratory experiments by Preston and Baratta (1948)". Economists Weitzman (1965) and Ali (1977) concluded that bettors must possess a utility function for money which is concave upward, indicating that they are risk lovers. Another thought is that bettors receive a level of utility from participating in betting that outweighs the disutility of the money they lose in the process. Hence, they have a net gain in utility. Henery (1985) attempted to explain the bettor bias by exploring the possibility that bettors discount a constant fraction of their losses, and therefore underestimate their true losses.

Golec and Tamarkin (1998) explain that the bettor bias exists because bettors love skewness and not variance. These two ideas will be discussed later.

All of these studies, except Swindler and Shaw (1995), conclude that there is a tendency for bettors to over-bet longshots and under-bet favourites. This anomaly is commonly referred to as the favourite longshot bias.

2.3 Favourite Longshot Bias

A favourite longshot bias exists where for favourite horses, the subjective probability is less than the objective probability, and for longshots, the subjective probability is higher than the objective probability. Hence, investors are not accurately determining each horses true chance of winning. The favourite horse is tending to win more often than bettors are predicting, while the longshot is winning less than predicted. Numerous studies have been performed regarding this phenomenon. Asch, Malkiel and Quandt (1982) determined subjective and objective probabilities for 5805 horses in 729 races at Atlantic City in 1978. They adopted the notation

$$B = \sum_{h=1}^{H} b_h$$

where B is the total amount in the win pool, b_h is the total amount bet on horse h where h = 1, 2, 3,...,H and H is the total number of horses in the race. The odds for each horse are determined by

$$O_h = \frac{B(1-t)-b_h}{b_h} = \frac{B(1-t)}{b_h} - 1$$

where t is the track take. The bettors subjective probability of horse h winning is

$$P_h = \frac{b_h}{B} = \frac{(1-t)}{(1-O_h)}$$

From here, it is easy to determine the subjective probabilities accurately for each horse. The track take and the odds can be readily observed. However, there is a problem in determining the objective probability of a horse winning. Without actually knowing the genetic composition of each horse and all the conditions affecting it, the true winning potential cannot be accurately determined. To approximate the objective probability, the horses must be categorised. Historically, researchers have chosen two methods to categorise horses; by order of favouritism or by subjective probability categories. Both methods have problems associated with them and these will be discussed in a later chapter. Asch et al (1982) categorised horses according to their order of favouritism. The objective probability is calculated using the equation

$$OB_h = \frac{W_i}{H_i}$$

where W_i is the number of winning horses in the i^{th} category, and H_i is the total number of horses in the i^{th} category. To measure the extent to which the subjective probability deviates from the objective probability, Asch et al (1982) determined a z statistic using

$$z = \frac{P_h - O_h}{\sigma_{O_h}}$$

where the standard error of the objective probability (σ_{Oh}) is $[O_h(1-O_h)/n]$ and n is the number of horses in each category. Asch et al (1982) computed these standard errors taking the objective probabilities as the 'true' probabilities and assumed a binomial process. The results are presented in Table 2.1.

Table 2.1: Subjective and Objective Probabilities of Winning in 729 Atlantic City (NJ) Races in 1978 (total number of horses = 5805)

Favourites	No. of Races	Objective Probability	Subjective Probability	z Statistic
1^{st}	729	0.361	0.325	-2.119*
2^{nd}	729	0.218	0.205	-0.903
3^{rd}	729	0.170	0.145	-1.972*
4^{th}	724	0.115	0.104	-0.961
5^{th}	692	0.071	0.072	0.074
6^{th}	598	0.050	0.048	-0.279
7^{th}	431	0.030	0.034	0.480
8^{th}	289	0.017	0.025	1.096
9^{th}	165	0.006	0.018	2.095*

* Significant at the 5 percent level.

For the favourite longshot bias to exist, the favourites need significant negative z values and the longshots need significant positive z values. The data follows this trend, however, note that in only 3 categories are the deviations significantly different from zero at the 5 percent level.

A more conclusive test was performed by Ali (1977) when he studied 20,247 harness horse races. Like Asch et al (1982), Ali (1977) categorised horses into order of favouritism to determine the objective probabilities. Ali's results are presented in Table 2.2

Table 2.2: Subjective and Objective Probabilities of Winning in 20,247 Harness Races by Ali (1977)

Favourites	No. of Races	Objective Probability	Subjective Probability	z Statistic
1st	20,247	0.3583	0.3237	-10.29*
2nd	20,247	0.2049	0.2077	0.99
3rd	20,247	0.1526	0.1513	-0.52
4th	20,247	0.1047	0.1121	3.45*
5th	20,231	0.0762	0.0827	3.49*
6th	20,088	0.0552	0.0601	3.01*
7th	19,281	0.0431	0.0417	5.80*
8th	15,749	0.0206	0.0276	6.20*
9th	299	0.0033	-	-
10th	71	0.0141	-	-

Note. Subjective probabilities for favourites 9 and 10 are not shown because they are based on a very small number of observations.
* Significant at the 5 percent level.

The 2nd and 3rd favourites are not significantly different from zero, however the 1st, 4th, 5th, 6th, 7th, and 8th favourites provide very strong evidence for the favourite longshot bias.

Bird, McCrae and Beggs (1987) took the favourite longshot bias analysis one step further. Using data from three Melbourne tracks on 1026 races in the 1983 and 1984 calendar years, they studied how the favourite longshot bias changed throughout the betting process. Their data came from a fixed odds bookmaking system, where bookmakers offer odds throughout the betting period. Over the length of the betting period, bookmakers continually reassess the odds. In doing this, they take into account the amount of money being placed on each horse. So even though each bookmaker determines the odds, these soon reflect the overall opinion of the bettors. Bird et al (1987) divided the betting period into four sub periods at which time the odds were observed. These periods were t_1 = -30 minutes, t_2 = -25 minutes, t_3 = -13 minutes and t_4 = -1 minute. They calculated the subjective and objective probabilities using the same method as Asch et al (1982) and Ali (1977). Their results are in Table 2.3.

Table 2.3: Subjective and Objective Probabilities of Winning in 1026 Races by Bird, McCrae and Beggs (1987)

	Period							
	t_1		t_2		t_3		t_4	
Favourite (Sample Size)	**Subj. Prob.**	**Obj. Prob.**	**Subj. Prob.**	**Obj. Prob.**	**Subj. Prob.**	**Obj. Prob.**	**Subj. Prob.**	**Obj. Prob.**
1 (1026)	0.2567	0.3000**	0.2598	0.3060**	0.2732	0.3171**	0.2806	0.3179**
2 (1026)	0.1686	0.1700	0.1693	0.1733	0.1728	0.1810	0.1740	0.1890
3 (1026)	0.1296	0.1540**	0.1301	0.1404	0.1288	0.1378	0.1296	0.1283
4 (1026)	0.1017	0.1014	0.1020	0.1104	0.1008	0.0940	0.1003	0.0933
5 (1026)	0.0811	0.0705	0.0809	0.0661*	0.0793	0.0644*	0.0793	0.0666
6 (1024)	0.0635	0.0540	0.0647	0.0570	0.0629	0.0632	0.0626	0.0607
7 (1009)	0.0524	0.0440	0.0520	0.0435	0.0528	0.0390**	0.0495	0.0403
8 (985)	0.0421	0.0311**	0.0415	0.0369	0.0399	0.0277**	0.0385	0.0334
9 (937)	0.0340	0.0236*	0.0333	0.0214**	0.0318	0.0219*	0.0303	0.0206*
10 (874)	0.0274	0.0284	0.0267	0.0273	0.0250	0.0305	0.0239	0.0266
11 (791)	0.0225	0.0180	0.0218	0.0220	0.0200	0.0174	0.0189	0.0144
12 (690)	0.0186	0.0095**	0.0180	0.0101*	0.0163	0.0098	0.0149	0.0116
13 (545)	0.0162	0.0117	0.0157	0.0131	0.0141	0.0121	0.0126	0.0115
14 (414)	0.0138	0.0099	0.0135	0.0078*	0.0118	0.0060*	0.0105	0.0060
15 (274)	0.0125	0.0063	0.0122	0.0069	0.0105	0.0085	0.0091	0.0091
16 (193)	0.0111	0.0043	0.0109	0.0041*	0.0088	0.0069	0.0075	0.0052

** Difference between subjective and objective probabilities significant at the 5 percent level
* Difference between subjective and objective probabilities significant at the 1 percent level

Like Asch et al (1982) and Ali (1977), a z statistic was computed to measure the extent to which the objective probability differed from the subjective probability. "The difference is positive and significant for all the first favourites with a scattering of significant negative differences for the 5th to 16th favourites" (Bird et al, 1987, p.239). Overall, the favourite longshot bias seemed to be more pronounced at t1, t2 and t3. As the close of the tote got closer, bettors seemed to eliminate the bias as there was little evidence of the favourite longshot bias at t4.

While there is a great deal of evidence to suggest that the favourite longshot bias frequently exists, Busche and Hall's (1988) study suggests otherwise. Busche and Hall studied 2653 races from the Happy Valley and Shatin racecourses in Hong Kong for the racing seasons of 1981-82 through to the middle of the 1986-87 season, consisting of more than 26,000 horses. They categorised horses by pari-mutuel odds groups and plotted betting odds (subjective probabilities) on the Y-axis versus observed win odds (objective probabilities) on the X-axis and ran a regression line through the observations. It is argued that if bettors are risk neutral and make unbiased and accurate predictions, returns must be equated across horses of differing win odds and the regression line would have a slope of one. A slope less than unity would provide evidence for the favourite longshot bias, and a slope greater than unity would provide evidence to reject the bias. Busche and Hall (1988) estimated the line for the collective six studies of Fabricant (1965), Griffith (1949), McGlothlin (1956), Seligman (1975), Snyder (1978), and Weitzman (1965) and found the equation to be (standard errors in parentheses)

$$\underset{}{\text{Betting odds}} = \underset{(0.403)}{1.144} + \underset{(0.023)}{0.747}\,\text{win odds} \qquad R^2 = 0.993$$

Note the slope was 11 standard deviations from one, which is significant beyond the 0.001 level and confirmed that the combined studies of Fabricant (1965), Griffith (1949), McGlothlin (1956), Seligman (1975), Snyder (1978), and Weitzman (1965) exhibited the favourite longshot bias. The estimated regression line for Busche and Hall's (1988) Hong Kong data was (standard errors in parentheses)

$$\text{Betting odds} = \underset{(1.40)}{-2.908} + \underset{(0.036)}{1.251}\,\text{win odds} \qquad R^2 = 0.99$$

The slope was approximately 7 standard deviations from one, which is also significant beyond the 0.001 level and hence provided strong evidence to suggest that the opposite of the favourite longshot bias existed in Hong Kong.

2.5 Henery's Test

In pursuit of trying to explain the seemingly risk loving behaviour of bettors, Henery (1985) presented a model which accounted for some of the empirical patterns of betting losses on horses. He proposed that punters discount a constant fraction (1-f) of their losing bets, so that they believe their chance of losing is fq, where q is the true chance of losing. Henery (1985) suggested that this behaviour could be explained by the natural tendency to discount certain losses as 'not typical', as external factors might have influenced the result, such factors that are not related to the ability of the horse. Examples may be where a horse is impeded or is disqualified after having won the race.

Lets say the true probability of losing the race is q so the true odds against winning are q/(1-q) and the punter only rates their chance of losing as Q=fq. The bettor will accept odds of Q/(1-Q) = fq/(1-fq) against the horse winning. To illustrate this, suppose a horse's true probability of losing is 90%, therefore its true odds against winning are 0.9/(1-0.9) = \$9. Now let's assume that the punter rates their chance of losing as Q = 0.98q. In this case, the punter would actually accept odds of 0.98(0.9)/(1-0.98(0.9)) = 0.882/1-0.882 = \$7.47. As you can see, this is a very powerful model. A punter discounting only 2% of his/her losses is willing to accept odds of \$7.47 on a horse that should be paying \$9. We have already seen evidence of bettors accepting odds that do not reflect a horse's true potential in the favourite longshot bias data presented earlier. The 9th group of favourites in the Asch et al (1985) study had average odds of \$45.25[8]. Given that the objective probability of these horses are 0.006, the odds that accurately reflect these horses true potential are \$135.80. It is this huge difference that Henery (1985) attempted to explain.

Henery (1985) examined two data sets. The first data set was taken from his earlier study (1984) which consisted of 883 selected races in the flat seasons of 1979 and 1980. The second was 22,109 horses in 1417 races from the 1973 flat season taken from Figgis (1976). The horses were grouped into odds classes, with class limits chosen on a logarithmic scale. He chose the mid-point X_j of the j^{th} class using the formula

$$1 + X_j = \exp\left(\frac{j - 0.5}{3}\right)$$

Two estimates of the probability of losing were considered. Firstly, the true probability of losing, q_j, was found by simply finding the proportion of horses that lost for each odds category. Secondly, the subjective losing probability, Q_j, was found by $Q_j = X_j/(1+X_j)$. Henery (1985) then plotted the q_j, Q_j combinations for each of the odds categories and estimated the equation $Q_j = fq_j$. He found that the function fitted the data well, however, he noted that in the 1973 data, there were two instances of significant deviations.

> "The fit to a straight line is much better for the 1979-1980 data, possibly because the races were selected from the middle to the end of the season (from July 6 onwards) and the merits of the horses would be better known, and also because only those horses were chosen where the prize money was more than 1000 pound." (Henery, 1985, p.345).

Henery (1985) found that f was 0.978 which suggested that bettors discount approximately 2.2% of their bets.

---

[8] Using a track take of 18.5%

2.4 Testing for Skewness

From the extensive work on the favourite longshot bias, many researchers have concluded that horse race gamblers are risk lovers (see Weitzman (1965), Ali (1977), Quandt (1986), Kanto, Rosenquist and Suvas (1992)). Researchers such as Golec & Tamarkin (1998), Bird, McCrae and Beggs (1987) agree with this but only if the first two moments around the distribution are analysed. They claim that this attitude can be explained by analysing the third moment around the distribution – skewness. They argue that bettors accept low-return, high-variance bets because they enjoy the high skewness offered by these bets. Bird et al (1987) tested the third moment in their study of three Melbourne racetracks. To calculate the skewness they used the equation

$$K_{mt} = \frac{T_{mt}}{(V_{mt})^{\frac{3}{2}}}$$

where

$$T_{mt} = \frac{1}{N_m} \sum_{n=1}^{N_m} (R_{mnt} - R_{mt})^3$$

and

$$V_{mt} = \frac{1}{N_m} \sum_{n=1}^{N_m} (R_{mnt} - R_{mt})^2$$

where R_{mnt} = return from a \$1 bet on the horse classified as the m[th] favourite at time t in race n

O_{mnt} = the bookmakers unadjusted odds quoted on the horse classified as the m[th] favourite at time t in race n

F_{mnt} = 1 if the horse classified as the m[th] favourite at time t wins race n

0 if the horse classified as the m[th] favourite at time t loses the race n

They then ran the regression

$$r_{mnt} = a + b_1 V_{mt} + b_2 K_{mt} + e_{mt}$$

where r_{mt} = average rate of return for the \$1 bet on the m[th] favourite at time t.

V_{mt} = variance of the return distribution of a bet of \$1 on the m[th] favourite at time t

K_{mt} = skewness of the return distribution of a \$1 bet on the m^{th} favourite at time t

e_{mt} = stochastic disturbance term for the m^{th} favourite at time t

Their results are reproduced in Table 2.4

Table 2.4: OLS Single Equation Estimates and Test Statistics by Bird, McCrae & Beggs (1987)

Variables	Period			
	t_1	t_2	t_3	t_4
Variances	0.029 (3.73)	0.026 (4.29)	0.032 (6.83)	0.022 (8.00)
Skewness	-0.063 (8.99)	-0.064 (6.22)	-0.091 (8.76)	-0.080 (10.20)
Constant	-0.222 (4.57)	-0.206 (3.41)	-0.113 (2.37)	-0.082 (2.03)

Note. The OLS t-statistics are given in the parentheses

The highly significant positive variance coefficients suggest that gamblers demand a higher return for accepting a higher variance in the return distribution. The negative value on the skewness coefficient suggests that bettors are willing to accept lower returns for bets that offer a lower probability of high returns. Bird et al (1987) concluded that it is skewness that bettors enjoy, not variance like other researchers claim.

Golec and Tamarkin (1998) found evidence that horse race bettors accept low-return, high-variance bets, because they enjoy the high skewness offered by these bets. They used the notation of Quandt (1986), where the amount bet on horse h in any race is b_h. The total amount bet on N horses in any race is normalised to unity so that

$$\sum_{h=1}^{n} b_h = 1$$

The subjective win odds for horse h are

$$O_h = \frac{1-t}{b_h} - 1$$

where t is the track take. The return mean, μ, and variance, σ^2, on horse h are given by

$$\mu_h = p_h\left(\frac{1-t}{b_h}\right) - 1$$

$$\sigma^2 = p_h(1-p_h)\left(\frac{1-t}{b_h}\right)^2$$

where p_h is the objective probability. The formula used to calculate skewness is

$$M_h^3 = \frac{p_h(1-p_h)(1-2p_h)(1-t)^3}{b_h^3}$$

It follows that if $b_h=p_h$, then $\mu=-t$. Hence, given a two moment world, bettors must be risk lovers since outcomes have positive variances and the expected return from their wagers is negative. While Golec and Tamarkin (1998) conceded that this seemingly risk lover attitude could be explained by the utility gained from entertainment, they believed this was only relevant for a proportion of bettors, since some punters take their betting very seriously and bet significant sums of money. Golec and Tamarkin (1998) believed that since most economic models assume that individuals follow risk averse preferences, then the challenge is to explain the observed longshot anomaly while maintaining the assumption of risk averse behaviour. Hence, they set out to explain the anomaly through the introduction of the third moment. They believed that "when bettors willingly select low-return, high-variance bets, both unattractive features, they do so because these bets also offer higher skewness than high-return, low-variance bets" (Golec and Tamarkin, 1998, p.208).

> "... a three moment model can explain the long shot anomaly under two conditions: (1) both variance and skewness increase as p_h decreases, and (2) the weights bettors place on mean, variance, and skewness are such that the utility derived from skewness appropriately offsets the disutility of negative expected returns and positive variance" (Golec and Tamarkin, 1998, p.208).

If $p_h=b_h$, then condition (1) almost always holds. As p_h decreases and we move from favourites to longshots, it can be verified from the mean and variance equations that they both increase monotonically as long as $p_h>2/3$[9]. It also follows that as long as $p_h<2/3$, skewness is a positive function of variance.

If p_h and b_h differ significantly, then it is the change in the ratio of p_h to b_h that determines whether variance and skewness increase together as p_h declines[10]. The

[9] Very few horses have an objective probability greater than 2/3.

[10] Note that the ratio of p_h to b_h is dependant upon condition (2), which is unknown.

favourite longshot bias implies that $p_h > b_h$ for favourites and $p_h < b_h$ for longshots, hence as Golec and Tamarkin point out "... whether variance and skewness increase together over the relevant probability range is an empirical issue." (1998, p.209)

Golec & Tamarkin (1998) studied several eastern racetracks for the years 1990-92. They broke their sample into two categories, an unconditional sample and a conditional sample. A race was included in the conditional sample if it contained a horse with a subjective win probability greater than 0.50[11]. This sample consisted of 840 races with 6,140 horses from 27 tracks. The unconditional sample had no restrictions and consisted of 1,469 races with 12,749 horses from 8 tracks. The reason that Golec and Tamarkin (1998) over-sampled high-probability horses in their conditional sample was because they claimed that traditional studies tended to underweight favourites. This is because significantly more money is bet on a favourite than a longshot, while each horse is given an equal weighting in any analysis. They believed that if a horse attracted a significantly greater amount of money, then it should be given greater weighting. By having a sample of heavy favourites (and therefore heavy longshots), Golec & Tamarkin divided their sample down the middle so they had two samples, one of heavy favourites and one of heavy longshots. The utility estimates for each sample were then compared to see how each type of horse effects utility.

Golec & Tamarkin (1998) developed the bettors utility function by extending the study of Ali (1977). Like Ali (1977), they assumed that bettors have identical utility functions, $u(\cdot)$; that they bet their full wealth, M; and that win odds are in units of M dollars. Further, a losing bet on horse h returns zero to the bettor and a winning bet returns X_h, where X_h equals one plus the win odds, O_h. The bettors expected utility is therefore

$$E(u) = p_h u(X_h) + (1 - p_h)u(0)$$

If $u(0) = 0$ and $u(X_H) = 1$ where H represents the highest odds horse, and if odds are set so bettors are indifferent between bets on any horse h in a race, then

$$E(u) = p_h u(X_h) = p_H u(X_H) = p_H$$

[11] It is more appropriate to select a race on the basis of the objective probability but since the objective probability of horses is not known, Golec & Tamarkin use subjective probabilities as a proxy.

Therefore, the utility function is defined as $p_H/p_h = u(X_h)$. Golec & Tamarkin (1998) approximated $u(X_h)$ by a Taylor series truncated to three terms and then took expectations[12]. The cubic utility function they estimated was

$$\frac{p_H}{p_h} = a + b_1 X_h + b_2 X_h^2 + b_3 X_h^3$$

where $b_2 < 0$ implies that bettors are risk averse, $b_2 = 0$ implies risk neutrality and $b_2 > 0$, a risk lover. Golec & Tamarkin (1998) hypothesised that $b_1 > 0$, $b_2 < 0$, $b_3 > 0$ which suggests bettors are risk averse and prefer mean and skewness. Golec & Tamarkin (1998) also fit a power utility function to the data of the form

$$u(X_h) = \alpha + \beta \log X_h + \varepsilon_h$$

where ε_h is a random error. It is the value of the β that determines whether a bettor is risk averse or not. A $\beta > 1$ implies u` > 0 and u`` > 0 which is the characteristic of a risk averse bettor. A $\beta > 1$ also implies skewness aversion which is consistent with risk loving behaviour. The results of Golec and Tamarkin (1998) are given below in Table 2.5

Table 2.5: Regression Estimates from Power and Cubic Utility Models using Conditioned Sample

Power Utility Model				
Sample	**Intercept**	**X**	**Observations**	**R^2**
All	-1.64*	1.20*	18	0.97
	(0.04)	(0.05)		
Favourites	-1.61*	1.08*	9	0.61
	(0.12)	(0.33)		
Longshots	-1.56*	1.13*	9	0.92
	(0.12	(0.13)		

Cubic Utility Model						
Sample	**Intercept**	**X**	**X^2**	**X^3**	**Observations**	**R^2**
All	-0.10*	0.089*	-0.008*	0.0003*	18	0.99
	(0.033	(0.017)	(0.002)	(0.00007)		
Favourites	-1.20**	1.65**	-0.71**	0.10**	9	0.86
	(0.59)	(0.77)	(0.33)	(0.05)		
Longshots	-0.15	0.11	-0.01	0.0004	9	0.98
	0.24)	(0.07	(0.007)	(0.0002)		

* Significant at the 5 percent level.
** Significant at the 10 percent level.

[12] Golec & Tamarkin indicate that function approximation and truncation of potentially important terms involve error by definition (see Loisetl 1976) and may also include some misspecification since they are applying a specific utility function when it may not be the true function.

When fitting the data with the cubic utility function, there is much support for the hypothesis that bettors prefer mean and skewness and are risk averse. With all the samples, Golec & Tamarkin (1998) obtained results supporting their hypothesis that $b_1 > 0$, $b_2 < 0$, and $b_3 > 0$. Only in the sample of the longshots were the coefficients not significant at least at the 10 percent level. However, when fitting the data with the power utility function, the results are rather conflicting. In all the samples, the β coefficient is greater than 1, implying risk-loving behaviour. This power utility function is consistent with the utility function proposed by Freidman and Savage (1948). The function starts out concave over the short odds and becomes convex over the long odds. The concave nature is consistent with risk averse behaviour, where bettors trade risk for return. At longer odds, bettors are willing to accept a lower return for enjoying high skewness, depicted by the convex nature of the utility function.

3. Research Methodology

3.1 Introduction

All tests in this thesis have been performed using standard statistical techniques previously employed by other researchers. Advanced spreadsheet techniques have been used to manipulate the vast quantity of data. All econometric analysis has been estimated using SHAZAM.

3.2 Calculation of Subjective and Objective Probabilities

The win odds of horses at the close of the tote and at 30 minutes before the close of the tote were supplied by the T.A.B (hereafter referred to as the closing and opening odds respectively). The subjective probability is measured as the proportion of the win pool that is bet on any one horse. Hence, if a horse attracts 15% of the win pool, it can be said that, collectively, the public believes the horse's probability of winning is 0.15. The subjective probability is thus

$$SP_i = \frac{b_i}{B_j}$$

where b_i is the amount bet on horse i and B_j is the total win pool for the j^{th} race in which horse i races. The odds on horse i are calculated as

$$OW_i = \frac{B_j(1-t)}{b_i}$$

where t is the track take. Rearranging this and substituting into the subjective probability equation above yields

$$SP_i = \frac{(1-t)}{OW_i}$$

The track take in New Zealand changed on the 1st of January, 1996. Pre 1996, the track take on the win pool was 18.4%. In 1996, this fell to 15.8%. The subjective probabilities have been calculated with the respective track take at the time. However, to calculate an approximate weak form efficiency return, the computer averages the track take over the whole sample, which equates to 16.9458%.

The calculation of the objective probability posses a few more difficulties. The objective probability is an estimate of each horse's true winning potential. It is impossible to try and estimate each horse's objective probability on a singular basis. Instead, to predict a horse's objective probability, horses are grouped such that all the horses in each group have a similar level of capability[13]. Once the groups are established, the objective probability is calculated as the proportion of winning horses per group, using the formula

$$OP_k = \sum_{i=1}^{n} \frac{W_i}{N_k}$$

where W_i equals 1 if the horse wins and 0 otherwise, and N_k is the total number of horses in category k. Using the assumption that each race is an independent trial, the variance of the objective probability is calculated using

$$Var(OP_k) = \frac{OP_k(1 - OP_k)}{N_k}$$

3.3 Grouping Horses

To conduct the tests in this thesis, the 118,101 horses must be divided into groups, such that each group contains horses of similar capabilities. Researchers have tended to do this in two ways; grouping by rank of favouritism and grouping by subjective probability categories. Problems arise with each method adopted.

3.3.1 Grouping by Rank of Favouritism

Grouping by rank of favouritism involves separating horses according to how they are ranked by the public. In the event that two horses from the same race have the same win odds, their favourite ranking is determined by their place prices, whereby the horse with the lower place price is deemed to be the more favoured of the two. However, the problem with separating horses with this method is that it is possible to have two horses in one group offering wildly different odds. For instance, in the group of 1st favourites studied in this thesis, one horse is paying $1.05 to win while another is paying $7.25. Given that the intention of grouping horses is to pool together horses of similar capabilities, this method may not be the most appropriate. Many researchers believe that two horses such as these do not have equal chances of winning, irrespective that

[13] The method of grouping, its problems and the methods used in this thesis are presented in the next section

the two horses are both ranked as the favourites when the tote closes. However, a horse's true chance of winning depends predominately on the opposition it faces. For the favourite to be paying $7.25, it needs to be in a race with equally capable horses. In this situation, the win pool would be divided relatively evenly amongst all of the horses. For a horse to be paying $1.05, it must stand out from its competitors so that a substantially higher proportion of the win pool is placed on it. Having said this, it is still possible for these two horses to have the same true chance of winning, since it is the calibre of its competitors that ultimately determines its odds, not its true winning potential. Irrespective of this potential problem, this method is adopted in this thesis. How it has been implemented is explained in the next section.

The horses are ranked according to their favouritism. This results in 18 categories, since the biggest race in the data set involves 18 horses. However, since there are only 185 18^{th} favourites and none of them won a race, they are amalgamated with the 274 17^{th} favourites to form the final grouping. Hence, this method consists of 17 groups (1^{st}, 2^{nd}, 3^{rd},...,16^{th} rank favourites and the 17^{th} and 18^{th} combined)

3.3.2 Grouping by Subjective Odd Categories

The other method of grouping horses is according to their subjective probability categories. Grouping horses this way means we assume that horses who receive similar support by the bettors have a similar chance of winning. The problem encountered with this method is that two or more horses could appear in the same category while racing each other. With the unlikely nature of dead heats, only one of these horses will win the race. Since the objective probability is computed as the proportion of winners in each category, this problem may understate the groups true winning potential. Golec and Tamarkin (1998) favoured this second method for two reasons. Firstly, they argued that bettors are aware that only one horse can win a race when they place their wagers and thus the odds are set appropriately. For this reason, they believed that the weaknesses of this method are less significant than the weaknesses of the favourite ranking method. Secondly, this method allowed them to create more categories and thus more observations for estimating utility functions. For the purpose of this thesis, this method has also been utilised to see how the results are affected by different grouping methods.

Three different methods of grouping horses according to their subjective probability categories are adopted. The first method is to group using even subjective probability categories. From 0 to 0.2, 16 categories are implemented at intervals of 0.0125. From 0.2 to 0.4, 8 categories are used at intervals of 0.025. 4 categories at intervals of 0.05 from 0.4 to 0.6 are used and finally 1 category incorporating subjective probabilities 0.6 through 1.

The second method is to allow the computer to derive 25 subjective probability categories such that each group has the same number of horses in it. This is achieved by allocating 4725 horses in the first group and 4724 to each of the 24 other groups.

The third method is to generate 50 subjective probability categories such that each group has the same number of horses in it. This is achieved by allocating 2363 in the first group and 2362 in each of the 49 other groups.

The grouping by rank of favouritism and the three methods by subjective probability categories are computed with the opening and closing odds. Hence, 8 groupings in total are calculated (4 at opening odds and 4 at closing odds).

3.4 Testing for Weak Form Efficiency

As previously discussed, a horse racing market is said to be weak form efficient when, irrespective of the strategy adopted, the expected return on a $1 bet is 1-t, where t is the track take. That is, under Fama's (1970) weak form efficiency definition, all historical information has been included in the odds. Using 4 horse groupings[14], expected returns are calculated on each category within the group. The equation used to find the expected return on placing a $1 bet on each horse in the k^{th} category is

$$ER_k = \sum_{i=1}^{n} \frac{W_i OW_i}{N_k}$$

where W_i equals 1 if the horse wins and 0 otherwise, OW_i are the win odds on horse i and N_k is the total number of horses in category k. To visualise whether there is weak form efficiency, average odds versus average returns are graphed for each group. Incorporated in these graphs is a horizontal line through one minus the track take

[14] Note that these expected returns were only calculated on the closing win price as it is these values that are returned to the bettors, not the open win prices.

depicting the line of weak form efficiency. If a strategy's expected return is not equal to one minus the track take, this is because the category has either been underbet (if expected return is greater than one minus the track take) or overbet (if expected return is less than one minus the track take). To formally measure the extent of underbetting and overbetting, the favourite longshot bias is examined.

3.5 Testing for the Favourite Longshot Bias on each Group

The favourite longshot bias is a term used to describe the phenomenon of favourites being underbet and longshots being overbet. Underbetting means the betting public believe that favourites have a worse chance of winning than evidence shows, that is, the subjective probability for favourites is less than the objective probability. For overbetting, the subjective probability is greater than the objective probability. To formally test for the existence of the favourite longshot bias, a z statistic is computed to estimate whether there is any statistical significance between the objective (OP_k) and subjective probabilities (SP_k). The z statistic is

$$z_k = \frac{(OP_k - SP_k)}{\left[\frac{OP_k(1-OP_k)}{N_k}\right]^{\frac{1}{2}}}$$

where N_k is the number of horses in the k^{th} category. The z statistic tests the hypothesis

$$H_0 : OP_k = SP_k$$
$$H_A : OP_k \neq SP_k$$

where the null hypothesis is the condition for weak form market efficiency and for the non-existence of the favourite longshot bias.

This test is conducted for all 8 groupings for two reasons, firstly, to ascertain whether the favourite longshot bias is more prevalent 30 minutes before the tote closes than when the tote closes and secondly, to determine whether the method of grouping has any effect on the favourite longshot bias results.

The computation of z statistics allows one to test the favourite longshot bias for a single category. To test for evidence throughout the entire group, the equation estimated is

$$OP_i = b_0 + b_1(SP_i) + e_i$$

where b_0 is the constant, b_1 is the slope and e_i is a random error term with usual least squares assumptions.

Of the 8 grouping methods, the favourite ranking method produces categories of different sizes. It is found that the variance of the dependant variable in the regression is inversely related to the number of horses in each subjective probability category. This produces heteroscedasticity in the Ordinary Least Squares (OLS) estimation but this is eliminated by using the Weight Least Squares (WLS) procedure where the weighting is $1/N_k$. The other grouping method producing categories of differing sizes is the grouping by even subjective probability ranges. However, with this method, OLS produces no heteroscedasticity and switching to WLS creates heteroscedasticity. Hence, the WLS method is only adopted for the grouping by rank of favouritism.

If the market is perfectly efficient with no evidence of the favourite longshot bias, then this would require $b_0=0$ and $b_1=1$. The magnitude of b_1 determines whether the favourite longshot bias exists. If $b_1>1$, there is evidence for the favourite longshot bias and if $b_1<1$, there is evidence against the favourite longshot bias. Two t tests are performed to test the hypothesis

$$H_0 : b_0 = 0$$
$$H_A : b_0 \neq 0$$

and

$$H_0 : b_1 = 1$$
$$H_A : b_1 \neq 1$$

Additionally, an F statistic is computed to test efficiency, where the hypothesis is

$$H_0 : b_0 = 0, b_1 = 1$$
$$H_A : b_0 \neq 0, b_1 \neq 1$$

3.6 Testing for the Favourite Longshot Bias throughout the Raceday

Researchers believe that the favourite longshot bias becomes more prevalent during the raceday[15]. This stems from the belief that bettors that start a day with a lot of money try to regain their losses by the end of the day. For some, the only way to regain all their

[15] As far as I am aware, this theory has been tested in a previous study. Despite an extensive search, I have not been able to find the relevant reference. Although the test here has been designed by me, the initial idea is credited to an unknown source.

losses is to bet on a winning longshot. For this reason, researchers believe that longshots will be more overbet by the final few races in the day and hence the favourite longshot bias will be more prevalent. To test this idea, two different methods are adopted. Firstly, the data is split according to race numbers[16] and then grouped according to their rank of favouritism to give 170 horse groupings (17 categories over 10 races). The subjective and objective probability pairs are calculated for each group. z values are calculated for each category and the equation estimated for each race using WLS is

$$OP_k = b_0 + b_1(SP_k) + e_k$$

This gives 10 estimates for b_1 corresponding to races 1-10. Recall that b_1 determines the existence of the favourite longshot bias where $b_1>1$ supports the bias. Let the 10 estimates of b_1 be defined as b_{1m}, where m indicates the race number. If b_{1m} increases throughout the raceday, this would indicate that the favourite longshot bias becomes more prevalent. To test for this, the equation estimated using OLS is

$$b_{1m} = \alpha + \beta(Race_m) + \varepsilon_i$$

To support the researchers claim that the b_{1m} coefficient increases throughout the raceday, the β coefficient needs to be significantly greater than zero. To formally test for this, the test performed is

$$H_0: \beta = 0$$
$$H_A: \beta \neq 0$$

To protect against the problem whereby the data set gets too heavily segregated, as in the previous method, a second method for examining this theory is adopted. This is to divide the data into two categories; races 1-7 and races 8-10. From here, the data is again grouped according to rank of favouritism. Once again, z values are calculated for each category to test the favourite longshot bias on a singular basis and the equation $OP_k=b_0+b_1(SP_k)+e_k$, which tests for the favourite longshot bias over all categories, is estimated for races 1-7 and races 8-10 using WLS. We now have estimates for

b_1 – for races 1-10

b_1 – for races 1-7

b_1 – for races 8-10

[16] The biggest race meetings in New Zealand consist of 10 races.

For simplicity, denote b_1 (races 1-10) as δ_0, b_1 (races 1-7) as δ_1 and b_1 (races 8-10) as δ_2. For evidence that the favourite longshot bias becomes more prevalent during the raceday we require δ_2 to be statistically significantly larger than δ_0 and δ_1. To test for this, three t statistics are calculated. Firstly, the hypothesis

$$H_0 : \delta_1 = \delta_2$$
$$H_A : \delta_1 < \delta_2$$

is tested using

$$\frac{\delta_1 - \delta_2}{\frac{[\sigma(\delta_1) + \sigma(\delta_2)]}{2}} \sim t_{17}$$

Secondly, the hypothesis

$$H_0 : \delta_0 = \delta_1$$
$$H_A : \delta_0 > \delta_1$$

is tested using

$$\frac{\delta_0 - \delta_1}{\sigma(\delta_0)} \sim t_{34}$$

Thirdly, the hypothesis

$$H_0 : \delta_0 = \delta_2$$
$$H_A : \delta_0 < \delta_2$$

is tested using

$$\frac{\delta_0 - \delta_2}{\sigma(\delta_0)} \sim t_{34}$$

Up until this stage, this thesis aims to answer two main questions. Firstly, is the New Zealand Racing Market weak form efficient and secondly, does there exist a favourite longshot bias? The next two tests by Henery (1985) and Golec and Tamarkin (1998) both attempt to explain why this bias exists. The study by Golec and Tamarkin (1998) also attempts to illustrate why bettors participate in a negative sum game.

3.7 Henery's Test

The interpretation of the favourite longshot bias is that punters are not correctly predicting each horse's true chance of winning. Punters tend to underestimate the favourite's ability and overestimate the longshot's ability. The implication of this is that for longshots, punters accept odds that do not reflect the horse's capabilities. We saw earlier that the 9th favourite in the Asch et al (1982) study had average odds of $45.25, yet computing its odds using the objective probability, reflecting its true chance of winning, yielded $135.80. It is these differences that Henery attempted to explain by comparing the empirical losing probability of horses with the subjective losing probability determined by punters.

Unlike Henery, who used only one grouping method, all 8 methods (4 at opening odds and 4 at closing odds) are utilised in this thesis to see if the results are dependant upon the grouping method used and how the results change between opening and closing odds. For each category k, the average win odds are computed using

$$OW_k = \sum_{i=1}^{n} \frac{OW_i}{N_k}$$

where OW_i are the win odds for horse i and N_k is the number of horses in category k. Suppose the true chance of a horse losing is q, so that the true odds against winning are q/(1-q). Henery believes punters rate their chances of losing as Q=fq where f is a constant between 0 and 1. Hence, bettors are discounting their losses by (1-f). It is the estimation of f that is central to Henery's test. To estimate the equation Q_k=fq_k, Q_k and q_k need to be estimated. q_k which is the empirical lose probability is computed as the fraction of horses which lose for the given odds group using

$$q_k = \frac{N_k}{\left(N_k - \sum_{i=1}^{n} W_i\right)}$$

Q_k which is the punters subjective losing probability is found using the average odds

$$Q_k = \frac{OW_k - 1}{OW_k}$$

Once Q_k and q_k are calculated for each category, f can be estimated by regressing the equation

$$Q_k = fq_k$$

We expect f to be a constant between 0 and 1, where 1 means bettors do not discount their losses at all, while 0 means they totally discount their losses. Any value of f between 0 and 1 offers evidence to explain the existence of the favourite longshot bias. The further f is from 1, the more the favourite longshot bias is explained by this model. To determine the spread of f, a 95% confidence interval is estimated around it.

3.8 Testing for Skewness

The following tests were conducted by Golec and Tamarkin (1998) to attempt to explain why bettors participate in a negative sum game. It was their intention to show that bettors are willing to accept low return bets because of the very high positive skewness associated with such tests. To conduct the tests, the first three moments around the distribution are calculated using the formulae

$$\mu_k = OP_k\left(\frac{1-t}{SP_k}\right)-1$$

$$\sigma_k^2 = OP_k(1-OP_k)\left(\frac{1-t}{SP_k}\right)^2$$

$$Sk_k^3 = \frac{OP_k(1-OP_k)(1-2OP_k)(1-t)^3}{SP_k^3}$$

where the track take is the average track take calculated at 0.169458.

For the three moment model to explain the favourite longshot bias, it is necessary that the variance and skewness increase as the objective probability declines. The first three moments will be calculated for each of the 8 groupings and are presented along with the objective probability to check for the condition above.

To attempt to explain why bettors participate in a negative sum game, the bettor's utility function is estimated to find evidence of risk averse behaviour.

Following Ali (1977), it is assumed that bettors have identical utility functions, u(·), and that they bet their full wealth, M. A losing bet on horse i returns zero and a winning bet returns OW_i. The bettor's expected utility is defined as

$$E(u) = OP_i u(OW_i) + (1-OP_i)u(0)$$

If u(0) = 0 and $u(OW_I)$ = 1, where I represents the highest-odds horse, and if odds are set so that bettors are indifferent between bets on any horse i in a race, then

$$E(u) = OP_i u(OW_k) = OP_I u(OW_I) = OP_I$$

where OW_k are the average win odds of category k. Rearranging we get

$$u(OW_k) = \frac{OP_I}{OP_i}$$

From here, Ali (1977) fits a power function to his data of the form

$$\frac{OP_K}{OP_k} = \alpha OW_k^{\beta}$$

where the size of β indicates the bettors preference. β<1 implies risk averse behaviour, β>1 risk loving behaviour and β=1 risk neutrality. The above estimation is conducted using the New Zealand data and the hypothesis that bettors are risk neutral is tested by

$$H_0 : \beta = 1$$
$$H_A : \beta \neq 1$$

Golec and Tamarkin (1998) approximated $u(OW_k)$ by a Taylor Series truncated to three terms and took expectations to get

$$\frac{OP_K}{OP_k} = \gamma_0 + \eta_1 OW_k + \eta_2 OW_k^2 + \eta_3 OW_k^3$$

This equation is estimated and tests are performed on the coefficients. η_1 is the mean coefficient, η_2 the variance coefficient and η_3 the skewness coefficient. The hypothesis that bettors are risk averse implies they prefer mean and skewness but dislike variance. Hence, if bettors are in fact risk averse, one would expect $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$. To test this hypothesis, three separate t tests are performed on each of the coefficients being

$$H_0 : b_1 = 0$$
$$H_A : b_1 > 0$$

$$H_0 : b_2 = 0$$
$$H_A : b_2 < 0$$

$$H_0 : b_3 = 0$$
$$H_A : b_3 > 0$$

Secondly, to establish that collectively the coefficients are different, a joint F statistic is estimated to test the hypothesis

$$H_0 : b_1 = b_2 = b_3 = 0$$
$$H_A : b_1 = b_2 = b_3 \neq 0$$

All of the tests above are linked to each other, and all aim to test whether the New Zealand horse racing betting market is efficient. There are several other tests that can be applied to racing data to test for efficiency, however time constraints have resulted in only these four being implemented in this thesis. The next chapter discloses the data set.

4. Data Set

This study uses the data from Turnball (1997) with the inclusion of an extra 4 months. Turnball's data covered all gallops in New Zealand for the period 01-08-1994 to 22-06-1997. The data set in this thesis covers the period 01-08-1994 to 30-10-1997 which involves 118,101 horses running in 10,361 races at 9,484 race meetings throughout New Zealand. The size of this data set is comparative with the previous studies of:

Busche & Hall (1988)	2,653 races.
Snyder (1978)	849 races.
Ali (1977)	20,247 races.
Ali (1979)	1,089 races.
Hausch et al (1981)	1,065 races.
Asch et al (1982)	729 races.
Fabricant (1965)	10,000 races.
Griffith (1949)	1,124 races.
McGlothin (1956)	9,248 races.
Seligman (1975)	1,183 races.
Weitzman (1965)	12,000 races.

Due to the vast differences in the ratio of racetrack numbers to population for betting countries, the average amount wagered per race varies considerably (see Busche & Hall (1988) for an illustration). In Hong Kong, only two racing tracks exist, which resulted in the average value wagered per race being $US 5.6 million during the 1985-86 season. Compare this with the United States of America, where the ratio of racetrack numbers to population is considerably lower, therefore the average money wagered per race is $US 152,000. New Zealand yields one of the lowest racetrack to population ratios in the world, having 58 racetracks for a population of 3.6 million. In 1996, the T.A.B.'s turnover from 3,172 gallop races was $NZ 644,700,000. Hence, the average total amount wagered per race was $NZ 203,247.

The number of meetings each year at New Zealand racetracks varies considerably. A small town on the west coast of New Zealand by the name of Kumara can only attract enough entries to hold one race meeting per year (typically early in January) while Ellerslie, New Zealand's most popular racetrack, holds approximately 25 race meetings per year.

5. Research Results

5.1 Weak Form Efficiency

Table 5.1 presents the average return results for horses grouped by their rank of favouritism. Column (1) is the ranking of horses. Column (2) gives the average win odds for each group. N_i in column (3) is the number of horses in each category. Column (4) is the total money wagered, given that $1 is placed on each horse in the category. Column (5) gives the total payout from all $1 bets computed as the total closing win prices of all winning horses. Finally, column (6) is the average return which is computed as column (5) divided by column (4).

Weak form efficiency is achieved if, irrespective of the strategy adopted, the average return is (1-t), where t is the track take including breakage. Given that the average track take is computed as 16.9458%, weak form efficiency is achieved when the average return is 0.830542. The closest average return to 0.830542 was with the 8th ranked favourite which yielded an average return of 0.8348. Overall, one could yield a return greater than one minus the track take by backing favourites 3, 4, 5, 7 and 8, while one could do no better than the weak form efficiency return by backing all other favourites. The best strategy to adopt is to back the 4th favourite horse. At a payoff of 0.8737, you would still not make any money, but at least you would minimise your losses.

Figure 5.1 presents a graph of average returns versus the favourite ranking. Included in this graph, and all subsequent graphs on average returns, is a straight line at 0.830542, indicating the weak form efficiency line. For the market to be perfectly weak form efficient, we would expect the data to fall upon this line. For favourites 1-8, there seems to be little deviation from this line. Beyond the 8th favourite, the expected return falls below the efficiency line and continually falls as the rank of favouritism rises. Apart from some minor isolation, there tends to be a negative relationship between the average return and the favourite ranking. This indicates that as horses of less favouritism are backed, the average return falls.

Table 5.2 presents average return results when horses are grouped by even odds categories. Each column is the same as those in Table 5.1 except column (1) is no longer the rank of favouritism, but subjective probability categories. The first 3

Table 5.1: Expected Return on $1 Bet by Rank of Favouritism

Favourite Ranking (1)	Average Odds (2)	N_k (3)	Total Wagered (4)	Total Payout (5)	Average Return (6)
1	$3.19	10361	$10361.00	$8334.75	$0.8044
2	$5.15	10361	$10361.00	$8496.70	$0.8201
3	$6.77	10361	$10361.00	$8810.95	$0.8504
4	$8.60	10359	$10359.00	$9050.25	$0.8737
5	$10.82	10333	$10333.00	$9020.35	$0.8730
6	$13.56	10246	$10246.00	$8200.50	$0.8004
7	$17.15	10006	$10006.00	$8676.00	$0.8671
8	$21.33	9448	$9448.00	$7887.55	$0.8348
9	$26.39	8585	$8585.00	$6472.40	$0.7539
10	$31.86	7394	$7394.00	$5011.60	$0.6778
11	$37.81	6151	$6151.00	$3810.60	$0.6195
12	$43.94	4920	$4920.00	$3675.60	$0.7471
13	$50.59	3767	$3767.00	$1995.20	$0.5297
14	$58.80	2809	$2809.00	$2105.25	$0.7495
15	$64.84	1601	$1601.00	$940.90	$0.5877
16	$72.58	942	$942.00	$646.70	$0.6865
17	$75.61	273	$273.00	$139.30	$0.5103
18	$87.06	184	$184.00	$0.00	$0.0000

Table 5.2: Expected Return on $1 Bet by Even Subjective Probability Categories

Lower-Upper	Average Odds	N_k	Total Wagered	Total Payout	Average Return
(1)	(2)	(3)	(4)	(5)	(6)
0-0.0125	$84.70	4073	$4,073.00	$1,765.30	$0.4334
0.0125-0.025	$45.35	16765	$16,765.00	$10,573.80	$0.6307
0.025-0.0375	$26.89	15351	$15,351.00	$11,159.75	$0.7270
0.0375-0.05	$19.07	12647	$12,647.00	$11,536.90	$0.9122
0.05-0.0625	$14.79	10794	$10,794.00	$8,906.90	$0.8252
0.0625-0.075	$12.08	8825	$8,825.00	$7,741.75	$0.8773
0.075-0.0875	$10.20	7464	$7,464.00	$6,304.80	$0.8447
0.0875-0.1	$8.84	6366	$6,366.00	$5,670.35	$0.8907
0.1-0.1125	$7.79	5306	$5,306.00	$4,474.10	$0.8432
0.1125-0.125	$6.97	4508	$4,508.00	$3,577.30	$0.7935
0.125-0.1375	$6.31	3709	$3,709.00	$3,019.10	$0.8140
0.1375-0.15	$5.75	3127	$3,127.00	$2,675.45	$0.8556
0.15-0.1625	$5.30	2519	$2,519.00	$2,149.05	$0.8531
0.1625-0.175	$4.90	2186	$2,186.00	$1,813.60	$0.8296
0.175-0.1875	$4.56	1814	$1,814.00	$1,511.65	$0.8333
0.1875-0.2	$4.26	1642	$1,642.00	$1,383.65	$0.8427
0.2-0.225	$3.89	2630	$2,630.00	$2,168.60	$0.8246
0.225-0.25	$3.48	2008	$2,008.00	$1,639.45	$0.8165
0.25-0.275	$3.15	1613	$1,613.00	$1,306.45	$0.8100
0.275-0.3	$2.87	1326	$1,326.00	$1,119.80	$0.8445
0.3-0.325	$2.63	912	$912.00	$691.55	$0.7583
0.325-0.35	$2.45	689	$689.00	$521.90	$0.7575
0.35-0.375	$2.28	499	$499.00	$427.60	$0.8569
0.375-0.4	$2.13	370	$370.00	$316.05	$0.8542
0.4-0.45	$1.95	516	$516.00	$418.00	$0.8101
0.45-0.5	$1.75	259	$259.00	$229.65	$0.8867
0.5-0.55	$1.58	114	$114.00	$103.60	$0.9088
0.55-0.6	$1.44	47	$47.00	$47.15	$1.0032
0.6-1	$1.26	22	$22.00	$21.35	$0.9705

categories from 0-0.0375 yield the lowest returns, the best being 0.7270 for the subjective odds interval 0.025-0.0375. For the 26 other categories from 0.0375-1, there is very little deviation from the weak form efficiency line. The most notable deviation being the interval 0.3-0.35 where the average return is approximately 0.75 and for the interval 0.55-0.6 where the average return is 1.0032. Of the 29 groupings, 16 yielded returns greater than one minus the track take while 13 yielded returns below one minus the track take.

Figure 5.2 presents a graph of average returns versus average odds when horses are grouped by even subjective probability categories. Looking at the graph, initially it could be easy to conclude that overall there is a downward sloping nature. However, this conclusion, upon further investigation, could be false. The nature of the grouping strategy means that horses with short odds are over-sampled at the detriment of horses with long odds[17]. It may be possible that there are subgroups within the last 3 categories that yield odds greater than one minus the track take and hence would give an entirely different slant to the conclusions drawn merely by the graph. Part of the reason for adopting the next two methods of grouping is to eliminate this over-sampling (under-sampling) of favourites (longshots).

Table 5.3 presents average return data when horses are grouped by subjective probability categories but such that there are 25 even sized groups. The first half of the groupings from 0-0.071462 vary wildly from the weak form efficiency return. For subjective odds categories 0.071462-1, the deviations are small, the biggest being 0.8821 for the interval 0.095918-0.10624. Overall, 13 average returns fall below 0.830542 and 12 are above. The highest return for this grouping is 0.9591 for the interval 0.039013-0.04345, where the average win odds are \$20.03. The lowest return is 0.4277 which is for the continued backing of horses within the range 0-0.01311, where average odds are \$81.92.

Figure 5.3 is a graph of average returns versus average odds for the 25 even sized groupings. The overall impression from this graph is of a downward sloping function. This trade-off between returns and odds is more conclusive for this grouping than in the previous grouping, since the sampling of long and short odds is more evenly spread.

[17] Note that this grouping method produces 26 estimates of average returns, that have average odds less than \$20.00, and 3 estimates of average returns, that have average odds greater than \$20.00.

Figure 5.1: Average Return versus Rank of Favouritism

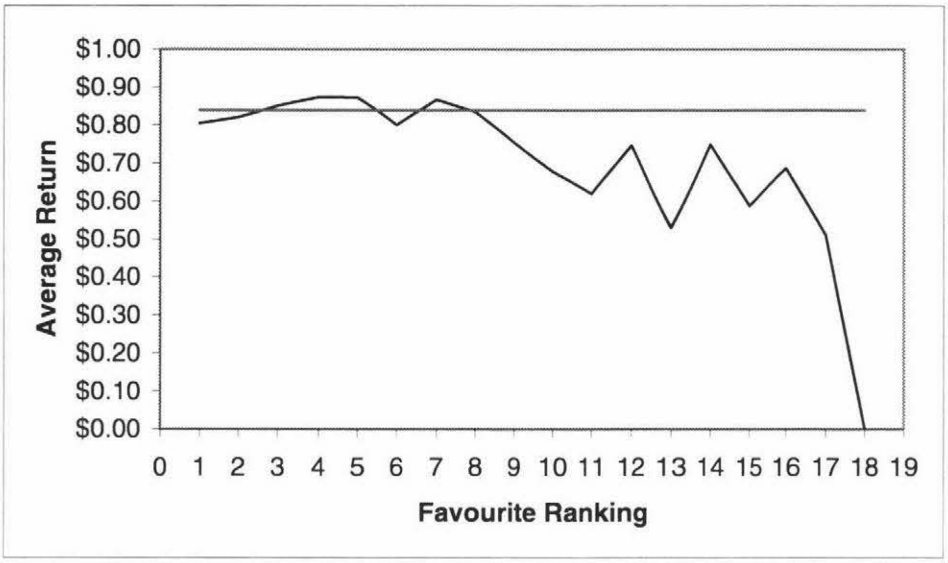


Figure 5.2: Average Return versus Average Odds for Even Subjective Probability Categories

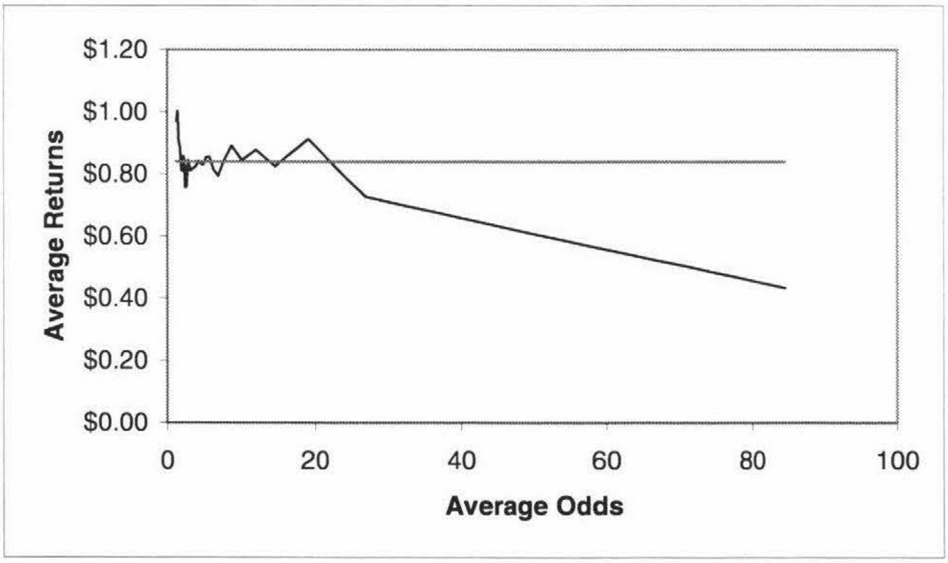


Table 5.3: Expected Return on $1 Bet by 25 Even Sized Subjective Probability Categories

Lower-Upper	Average Odds	N_k	Total Wagered	Total Payout	Average Return
(1)	(2)	(3)	(4)	(5)	(6)
0-0.01311	$81.92	4725	$4,725.00	$2,020.65	$0.4277
0.013111-0.016763	$55.30	4724	$4,724.00	$2,216.90	$0.4693
0.016764-0.020143	$44.82	4724	$4,724.00	$3,562.30	$0.7541
0.020144-0.023511	$37.91	4724	$4,724.00	$3,383.65	$0.7163
0.023512-0.027174	$32.64	4724	$4,724.00	$3,239.30	$0.6857
0.027175-0.030911	$28.49	4724	$4,724.00	$2,940.15	$0.6224
0.030912-0.034839	$25.15	4724	$4,724.00	$3,621.65	$0.7666
0.034840-0.039012	$22.41	4724	$4,724.00	$4,045.55	$0.8564
0.039013-0.04345	$20.03	4724	$4,724.00	$4,531.00	$0.9591
0.043451-0.048241	$18.02	4724	$4,724.00	$4,227.95	$0.8950
0.048242-0.053391	$16.26	4724	$4,724.00	$3,666.00	$0.7760
0.053392-0.058838	$14.72	4724	$4,724.00	$4,062.00	$0.8599
0.058839-0.064795	$13.38	4724	$4,724.00	$3,771.45	$0.7984
0.064796-0.071462	$12.14	4724	$4,724.00	$4,308.15	$0.9120
0.071463-0.078806	$11.00	4724	$4,724.00	$4,094.05	$0.8666
0.078807-0.086929	$9.97	4724	$4,724.00	$4,100.65	$0.8680
0.086930-0.095917	$9.05	4724	$4,724.00	$4,014.15	$0.8497
0.095918-0.10624	$8.19	4724	$4,724.00	$4,166.95	$0.8821
0.106241-0.118228	$7.37	4724	$4,724.00	$3,820.15	$0.8087
0.118229-0.132756	$6.60	4724	$4,724.00	$3,831.60	$0.8111
0.132757-0.151227	$5.84	4724	$4,724.00	$4,037.70	$0.8547
0.151228-0.176891	$5.07	4724	$4,724.00	$3,941.10	$0.8343
0.176892-0.215018	$4.26	4724	$4,724.00	$3,930.15	$0.8320
0.215019-0.275531	$3.42	4724	$4,724.00	$3,863.50	$0.8178
0.275532-1	$2.42	4724	$4,724.00	$3,877.90	$0.8209

Table 5.4: Expected Return on $1 Bet by 50 Even Sized Subjective Probability Categories

Lower-Upper	Average Odds	N_k	Total Wagered	Total Payout	Average Return
(1)	(2)	(3)	(4)	(5)	(6)
0-0.010795	$94.69	2363	$2,363.00	$983.00	$0.4160
0.010795-0.01311	$69.16	2362	$2,362.00	$1,037.65	$0.4393
0.01311-0.01503	$58.69	2362	$2,362.00	$1,097.80	$0.4648
0.01503-0.016763	$51.91	2362	$2,362.00	$1,119.10	$0.4738
0.016763-0.018482	$46.90	2362	$2,362.00	$1,586.45	$0.6717
0.018482-0.020143	$42.74	2362	$2,362.00	$1,975.85	$0.8365
0.020143-0.021798	$39.38	2362	$2,362.00	$1,781.30	$0.7541
0.021798-0.023511	$36.44	2362	$2,362.00	$1,602.35	$0.6784
0.023511-0.025332	$33.83	2362	$2,362.00	$1,451.95	$0.6147
0.025332-0.027174	$31.46	2362	$2,362.00	$1,787.35	$0.7567
0.027174-0.028985	$29.42	2362	$2,362.00	$1,689.55	$0.7153
0.028985-0.030911	$27.57	2362	$2,362.00	$1,241.60	$0.5257
0.030911-0.032848	$25.94	2362	$2,362.00	$1,965.15	$0.8320
0.032848-0.034839	$24.36	2362	$2,362.00	$1,656.50	$0.7013
0.034839-0.036845	$23.03	2362	$2,362.00	$1,682.25	$0.7122
0.036845-0.039012	$21.78	2362	$2,362.00	$2,363.30	$1.0006
0.039012-0.041227	$20.56	2362	$2,362.00	$2,215.25	$0.9379
0.041227-0.04345	$19.50	2362	$2,362.00	$2,315.75	$0.9804
0.04345-0.045876	$18.48	2362	$2,362.00	$2,082.10	$0.8815
0.045876-0.048241	$17.56	2362	$2,362.00	$2,145.85	$0.9085
0.048241-0.050845	$16.67	2362	$2,362.00	$1,833.50	$0.7762
0.050845-0.053391	$15.84	2362	$2,362.00	$1,832.50	$0.7758
0.053391-0.056033	$15.08	2362	$2,362.00	$2,064.35	$0.8740
0.056033-0.058838	$14.37	2362	$2,362.00	$1,997.65	$0.8457
0.058838-0.061699	$13.70	2362	$2,362.00	$2,027.70	$0.8585
0.061699-0.064795	$13.06	2362	$2,362.00	$1,743.75	$0.7383
0.064795-0.068008	$12.43	2362	$2,362.00	$2,126.75	$0.9004
0.068008-0.071462	$11.85	2362	$2,362.00	$2,181.40	$0.9235
0.071462-0.075083	$11.27	2362	$2,362.00	$2,120.75	$0.8979
0.075083-0.078806	$10.73	2362	$2,362.00	$1,973.30	$0.8354
0.078806-0.082783	$10.22	2362	$2,362.00	$2,024.80	$0.8572
0.082783-0.086929	$9.73	2362	$2,362.00	$2,075.85	$0.8789

0.086929-0.091264	$9.28	2362	$2,362.00	$1,986.65	$0.8411
0.091264-0.095917	$8.83	2362	$2,362.00	$2,027.50	$0.8584
0.095917-0.100908	$8.40	2362	$2,362.00	$2,075.50	$0.8787
0.100908-0.10624	$7.98	2362	$2,362.00	$2,091.45	$0.8855
0.10624-0.111957	$7.58	2362	$2,362.00	$2,073.45	$0.8778
0.111957-0.118228	$7.17	2362	$2,362.00	$1,746.70	$0.7395
0.118228-0.12505	$6.80	2362	$2,362.00	$1,951.80	$0.8263
0.12505-0.132756	$6.41	2362	$2,362.00	$1,879.80	$0.7959
0.132756-0.141188	$6.04	2362	$2,362.00	$1,943.75	$0.8229
0.141188-0.151227	$5.64	2362	$2,362.00	$2,093.95	$0.8865
0.151227-0.162937	$5.27	2362	$2,362.00	$1,990.05	$0.8425
0.162937-0.176891	$4.87	2362	$2,362.00	$1,951.05	$0.8260
0.176891-0.194337	$4.47	2362	$2,362.00	$2,013.90	$0.8526
0.194337-0.215018	$4.05	2362	$2,362.00	$1,916.25	$0.8113
0.215018-0.240894	$3.64	2362	$2,362.00	$1,982.10	$0.8392
0.240894-0.275532	$3.21	2362	$2,362.00	$1,881.40	$0.7965
0.275532-0.329301	$2.75	2362	$2,362.00	$1,893.10	$0.8015
0.329301-1	$2.10	2362	$2,362.00	$1,984.80	$0.8403

Figure 5.3: Average Return versus Average Odds for 25 Even Sized Subjective Probability Categories

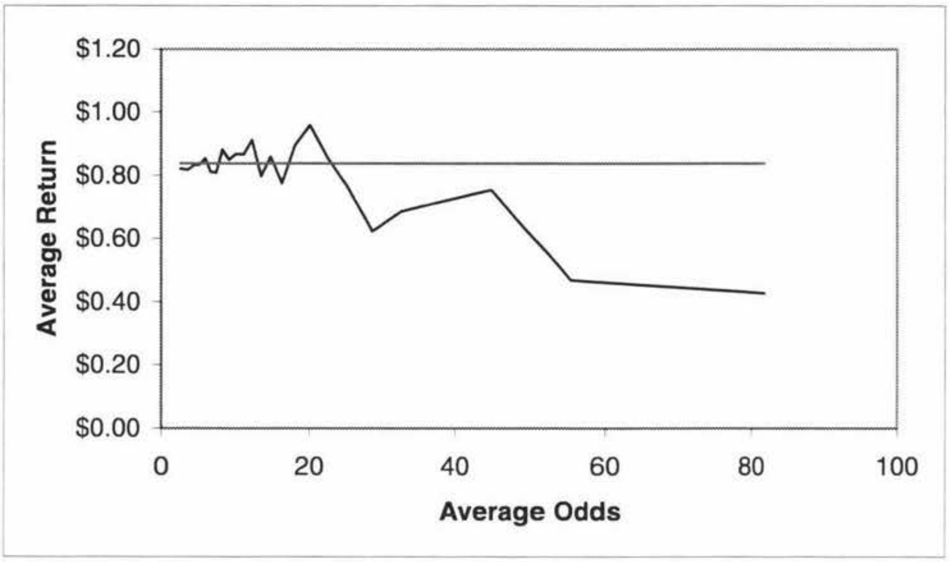
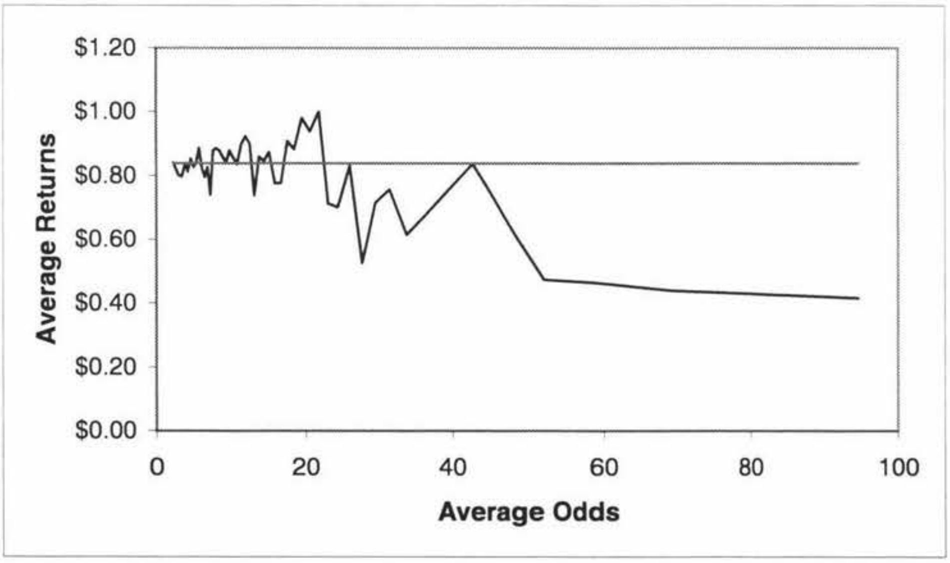


Figure 5.4: Average Returns versus Average Odds for 50 Even Sized Subjective Probability Categories



The conclusion that at low odds, the deviations are small and at long odds the deviations are large, is emphasised by this graph. Table 5.4 presents average return figures for horses grouped by 50 even sized subjective probability categories. The first point to note is that 13 of the first 15 categories from 0-0.036845 (which are for average odds $94.69 to $23.03) yield average returns lower than the weak form efficiency return with the two exceptions being slightly over 0.830542. The next 23 categories, from 0.036845-0.118228, which could be regarded as the middle group of categories, vary quite a lot from the weak form efficiency line but are neither predominately above or below the line, whereas the final 12 categories from 0.118228-1 vary little from the efficiency line. Once again, a strategy is found that returns a profit. For the subjective probability interval 0.036845-0.039012 a $1 bet on all 2362 horses would have yielded $2363.30, a mere 0.06% return, but nonetheless a positive return. The lowest return for this type of grouping was once again the group of biggest longshots at $0.4160.

The graphical representation of this table is presented in Figure 5.4. What is immediately obvious from this graph is the trade-off between average returns and average odds. By doubling the number of groupings to 50, the trade-off between average returns and average odds is even more conclusive.

5.2 Favourite Longshot Bias using Opening Odds

Tables 5.5-5.8 present the results of the favourite longshot bias analysis using the 4 different methods of grouping horses with opening odds. Column (1) in Table 5.5 is the ranking of favouritism (note that favourites 17 and 18 are summed to form a single category). Column (1) in Tables 5.6-5.8 are the lower and upper subjective probabilities, with favouritism increasing in descending order. Columns (2)-(7) are the same for each of the Tables 5.5-5.8. Column (2) shows the average odds for each category. Column (3) is the number of horses in each category. Column (4) is the number of winners in each category. Column (5) gives the objective probability, calculated as column (4) divided by column (3). Column (6) is the average subjective probability for each category. Column (7) is the z value which measures the extent to which the objective probability varies from the subjective probability.

For the favourite longshot bias to exist, we expect that for longshots, the objective probability to be less than the subjective probability and hence the z value will be negative. For favourites, we expect the objective probability to be less than the subjective probability, and to hence have positive z values. It is the size of the z values that determines the significance of the deviations.

Table 5.5 presents the favourite longshot bias analysis for the rank of favouritism grouping method. The first thing to note here is that favourites 1-4 all return positive z values indicating underbetting and favourites 5-18 all return negative z values which indicates overbetting. Of the first four favourites, the 1st and 4th are significant at the 5 percent level and the 2nd is significant at the 1 percent level. Of the categories 5-18, favourites 5 to 8 return insignificant z values, 9 to 13 return significant z values at the 1 percent level, yet the final group from 14 to 18, while showing evidence of overbetting, are not significant. The biggest deviation between the objective and subjective probabilities occurred for the 10th favourite. For this category, the subjective probability was 3.57% yet the horses true win percentage was a mere 2.41%. Despite the significant z values for longshots falling in the middle categories instead of the final categories, there is enough evidence to suggest that the favourite longshot bias exists at opening odds using this method of grouping.

Table 5.6 shows the favourite longshot bias analysis when grouping by even subjective probability classes. Given that the first four categories have z values which are significant well beyond the 1 percent level of significance, it is reasonable to conclude that, using this method at least, longshots are being overbet. Beyond the first four categories, the results are somewhat mixed. For the middle group of categories, there are more positive z values than negative z values, indicating that horses within this range are being slightly underbet. Contrary to the hypothesis, favourite horses, when grouped this way, show no evidence of being underbet. In fact for the horses that attract 50-55% of the total win pool, the z value was significantly negative at the 10 percent level, even though the average odds of $1.58 indicate they are heavy favourites.

Table 5.7 presents the favourite longshot bias analysis using 25 even sized subjective probability groups. This method of grouping indicates that there is overwhelming support for the favourite longshot bias. The first nine categories all return negative z values, 6 of which are significant at least at the 5 percent level. The final 16 categories all return positive z values, 12 of which are significant at least at the 5 percent level. It is interesting to note that if it is true that all favourites are underbet and all longshots are

Table 5.5: Favourite Longshot Bias Analysis by Rank of Favouritism using Opening Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.87	10361	2775	0.26783	0.256822	2.53060**
2	$6.00	10361	1775	0.17132	0.157299	3.78659***
3	$7.52	10361	1306	0.12605	0.12225	1.16527
4	$9.04	10359	1115	0.10764	0.099949	2.52440**
5	$10.79	10333	829	0.08023	0.083183	-1.10563
6	$12.81	10246	711	0.06939	0.069715	-0.12829
7	$15.41	10006	573	0.05727	0.058421	-0.49740
8	$18.28	9448	444	0.04699	0.049317	-1.06693
9	$21.96	8585	296	0.03448	0.04178	-3.70776***
10	$25.67	7394	178	0.02407	0.035786	-6.57064***
11	$29.79	6151	139	0.02260	0.031001	-4.43444***
12	$34.15	4920	92	0.01870	0.027026	-4.31170***
13	$39.00	3767	59	0.01566	0.023625	-3.93601***
14	$45.06	2809	50	0.01780	0.020382	-1.03499
15	$49.87	1601	29	0.01811	0.018204	-0.02710
16	$55.04	942	14	0.01486	0.016149	-0.32645
17/18	$63.49	457	5	0.01094	0.014016	-0.63194

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.6: Favourite Longshot Bias Analysis by Even Subjective Probability Categories using Opening Win Odds

Lower-Upper (1)	Average Odds (2)	N_k (3)	W_k (4)	OP_k (5)	SP_k (6)	z_k (7)
0-0.0125	$86.15	1836	8	0.00436	0.01005	-3.70325***
0.0125-0.025	$43.97	11370	155	0.01363	0.01949	-5.38588***
0.025-0.0375	$26.78	14927	356	0.02385	0.03133	-5.98992***
0.0375-0.05	$19.10	14401	497	0.03451	0.04366	-6.01570***
0.05-0.0625	$14.83	12744	689	0.05406	0.05610	-1.01441
0.0625-0.075	$12.12	10784	781	0.07242	0.068534	1.55774
0.075-0.0875	$10.23	9145	774	0.08464	0.08113	1.20301
0.0875-0.1	$8.86	7436	668	0.08983	0.09364	-1.14903
0.1-0.1125	$7.81	6227	700	0.11241	0.10612	1.57204
0.1125-0.125	$7.00	5058	623	0.12317	0.11845	1.02175
0.125-0.1375	$6.33	3992	499	0.12500	0.13095	-1.13724
0.1375-0.15	$5.78	3356	525	0.15644	0.14363	2.04178**
0.15-0.1625	$5.31	2541	427	0.16804	0.15622	1.59421
0.1625-0.175	$4.93	2334	425	0.18209	0.16841	1.71312*
0.175-0.1875	$4.58	1770	342	0.19322	0.18110	1.29186
0.1875-0.2	$4.27	1588	339	0.21348	0.19364	1.92927*
0.2-0.225	$3.91	2363	598	0.25307	0.21221	4.56874***
0.225-0.25	$3.50	1569	388	0.24729	0.23727	0.92014
0.25-0.275	$3.15	1228	342	0.27850	0.26246	1.25419
0.275-0.3	$2.88	877	249	0.28392	0.28771	-0.24894
0.3-0.325	$2.65	593	184	0.31029	0.31329	-0.15800
0.325-0.35	$2.46	484	160	0.33058	0.33737	-0.31767
0.35-0.375	$2.28	353	140	0.39660	0.36222	1.32046
0.375-0.4	$2.14	268	97	0.36194	0.38713	-0.85809
0.4-0.45	$1.95	365	161	0.44110	0.42323	0.68732
0.45-0.5	$1.75	195	96	0.49231	0.47359	0.52268
0.5-0.55	$1.58	124	55	0.44355	0.52202	-1.75893*
0.55-0.6	$1.45	78	43	0.55128	0.57094	-0.34909
0.6-1	$1.23	95	69	0.72632	0.67722	1.07321

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.7: Favourite Longshot Bias Analysis by 25 Even Sized Subjective Probability Categories using Opening Win Odds

Lower-Upper	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-0.016754	$67.99	4725	36	0.00762	0.01293	-4.19885***
0.016754-0.021614	$43.07	4274	64	0.01497	0.01931	-2.33505**
0.021614-0.025826	$34.91	4274	77	0.01802	0.02377	-2.82894***
0.025826-0.029891	$29.71	4274	97	0.02270	0.02790	-2.28499**
0.029891-0.033777	$26.04	4274	124	0.02901	0.03184	-1.10140
0.033777-0.037677	$23.18	4274	123	0.02878	0.03574	-2.72213***
0.037677-0.041647	$20.90	4274	128	0.02995	0.03965	-3.72096***
0.041647-0.045719	$18.97	4274	173	0.04048	0.04369	-1.06531
0.045719-0.04999	$17.30	4274	193	0.04516	0.04789	-0.86204
0.04999-0.054378	$15.87	4274	225	0.05264	0.05218	0.13608
0.054378-0.05909	$14.63	4274	277	0.06481	0.05671	2.15188**
0.05909-0.063975	$13.48	4274	284	0.06645	0.06150	1.29786
0.063975-0.069294	$12.45	4274	319	0.07464	0.06657	2.00589**
0.069294-0.075138	$11.48	4274	368	0.08610	0.07219	3.24304***
0.075138-0.081437	$10.59	4274	387	0.09055	0.07820	2.81325***
0.081437-0.088176	$9.79	4274	419	0.09803	0.08470	2.93165***
0.088176-0.095942	$9.01	4274	398	0.09312	0.09197	0.25919
0.095942-0.104701	$8.27	4274	489	0.11441	0.10017	2.92526***
0.104701-0.114484	$7.57	4274	540	0.12635	0.10942	3.33006***
0.114484-0.126345	$6.89	4274	604	0.14132	0.12023	3.95721***
0.126345-0.142006	$6.19	4274	608	0.14226	0.13395	1.55415
0.142006-0.162878	$5.47	4274	798	0.18671	0.15192	5.83698***
0.162878-0.192736	$4.70	4274	887	0.20753	0.17665	4.97946***
0.192736-0.248828	$3.84	4274	1155	0.27024	0.21722	7.80588***
0.248828-1	$2.59	4274	1617	0.37833	0.33692	5.58279***

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.8: Favourite Longshot Bias Analysis by 50 Even Sized Subjective Probability Categories using Opening Win Odds

Lower-Upper (1)	Average Odds (2)	N_k (3)	W_k (4)	OP_k (5)	SP_k (6)	z_k (7)
0-0.013394	$81.20	2363	12	0.00508	0.01070	-3.84196***
0.013394-0.016754	$54.78	2362	24	0.01016	0.01517	-2.42548**
0.016754-0.01939	$45.80	2362	29	0.01228	0.01811	-2.57376**
0.01939-0.021614	$40.34	2362	35	0.01482	0.02051	-2.29147**
0.021614-0.023815	$36.47	2362	35	0.01482	0.02271	-3.17430***
0.023815-0.025826	$33.35	2362	42	0.01778	0.02483	-2.59334***
0.025826-0.027908	$30.79	2362	46	0.01948	0.02688	-2.60607***
0.027908-0.029891	$28.64	2362	51	0.02159	0.02892	-2.44896**
0.029891-0.031847	$26.82	2362	68	0.02879	0.03086	-0.60234
0.031847-0.033777	$25.26	2362	56	0.02371	0.03282	-2.90989***
0.033777-0.035756	$23.83	2362	64	0.02710	0.03475	-2.29211**
0.035756-0.037677	$22.53	2362	59	0.02498	0.03673	-3.65828***
0.037677-0.039649	$21.42	2362	49	0.02075	0.03864	-6.10121***
0.039649-0.041647	$20.38	2362	79	0.03345	0.04066	-1.95023*
0.041647-0.043723	$19.40	2362	82	0.03472	0.04268	-2.11323**
0.043723-0.045719	$18.53	2362	91	0.03853	0.04470	-1.55930
0.045719-0.04792	$17.68	2362	92	0.03895	0.04684	-1.98078**
0.04792-0.04999	$16.92	2362	101	0.04276	0.04896	-1.48808
0.04999-0.052186	$16.20	2362	105	0.04445	0.05110	-1.56777
0.052186-0.054378	$15.54	2362	120	0.05080	0.05326	-0.54283
0.054378-0.056686	$14.93	2362	123	0.05207	0.05553	-0.75695
0.056686-0.05909	$14.33	2362	154	0.06520	0.05788	1.44092
0.05909-0.061499	$13.74	2362	138	0.05843	0.06029	-0.38640
0.061499-0.063975	$13.21	2362	146	0.06181	0.06272	-0.18311
0.063975-0.066558	$12.70	2362	168	0.07113	0.06528	1.10608
0.066558-0.069294	$12.20	2362	151	0.06393	0.06787	-0.78379
0.069294-0.072212	$11.72	2362	174	0.07367	0.07072	0.54858
0.072212-0.075138	$11.24	2362	194	0.08213	0.07365	1.50099
0.075138-0.07814	$10.80	2362	188	0.07959	0.07661	0.53583
0.07814-0.081437	$10.38	2362	199	0.08425	0.07979	0.78077
0.081437-0.084693	$9.98	2362	194	0.08213	0.08302	-0.15661
0.084693-0.088176	$9.59	2362	225	0.09526	0.08638	1.46930

0.088176-0.091907	$9.20	2362	199	0.08425	0.09002	-1.00956
0.091907-0.095942	$8.82	2362	199	0.08425	0.09392	-1.69202*
0.095942-0.100064	$8.45	2362	237	0.10034	0.09799	0.37932
0.100064-0.104701	$8.09	2362	252	0.10669	0.10235	0.68376
0.104701-0.109343	$7.74	2362	261	0.11050	0.10696	0.54799
0.109343-0.114484	$7.41	2362	279	0.11812	0.11188	0.93958
0.114484-0.119978	$7.07	2362	317	0.13421	0.11719	2.42577**
0.119978-0.126345	$6.72	2362	287	0.12151	0.12327	-0.26274
0.126345-0.133768	$6.37	2362	295	0.12489	0.13005	-0.75845
0.133768-0.142006	$6.01	2362	313	0.13251	0.13785	-0.76422
0.142006-0.151341	$5.66	2362	390	0.16511	0.14667	2.41380**
0.151341-0.162878	$5.28	2362	408	0.17273	0.15717	2.00162**
0.162878-0.175907	$4.91	2362	426	0.18036	0.16909	1.42350
0.175907-0.192736	$4.50	2362	461	0.19517	0.18419	1.34648
0.192736-0.215233	$4.07	2362	584	0.24725	0.20370	4.90602***
0.215233-0.248828	$3.60	2362	571	0.24174	0.23072	1.25143
0.248828-0.310036	$3.02	2362	675	0.28577	0.27498	1.16173
0.310036-1	$2.16	2362	942	0.39881	0.39884	-0.00232

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

overbet then, according to this method, the odds at which horses change from being considered favourites to longshots is at approximately $16.00.

The final favourite longshot bias analysis using opening odds is for 50 even sized subjective probability categories and is presented in Table 5.8. With this method it is apparent that longshots are being overbet, since the first 21 categories all return negative z values, 15 of which are significant at least at the 10 percent level. The next 15 categories provide a sprinkling of both positive and negative z values, of which only 1 is significant. The final 16 categories which can all be considered favourites have 12 positive values of which only 4 are significant. Interestingly, despite this apparent support for the favourite longshot bias, the category that contains the biggest favourites (subjective probability interval 0.310036-1) is actually being overbet, although the difference in the objective and subjective probabilities is not statistically significant.

5.3 Favourite Longshot Bias using Closing Odds

Tables 5.9-5.12 present the favourite longshot bias analysis for the four different grouping methods, using closing odds. Columns (1)-(7) are the same as columns (1)-(7) for Tables 5.5-5.8, which have been previously discussed.

Table 5.9 presents the results when grouping horses according to their rank of favouritism. Of the 17 categories specified, only 4 produced z values that were significant at least at the 10 percent level. These were for favourites 4, 10, 11 and 13 (4 is positive, 10, 11 and 13 are negative). Apart from the 1st ranking favourite returning a negative z value, the typical pattern is that favourites had positive values and longshots had negative values, as we would expect for the favourite longshot bias. However, because less than 25% of the categories had significant z values, it is impossible to draw any concrete conclusion as to the existence of the favourite longshot bias with this method.

Table 5.10 presents the results when grouping using even subjective probability categories. These results behave in much the same way as the opening odds, using the same grouping method. That is, the first 3 categories provide very strong evidence that these horses are being overbet. The fourth category, 0.0375-0.05, proved to be very interesting since using opening odds, the z value was –6.01570, which indicates overbetting, but using closing odds, the z value became 2.32947, which indicates

Table 5.9: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.19	10361	2852	0.27526	0.27986	-1.04760
2	$5.15	10361	1763	0.17016	0.16882	0.36186
3	$6.77	10361	1349	0.13020	0.12754	0.80484
4	$8.60	10359	1108	0.10696	0.10045	2.14458**
5	$10.82	10333	871	0.08429	0.08030	1.46133
6	$13.56	10246	654	0.06383	0.06490	-0.44350
7	$17.15	10006	549	0.05487	0.05233	1.11304
8	$21.33	9448	417	0.04414	0.04271	0.67443
9	$26.39	8585	296	0.03448	0.03512	-0.32683
10	$31.86	7394	190	0.02570	0.02930	-1.95976*
11	$37.81	6151	120	0.01951	0.02474	-2.96770***
12	$43.94	4920	101	0.02053	0.02123	-0.34525
13	$50.59	3767	47	0.01248	0.01839	-3.26704***
14	$58.80	2809	41	0.01460	0.01584	-0.55137
15	$64.84	1601	18	0.01124	0.01423	-1.13329
16	$72.58	942	11	0.01168	0.01253	-0.24392
17/18	$80.22	457	3	0.00656	0.01137	-1.27195

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.10: Favourite Longshot Bias Analysis by Even Subjective Probability Categories using Closing Win Odds

Lower-Upper (1)	Average Odds (2)	N_k (3)	W_k (4)	OP_k (5)	SP_k (6)	z_k (7)
0-0.0125	$84.70	4073	22	0.00540	0.01010	-4.09024***
0.0125-0.025	$45.35	16765	248	0.01479	0.01888	-4.38475***
0.025-0.0375	$26.89	15351	423	0.02756	0.03111	-2.69390***
0.0375-0.05	$19.07	12647	607	0.04800	0.04357	2.32947**
0.05-0.0625	$14.79	10794	606	0.05614	0.05605	0.04342
0.0625-0.075	$12.08	8825	644	0.07297	0.06853	1.60460
0.075-0.0875	$10.20	7464	620	0.08307	0.08109	0.61685
0.0875-0.1	$8.84	6366	643	0.10101	0.09361	1.95778*
0.1-0.1125	$7.79	5306	575	0.10837	0.10608	0.53707
0.1125-0.125	$6.97	4508	514	0.11402	0.11858	-0.96365
0.125-0.1375	$6.31	3709	479	0.12915	0.13100	-0.33651
0.1375-0.15	$5.75	3127	465	0.14870	0.14363	0.79691
0.15-0.1625	$5.30	2519	406	0.16118	0.15595	0.71378
0.1625-0.175	$4.90	2186	370	0.16926	0.16854	0.08966
0.175-0.1875	$4.56	1814	332	0.18302	0.18099	0.22320
0.1875-0.2	$4.26	1642	324	0.19732	0.19377	0.36169
0.2-0.225	$3.89	2630	559	0.21255	0.21274	-0.02441
0.225-0.25	$3.48	2008	471	0.23456	0.23734	-0.29430
0.25-0.275	$3.15	1613	415	0.25728	0.26277	-0.50374
0.275-0.3	$2.87	1326	391	0.29487	0.28779	0.56580
0.3-0.325	$2.63	912	263	0.28838	0.31332	-1.66246*
0.325-0.35	$2.45	689	214	0.31060	0.33715	-1.50652
0.35-0.375	$2.28	499	188	0.37675	0.36205	0.67783
0.375-0.4	$2.13	370	149	0.40270	0.38698	0.61664
0.4-0.45	$1.95	516	215	0.41667	0.42205	-0.24819
0.45-0.5	$1.75	259	131	0.50579	0.47096	1.12117
0.5-0.55	$1.58	114	66	0.57895	0.52008	1.27292
0.55-0.6	$1.44	47	33	0.70213	0.57153	1.95778*
0.6-1	$1.26	22	17	0.77273	0.64734	1.40334

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.11: Favourite Longshot Bias Analysis by 25 Even Sized Subjective Probability Categories using Closing Win Odds

Lower-Upper	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-0.01311	$81.92	4725	26	0.00550	0.01047	-4.61785***
0.013111-0.016763	$55.30	4724	41	0.00868	0.01500	-4.68171***
0.016764-0.020143	$44.82	4724	80	0.01693	0.01847	-0.81671
0.020144-0.023511	$37.91	4724	89	0.01884	0.02181	-1.50324
0.023512-0.027174	$32.64	4724	100	0.02117	0.02532	-1.98305**
0.027175-0.030911	$28.49	4724	103	0.02180	0.02901	-3.39291***
0.030912-0.034839	$25.15	4724	144	0.03048	0.03286	-0.94972
0.034840-0.039012	$22.41	4724	181	0.03831	0.03687	0.51885
0.039013-0.04345	$20.03	4724	227	0.04805	0.04121	2.19901**
0.043451-0.048241	$18.02	4724	235	0.04975	0.04584	1.23344
0.048242-0.053391	$16.26	4724	226	0.04784	0.05083	-0.96385
0.053392-0.058838	$14.72	4724	276	0.05843	0.05607	0.68934
0.058839-0.064795	$13.38	4724	282	0.05970	0.06176	-0.59997
0.064796-0.071462	$12.14	4724	355	0.07515	0.06807	1.84630*
0.071463-0.078806	$11.00	4724	372	0.07875	0.07510	0.92951
0.078807-0.086929	$9.97	4724	412	0.08721	0.08279	1.07671
0.086930-0.095917	$9.05	4724	444	0.09399	0.09135	0.62171
0.095918-0.10624	$8.19	4724	509	0.10775	0.10093	1.51197
0.106241-0.118228	$7.37	4724	518	0.10965	0.11207	-0.53105
0.118229-0.132756	$6.60	4724	580	0.12278	0.12518	-0.50250
0.132757-0.151227	$5.84	4724	693	0.14670	0.14166	0.97866
0.151228-0.176891	$5.07	4724	779	0.16490	0.16329	0.29822
0.176892-0.215018	$4.26	4724	952	0.20152	0.19455	1.19479
0.215019-0.275531	$3.42	4724	1132	0.23963	0.24231	-0.43211
0.275532-1	$2.42	4724	1661	0.35161	0.35042	0.17105

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.12: Favourite Longshot Bias Analysis by 50 Even Sized Subjective Probability Categories using Closing Win Odds

Lower-Upper	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-0.010795	$94.69	2363	11	0.00466	0.00896	-3.07431***
0.010795-0.01311	$69.16	2362	15	0.00635	0.01199	-3.44735***
0.01311-0.01503	$58.69	2362	19	0.00804	0.01409	-3.28727***
0.01503-0.016763	$51.91	2362	22	0.00931	0.01591	-3.33637***
0.016763-0.018482	$46.90	2362	34	0.01439	0.01761	-1.31274
0.018482-0.020143	$42.74	2362	46	0.01948	0.01932	0.05407
0.020143-0.021798	$39.38	2362	45	0.01905	0.02097	-0.68246
0.021798-0.023511	$36.44	2362	44	0.01863	0.02266	-1.44795
0.023511-0.025332	$33.83	2362	43	0.01820	0.02441	-2.25397
0.025332-0.027174	$31.46	2362	57	0.02413	0.02624	-0.66686
0.027174-0.028985	$29.42	2362	58	0.02456	0.02807	-1.10483
0.028985-0.030911	$27.57	2362	45	0.01905	0.02995	-3.87449***
0.030911-0.032848	$25.94	2362	76	0.03218	0.03186	0.08644
0.032848-0.034839	$24.36	2362	68	0.02879	0.03385	-1.47144
0.034839-0.036845	$23.03	2362	73	0.03091	0.03583	-1.38267
0.036845-0.039012	$21.78	2362	108	0.04572	0.03790	1.82025*
0.039012-0.041227	$20.56	2362	108	0.04572	0.04009	1.30990
0.041227-0.04345	$19.50	2362	119	0.05038	0.04232	1.79003*
0.04345-0.045876	$18.48	2362	113	0.04784	0.04465	0.72659
0.045876-0.048241	$17.56	2362	122	0.05165	0.04704	1.01302
0.048241-0.050845	$16.67	2362	110	0.04657	0.04955	-0.68705
0.050845-0.053391	$15.84	2362	116	0.04911	0.05212	-0.67582
0.053391-0.056033	$15.08	2362	137	0.05800	0.05472	0.68297
0.056033-0.058838	$14.37	2362	139	0.05885	0.05743	0.29347
0.058838-0.061699	$13.70	2362	148	0.06266	0.06026	0.48090
0.061699-0.064795	$13.06	2362	134	0.05673	0.06327	-1.37294
0.064795-0.068008	$12.43	2362	171	0.07240	0.06642	1.12162
0.068008-0.071462	$11.85	2362	184	0.07790	0.06972	1.48399
0.071462-0.075083	$11.27	2362	188	0.07959	0.07327	1.13539
0.075083-0.078806	$10.73	2362	184	0.07790	0.07694	0.17443
0.078806-0.082783	$10.22	2362	199	0.08425	0.08078	0.60673
0.082783-0.086929	$9.73	2362	213	0.09018	0.08481	0.91126

0.086929-0.091264	$9.28	2362	214	0.09060	0.08909	0.25571
0.091264-0.095917	$8.83	2362	230	0.09738	0.09361	0.61800
0.095917-0.100908	$8.40	2362	247	0.10457	0.09835	0.98814
0.100908-0.10624	$7.98	2362	262	0.11092	0.10346	1.15511
0.10624-0.111957	$7.58	2362	274	0.11600	0.10905	1.05517
0.111957-0.118228	$7.17	2362	244	0.10330	0.11508	-1.88041*
0.118228-0.12505	$6.80	2362	287	0.12151	0.12161	-0.01552
0.12505-0.132756	$6.41	2362	293	0.12405	0.12874	-0.69168
0.132756-0.141188	$6.04	2362	322	0.13633	0.13697	-0.09158
0.141188-0.151227	$5.64	2362	371	0.15707	0.14634	1.43265
0.151227-0.162937	$5.27	2362	378	0.16003	0.15685	0.42232
0.162937-0.176891	$4.87	2362	401	0.16977	0.16973	0.00480
0.176891-0.194337	$4.47	2362	452	0.19136	0.18496	0.79143
0.194337-0.215018	$4.05	2362	473	0.20025	0.20415	-0.47347
0.215018-0.240894	$3.64	2362	545	0.23074	0.22700	0.43146
0.240894-0.275532	$3.21	2362	587	0.24852	0.25762	-1.02351
0.275532-0.329301	$2.75	2362	688	0.29128	0.30044	-0.97999
0.329301-1	$2.10	2362	973	0.41194	0.40036	1.14351

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

underbetting and is significant at the 5 percent level. The next 25 categories provide very little evidence for the bias since only 3 categories had significant z values.

Table 5.11 presents the favourite longshot bias results when grouping by 25 even sized subjective probability categories. These results are similar to the previous table. There is very strong evidence that longshots are overbet, since the first 6 intervals all return negative z values, of which 4 are highly significant. The remaining 19 categories only produce two other significant z values, the first being the interval 0.039013-0.04345. The average odds for this interval are $20.03, hence the horses could be considered longshots, yet the z value of 2.19901 indicates that they are significantly underbet. The second interval is 0.064796-0.071462, which has average odds of $12.14. The significant z value of 1.84630 indicates that these favourites are underbet. Since there is a lack of significant z values for the favourite categories, no conclusions can be drawn as to whether overall, these categories are underbet or overbet.

As we would expect, doubling the amount of categories to 50 in Table 5.12 leads to the same inferences as in Table 5.11. This table shows that heavy longshots are extremely overbet as the first four categories return negative z values significant beyond the 1 percent level. These four categories cover horses with average odds from $94.69 to $51.91. Once again, after this point, there are very few significant z values so no conclusions can be made as to whether favourites are underbet or overbet.

5.4 Favourite Longshot Bias throughout the Raceday

To examine whether the favourite longshot bias becomes more prevalent throughout the raceday, two different methods are adopted. Firstly, the data was divided into separate races 1, 2, 3,..., 10, where each race was divided based on their favourite ranking. Once again, favourites 17 and 18 were summed to form one category. The second method was to divide the data into two groups; races 1-7, and races 8-10, and then categorise these according to rank of favouritism. The results are as follows.

5.4.1 Separating Races 1, 2, 3,...,10

For this analysis, 170 separate objective-subjective probability pairings were calculated. The results from race 1 are presented in Table 5.13. The results from races 2, 3,...,10 are in Tables A.1-A.9 in Appendix A. Because the data was split amongst so many categories, this tended to effect the results through the size of the objective probability variance. To get an idea of the size of the categories, 61 had just over 1000 horses in them, while 60 had 600 horses or less. Because the objective probability variance is computed as $OP_i (1-OP_i) / N_k$, the small N_k values have caused relatively high variances. This in turn led to very few significant z values. In fact, of the 170 categories, only 14 z values were significant at least at the 10 percent level. Interestingly, those categories that did have significant z values tended to be neither strong favourites nor strong longshots. This is quite the opposite to what was found when looking at all races collectively. In the collective groupings, the significant z values tended to be amongst the heavy favourites and longshots, with very little evidence of underbetting or overbetting of the middle horses.

By simply glancing across the z value signs, it is apparent that favourites tend to have positive z values while longshots have predominately negative z values. This tends to suggest that the favourite longshot bias does exist but the lack of significant z values inhibits any strong conclusions to be made. A more appropriate way of determining the presence of the favourite longshot bias is to examine the weighted least squares regression test results which are presented in Table 5.14. The b_{1m} coefficient determines whether the favourite longshot bias is prevalent in the data or not. A value greater than unity confirms the favourite longshot bias while a value less than unity provides support for the opposite bias. 6 of the 10 b_{1m}'s are greater than 1, however only 2 of these are significantly greater than 1 at the 10 percent level. These are for races 4 and 8. To determine whether the favourite longshot bias becomes more prevalent throughout the raceday the equation $b_{1m} = \alpha + \beta(Race_m) + \varepsilon_i$ is estimated. The results of this estimation are presented in Table 5.15. A β value significantly greater than zero is required to conclude that the favourite longshot bias becomes more pronounced throughout the raceday. The β value for this data is –0.00124 and is not significantly different from zero. Hence we can conclude that there is no evidence to suggest, at least with this method, that the favourite longshot bias becomes more prevalent during the raceday with the New Zealand data.

Table 5.13: Race 1: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.10	1051	312	0.29686	0.28675	0.71716
2	$5.01	1051	193	0.18363	0.17342	0.85527
3	$6.63	1051	147	0.13987	0.13027	0.89740
4	$8.58	1051	106	0.10086	0.10074	0.01262
5	$11.00	1051	82	0.07802	0.07921	-0.14336
6	$13.87	1043	54	0.05177	0.06357	-1.71893*
7	$17.99	1020	39	0.03824	0.05015	-1.98465**
8	$22.47	945	37	0.03915	0.04083	-0.26585
9	$28.19	855	31	0.03626	0.03322	0.47437
10	$34.11	726	9	0.01240	0.02784	-3.76122***
11	$39.83	588	13	0.02211	0.02372	-0.26523
12	$46.73	464	17	0.03664	0.02026	1.87840*
13	$51.93	344	7	0.02035	0.01784	0.33007
14	$59.03	254	3	0.01181	0.01576	-0.58242
15	$67.38	133	3	0.02256	0.01394	0.66887
16	$76.21	80	0	0.00000	0.01224	-
17/18	$88.63	33	0	0.00000	0.01055	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.14: Weighted Least Squares Regression Test Results using Closing Win Odds

Description of Grouping	b_0	P-value$_0$	b_{1m}	P-value$_1$	F Statistic
Grouping by rank of favouritism Race 1	-0.00759 (0.00235)	0.00565	1.08209 (0.04975)	0.11968	5.20471 [0.01920]
Grouping by rank of favouritism Race 2	-0.00656 (0.00206)	0.00607	1.03133 (0.03890)	0.43313	5.64926 [0.01483]
Grouping by rank of favouritism Race 3	0.00597 (0.00246)	0.02803	0.95724 (0.04784)	0.38549	3.06483 [0.07656]
Grouping by rank of favouritism Race 4	-0.00811 (0.00175)	0.00032	1.06364 (0.03503)	0.08926	11.05509 [0.00112]
Grouping by rank of favouritism Race 5	-0.00049 (0.00258)	0.85202	0.99758 (0.04870)	0.96104	0.03409 [0.96656]
Grouping by rank of favouritism Race 6	-0.00536 (0.00188)	0.01216	1.01887 (0.03987)	0.64284	4.74649 [0.02529]
Grouping by rank of favouritism Race 7	0.00535 (0.00220)	0.02796	0.94248 (0.04319)	0.20280	2.95956 [0.08253]
Grouping by rank of favouritism Race 8	-0.00866 (0.00222)	0.00139	1.07746 (0.04372)	0.09677	7.72423 [0.00494]
Grouping by rank of favouritism Race 9	-0.00154 (0.00320)	0.63602	0.99443 (0.06597)	0.93385	0.20155 [0.81964]
Grouping by rank of favouritism Race 10	-0.00635 (0.00202)	0.00666	1.05936 (0.03500)	0.11049	4.96799 [0.02210]

Note. Regression results are WLS estimates for the function, $OP_k = b_0 + b_{1m}(SP_k) + e_k$, where e_k is a random error term. The P-value$_0$ is from the t test $b_0=0$, and P-value$_1$ is from the t test $b_1=1$. The F-statistic is the joint test $b_0=0$, $b_1=1$, with the appropriate P-value below it in brackets. Standard errors are shown in parentheses.

Table 5.15: Ordinary Least Squares Regression Test Results on how the Favourite Longshot Bias changes throughout the Raceday using Closing Win Odds

Description	α	β	P-value$_0$	F Statistic	R^2
Favourite longshot bias throughout the raceday	1.0292 (0.03554)	-0.00124 (0.00572)	0.83455	0.95394 [0.42505]	0.0053

Note. Regression results are OLS estimates for the function, $b_{1m} = \alpha + \beta(Race_m) + \varepsilon_i$, where ε_i is a random error term. The P-value$_0$ is from the t test $\beta=0$. The F-statistic is the joint test $\alpha=1$, $\beta=0$, with the appropriate P-value below it in brackets. Standard errors are shown in parentheses.

5.4.2 Separating Races 1-7, 8-10

The second method was to separate all the horses into two categories; those that ran in races 1-7 and those that ran in races 8-10. For simplicity, only 1 method of grouping was performed on each category, that is the grouping by rank of favouritism method. The results of races 1-7 are presented in Table 5.16 and races 8-10 are presented in Table 5.17.

In Table 5.16, out of 17 categories, only 4 have significant z values. Apart from the 1st and 6th favourites (which both have negative z values), the pattern is that favourites 1-9 have positive z values and 10-18 have negative z values. Note that the significant z values occurred at favourites 4, 10, 11, and 15, which tend to be the middle categories rather than the ends.

Table 5.17 had the problem of few horses in each category, resulting in the achievement of only one significant z value. The scattering of positive and negative z values meant no conclusions on the existence of the favourite longshot bias could be made for races 8-10 when looking at each category individually.

Table 5.18 presents the weighted least squares regression test results for the two previous groupings. The b_1 value for races 1-7 is 1.00860 and is not statistically significant from 1. The b_1 value for races 8-10 is 1.04421 and again is not significant from 1. To formally test whether the favourite longshot bias becomes more prevalent during the raceday, three separate tests are conducted. The first test is to establish whether the b_1 coefficient for races 8-10 (denoted as δ_2) is significantly greater than the b_1 coefficient for races 1-7 (denoted as δ_1). The test statistic is –1.5333, which is significant at the 10 percent level. The next two tests establish whether each coefficient, δ_1 & δ_2, are significantly different from δ_0 (the b_1 coefficient for races 1-10). Test statistics show that δ_1 and δ_0 are significantly different at the 20 percent level and δ_2 and δ_0 are significantly different at the 10 percent level. Because we would expect the difference $\delta_1 - \delta_2$ to be negative in 90% of tests, we can conclude that, with this method, there is sufficient evidence to support the hypothesis that the favourite longshot bias becomes more prevalent during the day.

Table 5.16: Races 1-7: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.15	7323	2043	0.27898	0.28404	-0.96427
2	$5.12	7323	1251	0.17083	0.17016	0.15213
3	$6.78	7322	950	0.12975	0.12764	0.53696
4	$8.64	7321	777	0.10613	0.10000	1.70409*
5	$10.89	7300	622	0.08521	0.07987	1.63326
6	$13.74	7237	463	0.06398	0.06423	-0.08876
7	$17.45	7061	371	0.05254	0.05157	0.36435
8	$21.75	6651	290	0.04360	0.04204	0.62543
9	$26.99	6066	209	0.03445	0.03442	0.01531
10	$32.63	5219	127	0.02433	0.02865	-2.02561**
11	$38.70	4323	79	0.01827	0.02421	-2.91542***
12	$45.04	3439	71	0.02065	0.02075	-0.04378
13	$51.69	2618	39	0.01490	0.01797	-1.29998
14	$59.70	1943	28	0.01441	0.01558	-0.43347
15	$65.51	1087	10	0.00920	0.01404	-1.67318*
16	$73.74	646	7	0.01084	0.01234	-0.36808
17/18	$80.69	295	3	0.01017	0.01127	-0.18855

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.17: Races 8-10: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking (1)	Average Odds (2)	N_k (3)	W_k (4)	OP_k (5)	SP_k (6)	z_k (7)
1	$3.31	3038	809	0.26629	0.26979	-0.43585
2	$5.24	3038	512	0.16853	0.16553	0.44160
3	$6.77	3038	399	0.13134	0.12729	0.66029
4	$8.49	3038	331	0.10895	0.10153	1.31309
5	$10.63	3033	249	0.08210	0.08134	0.15283
6	$13.12	3009	191	0.06348	0.06651	-0.68236
7	$16.42	2945	178	0.06044	0.05415	1.43230
8	$20.34	2797	127	0.04541	0.04432	0.27691
9	$24.96	2519	87	0.03454	0.03682	-0.62660
10	$30.02	2175	63	0.02897	0.03086	-0.52625
11	$35.71	1828	41	0.02243	0.02599	-1.02922
12	$41.39	1481	30	0.02026	0.02233	-0.56606
13	$48.08	1149	8	0.00696	0.01932	-5.03796***
14	$56.79	866	13	0.01501	0.01643	-0.34305
15	$63.42	514	8	0.01556	0.01462	0.17305
16	$70.05	296	4	0.01351	0.01296	0.08270
17/18	$79.37	162	0	0.00000	0.01155	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table 5.18: Weighted Least Squares Regression Test Results using Closing Win Odds

Description of Grouping	b_0	P-value$_0$	b_1	P-value$_1$	F Statistic
Grouping by rank of favouritism Race 1-7	-0.00171 (0.00064)	0.01763	1.00860 (0.01266)	0.50748	3.92773 [0.04249]
Grouping by rank of favouritism Race 8-10	-0.00567 (0.00179)	0.00642	1.04421 (0.03384)	0.21109	5.15597 [0.01976]

Note. Regression results are WLS estimates for the function, $OP_k = b_o + b_1(SP_k) + e_k$, where e_k is a random error term. The P-value$_0$ is from the t test $b_0=0$, and P-value$_1$ is from the t test $b_1=1$. The F-statistic is the joint test $b_o=0$, $b_1=1$, with the appropriate P-value below it in brackets. Standard errors are shown in parentheses.

5.5 Henery's Test

Table 5.19 presents the results of Henery's test when grouping by rank of favourtism using opening win odds. The three other methods using opening win odds and the four methods using closing win odds are presented in Tables B.1-B.7 in Appendix B. Column (1) is the rank of favouritism (except in B.1-B.3 and B.5-B.6, where it is the subjective probability intervals). Column (2) are the average odds for the particular category. Column (3) is the number of horses in each category. Column (4) is the number of winners in each category. Column (5) is the estimate of the categories true chance of losing while column (6) is the punters subjective probability of losing. By comparing the q_k figure (empirical lose probability) with the Q_k figure (subjective lose probability), one can compare the true probability of a horse losing with what the bettor perceives the probability to be. In effect, if $Q_k > q_k$, then the punter believes the horses in that particular category have a worse chance of winning than historical data shows and vice versa.

Looking at all eight tables, the first interpretation is that punters are a lot more optimistic about a horse's chance than the historical data indicates. From all 8 groupings, a total of 242 Q_k, q_k pairings were calculated. Of these, in only two cases do we find that $Q_k>q_k$. For the remaining 240 categories, the punters believe the horses have a better chance of winning than historical data indicates. This result is not unexpected. Since bettors are participating in a system where they lose (on average), then surely it must be the case that they lose more often than they expect. Essentially, it is from this difference that the T.A.B. makes its money.

In general, the difference between the Q_k's and q_k's is greater for favourites and lower for longshots. For example, for the most favoured category in Table B.3, the true probability of losing is 0.27368, whereas the punter's subjective probability of losing is 0.18420. Compare this with the least favoured category in the same table, where the true probability is 0.99564 and the punter's subjective probability of losing is 0.98839. Interestingly, the exception to this trend is the two Q_k, q_k pairings for which $Q_k>q_k$. Contrary to the trend, these two categories are for favourites. The first occurrence is in Table B.2. The Q_k for the subjective probability interval 0.192736-0.215233 is 0.75433, while the true probability of losing is 0.75275. The second occurrence is in Table B.7, with the subjective probability interval 0.55-0.6.

Table 5.19: Henery's Test by Rank of Favouritism using Opening Win Odds

Ranking (1)	Average Odds (2)	N_k (3)	W_k (4)	q_k (5)	Q_k (6)
1	$3.62	10361	2775	0.73217	0.72399
2	$5.60	10361	1775	0.82868	0.82142
3	$7.14	10361	1306	0.87395	0.86002
4	$8.71	10359	1115	0.89236	0.88523
5	$10.50	10333	829	0.91977	0.90476
6	$12.62	10246	711	0.93061	0.92078
7	$15.25	10006	573	0.94273	0.93444
8	$18.29	9448	444	0.95301	0.94532
9	$21.99	8585	296	0.96552	0.95452
10	$25.89	7394	178	0.97593	0.96138
11	$30.15	6151	139	0.97740	0.96683
12	$34.59	4920	92	0.98130	0.97109
13	$39.76	3767	59	0.98434	0.97485
14	$46.49	2809	50	0.98220	0.97849
15	$51.85	1601	29	0.98189	0.98071
16	$57.64	942	14	0.98514	0.98265
17/18	$66.99	457	5	0.98906	0.98507

Table 5.20: OLS & WLS Regression Test Results on Henery's Test using Opening Win Odds

Description of Grouping	Coefficient f	Lower	Upper	r^2
Grouping by even odd range categories	0.95468 (0.01213)	0.9298	0.9795	0.9429
Grouping by 25 even sized odds categories	0.98207 (0.00199)	0.9780	0.9862	0.9886
Grouping by 50 even sized odds categories	0.98161 (0.00185)	0.9779	0.9853	0.9810
Grouping by rank of favouritism	0.99475 (0.00094)	0.9928	0.9967	0.9931

Note. Regression results are OLS & WLS estimates for the function, $Q_k = fq_k + e_i$, where e_i is a random error term. The dependent variable, Q_k, equals the punters subjective losing probability for horses in the k^{th} category, and the independent variable is the empirical lose probability for horses in the k^{th} category. Lower is the lower 95% confidence interval, upper is the upper 95% confidence interval. Standard errors are shown in parentheses.
The WLS procedure is used for the grouping by rank of favouritism method.

Table 5.21: OLS & WLS Regression Test Results on Henery's Test using Closing Win Odds

Description of Grouping	Coefficient f	Lower	Upper	r^2
Grouping by even odd range categories	0.95753 (0.00783)	0.9415	0.9736	0.9760
Grouping by 25 even sized odds categories	0.98234 (0.00354)	0.9750	0.9897	0.9715
Grouping by 50 even sized odds categories	0.98177 (0.00275)	0.9763	0.9873	0.9662
Grouping by rank of favouritism	0.99429 (0.00145)	0.9912	0.9974	0.9868

Note. Regression results are OLS & WLS estimates for the function, $Q_k = fq_k + e_i$, where e_i is a random error term. The dependent variable, Q_k, equals the punters subjective losing probability for horses in the k^{th} category, and the independent variable is the empirical lose probability for horses in the k^{th} category. Lower is the lower 95% confidence interval, upper is the upper 95% confidence interval. Standard errors are shown in parentheses.
The WLS procedure is used for the grouping by rank of favouritism method.

Table 5.20 presents the ordinary least squares and weighted least squares regression test results on Henery's test using opening odds. The first thing to note is that the equation fits the data extremely well, indicated by the high r^2 values (the smallest is 0.9429 while the largest is 0.9931). Secondly, the estimated f coefficients are all significantly below zero. Hence, there is strong evidence to suggest that bettors are discounting a percentage of their losses. However, the actual percentage of discounting is a little more debatable. We have 4 estimates for f (essentially 3 estimates since the estimates for f are nearly identical for the 25 and 50 grouping methods) The estimate for the grouping by even subjective probability categories is very low at 0.95468 and is in fact significantly different from all other estimates of f at the 5 percent level. The estimate for the grouping by rank of favouritism is very high at 0.99475 and while it is still significantly different from 1 at the 5 percent level, it would not explain much of the variation between the subjective and objective odds. The two estimates that seem more realistic and are close to the estimates Henery got for his study are 0.98207 for 25 even sized groups and 0.98161 for 50 even sized groups. These suggest that bettors are discounting approximately 2 percent of their losses.

Table 5.21 presents the same ordinary and weighted least squares regression analysis as Table 5.20 but this time using closing odds. Again the equation fits the data well as indicated by the high r^2 values, and again we get three estimates of the f coefficient. The results are very similar to those in Table 5.20. The f estimate for the even ranged category is very low at 0.95753, indicating bettors discount approximately 4.3 percent of their losses. The high f estimate is again for the favourite ranking method at 0.99429 and the estimates consistent with Henery are 0.98234 for the 25 even sized grouping method and 0.98177 for the 50 even sized grouping method. As with the opening odds, all f estimates are significantly different from 1 at the 5 percent level.

5.6 Testing for Skewness

The mean, variance and skewness results when grouped by rank of favouritism using opening odds are presented in Table 5.22. The three other methods using opening odds and the four methods using closing odds are presented in Tables C.1-C.7 in Appendix C. Column (1) is the favourite ranking (except in C.1-C.3 and C.5-C.7 where it is the subjective probability intervals), column (2) is the number of horses in each category, column (3) is the objective probability, and columns (4), (5), and (6) are the mean, variance and skewness respectively.

Recall that one condition necessary for the three moment model to explain the favourite longshot bias is that the variance and skewness must increase as the objective probability decreases. Therefore, it was examined how many times this occurred by calculating the percentage of times the variance and skewness estimates moved in the opposite direction to the objective probability.

Looking at the opening odds, which are covered in Tables 5.22 and C.1-C-3, the first impression is that overall, there seems to be a negative relationship between the objective probability and the variance/skewness coefficients. For the variance – objective probability relationship in Table 5.22, 75% of the objective probability movements were coupled with the variance moving in the opposite direction. The other percentages were 83%, 51%, and 86% for the 25 even sized grouping method, 50 even sized grouping method and the even subjective probability categories method respectively. For the skewness – objective probability relationship, 88% of the objective probability movements were associated with skewness movements in the opposite direction when grouping by rank of favouritism. The other percentages were 92%, 73%, and 93% for the 25 even sized grouping method, 50 even sized grouping method and the even subjective probability categories method respectively.

As for the closing odds, which are presented in Tables C.4-C.7, again there seems to be a distinct negative relationship between the objective probability and the variance/skewness coefficients. For the variance – objective probability relationship, the percentages were 63%, 83%, 61%, and 93% for the favourite ranking method, 25 even sized groups, 50 even sized groups and the even subjective probability categories method respectively. The skewness – objective probability relationship produced similar percentages at 75%, 92%, 65%, and 90% respectively.

Note that of the 8 Tables, the lowest percentages always occurred for the grouping by 50 even sized method – the most segregated method. This is consistent with what Golec & Tamarkin (1998) found when comparing the studies of Ali (1977), Hausch, Ziemba, William & Rubenstein (1981) and Asch & Quandt (1987). The studies of Ali (1977) and Hausch et al (1981) only had 8 and 10 categories respectively while Asch &

Table 5.22: Mean, Variance and Skewness of Unit Bet Returns by Rank of Favouritism using Opening Win Odds

Ranking	N_k	OP_k	Mean	Variance	Skewness
(1)	(2)	(3)	(4)	(5)	(6)
1	10361	0.26783	-0.13385	2.050837	3.0796
2	10361	0.17132	-0.09545	3.95783	13.73732
3	10361	0.12605	-0.14364	5.084564	25.83512
4	10359	0.10764	-0.10558	6.632319	43.24818
5	10333	0.08023	-0.19896	7.356346	61.66407
6	10246	0.06939	-0.17329	9.165419	94.03702
7	10006	0.05727	-0.18588	10.9111	137.3519
8	9448	0.04699	-0.20858	12.7019	193.8061
9	8585	0.03448	-0.3146	13.15526	243.4794
10	7394	0.02407	-0.44129	12.65478	279.5586
11	6151	0.02260	-0.39458	15.8531	405.5219
12	4920	0.01870	-0.42535	17.32944	512.6382
13	3767	0.01566	-0.44939	19.05373	648.8556
14	2809	0.01780	-0.27467	29.03005	1140.827
15	1601	0.01811	-0.17358	37.02179	1627.897
16	942	0.01486	-0.23565	38.72631	1932.491
17/18	457	0.01094	-0.35168	37.99715	2202.317

Quandt (1987) had 20 categories and hence was the most segregated. Golec & Tamarkin (1998) found that only in Asch & Quandt's (1987) study were there few instances where the variance and skewness coefficients moved in the same direction as the objective probability. This may have been due to the fact that for the very segregated methods, there was not a significant movement in the objective probabilities from one category to the next. Overall, the negative relationship pattern may exist, but it is possible with small deviations in the objective probability that the variance and skewness coefficients could move in the same direction.

Across all the grouping methods, despite some minor exceptions, there is a clear conclusion that a negative relationship exists between the objective probability and the variance and skewness coefficients.

Of the 242 skewness coefficient estimates across all 8 grouping methods, a mere 6 indicated negative skewness. These instances only occurred with the even subjective probability grouping method (Tables C.3 and C.7) and with extreme favourites (objective probabilities greater than 0.5). Note however that in these instances, the number of observations are relatively small (259 observations is the highest, while 22 the lowest). Without a doubt, there is a positive skewness in the New Zealand horse racing data and this positive skewness becomes more pronounced as we move from favourites to longshots.

5.7 The Power Utility Model

The ordinary least squares and weighted least squares regression estimates of the power utility model, $u(OW_k) = \alpha OW_k^{\beta} + \varepsilon_k$, are presented in Tables 5.23 (opening odds) and 5.24 (closing odds). Recall that it is the value of β that determines the bettor's risk preference. If $\beta>1$, then bettors are risk lovers since $u'>0$, $u''>0$, and $u'''<0$ and if $\beta<1$, then this produces $u'>0$, $u''<0$, and $u'''>0$ which would imply risk averse behaviour.

The two inferences from the tables are firstly, the New Zealand data is consistent with bettors having risk loving preferences since the β estimates for all 8 groupings methods are all significantly greater than zero at the 5 percent level (except for the grouping by rank of favouritism using closing odds which is significant at the 10 percent level).

Table 5.23: OLS & WLS Regression Estimates for Power Utility Model using Opening Win Odds

Description of Grouping	Coefficient Estimates α	β	P-value	r^2
Grouping by rank of favouritism	-4.4963 (0.1676)	1.05846 (0.04394)	0.20331	0.9748
Grouping by 25 even sized odds categories	-5.2365 (0.06477)	1.2133 (0.02443)	0.00000	0.9908
Grouping by 50 even sized odds categories	-5.5309 (0.05749)	1.2103 (0.02169)	0.00000	0.9848
Grouping by even odd range categories	-5.4149 (0.04992)	1.1188 (0.02544)	0.00014	0.9862

Note. Regression results are OLS & WLS estimates for the power utility function, $u(OW_k) = \alpha OW_k^{\beta} + \varepsilon_K$, where ε_K is a random error term. The dependent variable, $u(OW_k)$, equals the ratio of the objective win probability of horses in the lowest probability category to the objective win probability for horses in the k^{th} category, and the independent variable is the average win price for horses in the k^{th} category. The P-value is from the t test $\beta>1$. Standard errors are shown in parentheses.
The WLS procedure is used for the grouping by rank of favouritism method.

Table 5.24: OLS & WLS Regression Estimates for Power Utility Model using Closing Win Odds

Description of Grouping	Coefficient Estimates α	β	P-value	r^2
Grouping by rank of favouritism	-5.33993 (0.2641)	1.17210 (0.06598)	0.01977	0.9546
Grouping by 25 even sized odds categories	-5.3111 (0.08777)	1.1351 (0.03184)	0.00062	0.9822
Grouping by 50 even sized odds categories	-5.4711 (0.06891)	1.1347 (0.02501)	0.00000	0.9772
Grouping by even odd range categories	-5.1706 (0.03754)	1.0871 (0.01915)	0.00020	0.9917

Note. Regression results are OLS & WLS estimates for the power utility function, $u(OW_k) = \alpha OW_k^{\beta} + \varepsilon_K$, where ε_K is a random error term. The dependent variable, $u(OW_k)$, equals the ratio of the objective win probability of horses in the lowest probability category to the objective win probability for horses in the k^{th} category, and the independent variable is the average win price for horses in the k^{th} category. The P-value is from the t test $\beta>1$. Standard errors are shown in parentheses.
The WLS procedure is used for the grouping by rank of favouritism method.

Secondly, the New Zealand data fits the power utility model well since the r^2 values are very high (the lowest being 0.9546).

5.8 The Cubic Utility Model

Tables 5.25 (opening odds) and 5.26 (closing odds) present the ordinary least squares and weighted least squares regression estimates for the cubic utility model of the form $u(OW_k) = \gamma_0 + \eta_1 OW_k + \eta_2 OW_k + \eta_3 OW_k + e_k$. If $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$, then the bettors are risk averse, that is they prefer mean and skewness but dislike variance.

The mean coefficients, (η_1) for all 8 grouping methods are positive and 6 are significantly positive at least at the 10 percent level. The two insignificant positive coefficients are for the grouping by 25 even sized categories and grouping by 50 even sized categories, and are both at closing odds. The results suggest that bettors do in fact prefer higher means.

The variance coefficients (η_2) provide somewhat mixed results. 2 out of the 8 coefficients are negative, while only one of these is significant. The remaining 6 are positive, with 3 being significant.

The skewness coefficients (η_3) are predominately positive (6 out of the 8 groupings) and once again the two instances where the sign is negative occur at closing odds. Also, of the 6 positive coefficients, 4 are significant at least at the 10 percent level.

Overall, the grouping by 50 even sized categories (opening odds) and the grouping by rank of favouritism (opening odds) are the only instances where the results $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$ are achieved. However, only in the grouping by rank of favouritism (opening odds) were all coefficients significant.

By far, the opening odds return coefficients that are closer to the proposed hypothesis that $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$ than the closing odds do.

The next chapter revisits the results and offers explanations as to their size, significance and interpretation.

Table 5.25: OLS & WLS Regression Estimates for Cubic Utility Model using Opening Win Odds

Description of Grouping	Coefficient Estimates γ_0	η_1	η_2	η_3	F-statistic	r^2
Grouping by rank of favouritism	-0.21690 (0.09172)	0.04712** (0.01037)	-0.00109** (0.00030)	0.000010** (0.000003)	150.028	0.9719
Grouping by 25 even sized odds categories	-0.01202 (0.01122)	0.00899** (0.00172)	0.00008 (0.00007)	0.00000003 (0.0000007)	1625.757	0.9957
Grouping by 50 even sized odds categories	-0.01841 (0.00833)	0.00850** (0.00115)	-0.00002 (0.00004)	0.0000008* (0.0000003)	1463.875	0.9896
Grouping by even odd range categories	-0.00244 (0.00181)	0.00561** (0.00040)	0.00001 (0.00001)	0.0000006** (0.0000001)	142932.497	0.9994

Note. Regression results are OLS & WLS estimates for the cubic utility function, $u(OW_k) = \gamma_o + \eta_1 OW_k + \eta_2 OW_k^2 + \eta_3 OW_k^3 + e_k$, where e_k is a random error term. The dependent variable, $u(OW_k)$, equals the ratio of the objective win probability of horses in the lowest probability category to the objective win probability for horses in the k[th] category, and the independent variable is the average win price for horses in the k[th] category. The F-statistic tests $\eta_1=\eta_2=\eta_3=0$. Standard errors are shown in parentheses.
** single t tests that $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$ significant at 1 percent level
* single t tests that $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$ significant at 10 percent level
The WLS procedure is used for the grouping by rank of favouritism method.

Table 5.26: OLS & WLS Regression Estimates for Cubic Utility Model using Closing Win Odds

Description of Grouping	Coefficient Estimates γ_0	η_1	η_2	η_3	F-statistic	r^2
Grouping by rank of favouritism	0.01103 (0.00605)	0.01160** (0.00061)	0.000008 (0.00002)	0.00000026 (0.0000009)	28743.252	0.9998
Grouping by 25 even sized odds categories	0.01855 (0.01723)	0.00258 (0.00233)	0.00016* (0.00007)	-0.0000006 (0.0000006)	511.499	0.9865
Grouping by 50 even sized odds categories	0.02471 (0.01073)	0.00054 (0.00133)	0.00020** (0.00004)	-0.0000010** (0.0000003)	855.730	0.9824
Grouping by even odd range categories	0.00162 (0.00135)	0.00582** (0.00295)	0.00002* (0.00001)	0.0000005** (0.00000009)	26869.432	0.9997

Note. Regression results are OLS & WLS estimates for the cubic utility function, $u(OW_k) = \gamma_o + \eta_1 OW_k + \eta_2 OW_k^2 + \eta_3 OW_k^3 + e_k$, where e_k is a random error term. The dependent variable, $u(OW_k)$, equals the ratio of the objective win probability of horses in the lowest probability category to the objective win probability for horses in the k[th] category, and the independent variable is the average win price for horses in the k[th] category. The F-statistic tests $\eta_1=\eta_2=\eta_3=0$. Standard errors are shown in parentheses.
** single t tests that $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$ significant at 1 percent level
* single t tests that $\eta_1>0$, $\eta_2<0$ and $\eta_3>0$ significant at 10 percent level
The WLS procedure is used for the grouping by rank of favouritism method.

6. Discussion of Results

6.1 Weak Form Efficiency

Recall Fama's (1970) definition of weak form efficiency was that historical information is fully reflected in the prices. Hence, for the New Zealand racing industry to be weak form efficient, we require an expected return of (1-t), irrespective of the strategy adopted. Tables 5.1-5.4 present 122 differing strategies adopted from the data set. These strategies fell into two categories; consistent betting on the n^{th} favourite, or consistent betting on horses within a predetermined subjective probability category.

Of the 122 strategies, 120 had negative returns with a mere 2 strategies yielding positive returns. The first occurrence is in Table 5.2 for the continual backing of horses that fell in the subjective probability range 0.55-0.6. In this entire data set, a horse fell into this category only 47 times. A dollar bet on these horses would have yielded $47.15, a 0.32% return. The second occurrence was for the continual backing of horses in the 0.036845-0.039012 subjective probability range in Table 5.4. A dollar bet on the 2362 horses in this category would have yielded $2363.30, a 0.06% return. As you can see, these positive returns are very small and are definitely not significantly different from a zero return. However, these returns are significantly different from the weak form efficiency return of 0.830542. Any financial advisor would surely be glad to find a strategy that yields a return 17.0058% higher than the weak form efficiency return.

If advice to bettors could be offered it would be this. Once you establish that over time you cannot gain a positive return from betting but enjoy the experience of trying to do so, why not try to minimise your losses. That way you can still participate in the activity that returns a positive level of utility but at a minimum cost to you. Such strategies explained above may achieve this goal.

Looking at Figures 5.1-5.4, it seems that the New Zealand data does not fit Fama's (1970) weak form efficiency criteria. Instead, there seems to be a pattern that occurs irrespective of the grouping method adopted. This pattern is that there is a negative trade-off between average odds and the average return. It seems that if bettors place their money at higher odds, they can expect a lower return. To verify this, a regression line is estimated through the data points for the 4 graphs. A significant negative slope coefficient would confirm the hypothesis that for the New Zealand horse racing data, the risk-return trade-off is negative. The results of the ordinary least squares regression estimates are presented below in Table 6.1

Table 6.1: OLS Regression Estimates of the Risk-Return Trade-off

Description of Grouping	Coefficient Estimates Constant	Slope	P-value	r^2
Grouping by rank of favouritism	0.97301 (0.07212)	-0.02882 (0.00666)	0.00052	0.5391
Grouping by 25 even sized odds categories	0.90042 (0.02089)	-0.00557 (0.00077)	0.00000	0.6938
Grouping by 50 even sized odds categories	0.89870 (0.01789)	-0.00549 (0.00066)	0.00000	0.5922
Grouping by even odd range categories	0.87578 (0.01274)	-0.00499 (0.00065)	0.00000	0.6873

The P-value is from the t test slope = 0. Standard errors are given in parentheses.

Every grouping method gives a significant negative slope at the 1 percent level. Hence, we can conclude that there exists a negative risk-return trade-off in the New Zealand data. Contrary to the risk-return trade-off experienced in a stock market, where one accepts more risk for a compensating higher return, the racing track is not so forgiving. In a stock market you are compensated for withstanding risk, however, in a betting market you are punished for doing so.

So why does this trade-off occur? The trade-off shows that favourites tend to have expected returns greater than the weak form efficiency return, while longshots have expected returns lower than the weak form efficiency return. For this to occur, it is necessary that favourites are bet on less than they should be, hence they are underbet, and longshots are bet on more than they should be, hence they are overbet. Readers will recognise this phenomenon as the favourite longshot bias. We can conclude that the negative trade-off between risk and return experienced in the New Zealand data is a direct result of the favourite longshot bias being present in the data. The favourite longshot bias, its results and explanations are discussed in the following section.

6.2 Favourite Longshot Bias

The favourite longshot bias is said to exist in a betting market where for favourites, the subjective probability is less than the objective probability (underbetting) and for longshots, the subjective probability is greater than the objective probability (overbetting).

6.2.1 Opening Odds

Results from the 4 grouping methods for opening odds are presented in Tables 5.5-5.8. The z value measured the extent to which the subjective probability differed from the objective probability. The z value is therefore an indication of underbetting (positive values) and overbetting (negative values).

The z values for the first few longshot categories, over all 4 methods, were negative, and most were significant well beyond the 1 percent level. For this reason, it is possible to conclude that, at opening odds, longshots are overbet in the New Zealand racing data. The z values for favourites, on the other hand, offered limited evidence that they were underbet. For the two grouping methods, grouping by rank of favouritism and grouping by 25 even sized subjective probability categories, the z values for favourites were positive and significant. However, for the other two methods, the z values seemed to follow no distinct pattern. Hence, we can only say that whether the favourites are being underbet is dependant upon the grouping method adopted.

An ordinary least squares and weighted least squares procedure is used to estimate the equation $OP_k = b_0 + b_1(SP_k) + e_k$ for each grouping method to test for the existence of the favourite longshot bias over an entire grouping. The results of these estimations for opening odds are presented in Table 6.2. For market efficiency and for the non-existence of the favourite longshot bias, we require $b_0=0$, $b_1=1$. If $b_1>1$, then this indicates that the favourite longshot bias exists since it shows that as we move from longshots to favourites, the objective probability increases faster than the subjective probability. Hence, at low levels of subjective probability, $SP_k>OP_k$ and at high levels of subjective probabilities, $SP_k<OP_k$.

For all of the methods, except the grouping by even subjective probability categories, there is strong evidence to suggest that the favourite longshot bias is prevalent in the New Zealand data 30 minutes before the tote closes. Excluding the grouping by even subjective probability categories, the t tests on $b_0=0$ were all significant beyond the 1 percent level as were the t tests on $b_1=1$. It then follows that the joint F test that $b_0=0$ and $b_1=1$ were all significant beyond the 1 percent level. The insignificance of the t tests for the grouping by even subjective probability categories meant no conclusions could be reached for this method.

Table 6.2: OLS & WLS Regression Test Results using Opening Win Odds

Description of Grouping	b_0	P-value$_0$	b_1	P-value$_1$	F-statistic
Grouping by rank of favouritism	-0.00437 (0.00102)	0.00066	1.06636 (0.02041)	0.00536	9.48007 [0.00218]
Grouping by 25 even sized odds categories	-0.00796 (0.00242)	0.00321	1.19876 (0.02137)	0.00000	62.47340 [0.00000]
Grouping by 50 even sized odds categories	-0.00653 (0.00189)	0.00117	1.07723 (0.01660)	0.00003	10.84784 [0.00013]
Grouping by even odd range categories	0.00336 (0.00713)	0.64097	1.00018 (0.02470)	0.99413	0.31155 [0.73491]

Note. Regression results are OLS & WLS estimates for the function, $OP_k = b_0 + b_1(SP_k) + e_k$, where e_k is a random error term. The P-value$_0$ is from the t test $b_0=0$, and P-value$_1$ is from the t test $b_1=1$. The F-statistic is the joint test $b_0=0$, $b_1=1$, with the appropriate P-value below it in brackets. Standard errors are shown in parentheses.

The WLS procedure is used for the grouping by rank of favouritism method.

6.2.2 Closing Odds

The results from the 4 grouping methods using closing odds are presented in Tables 5.9-5.12. Like the opening odds, it was possible to conclude that longshots were still being overbet when the tote closed, since the z values for the first few longshot categories in each method were negative and most were significant at the 1 percent level. As for the favourites, there was even less evidence to suggest that they were being underbet when the tote closed, than 30 minutes before the tote closed.

The ordinary least squares and weighted least squares regression estimates of the equation $OP_k = b_0 + b_1(SP_k) + e_k$ for the closing odds are presented in Table 6.3. Recall that for a betting market to exhibit the favourite longshot bias, it requires $b_0<0$ and $b_1>1$. This was achieved across every grouping method for the New Zealand data at the closing odds. However, only for the even subjective probability categories method were the coefficients significantly different from the market efficiency values. Looking at the F statistics which jointly tested $b_0=0$ and $b_1=1$, only the even subjective probability categories method was significant.

Overall, there seemed to be a favourite longshot bias 30 minutes before the tote closed, that is eliminated in the final 30 minutes. This can be explained by examining when the betting takes place and what type of bettors participate. Firstly, most bets are placed in the last few minutes before a race. The amount of betting up until 30 minutes before the tote closes makes up a very small proportion of the total bets placed. Hence, when examining the opening odds in this thesis, we are examining a small and maybe unrepresentative proportion of the bets. Additionally, the more experienced bettor knows that in order to have a greater chance of predicting a horse's objective probability, they have to accumulate as much information about the horses as possible before placing their bets. That is, if you can be the last person to place a bet on a race, then you have the advantage of potentially absorbing the most amount of information. For this reason, if there exists a section of bettors with superior information, then one would expect these bettors to enter the market as late as possible, and definitely not before the 30 minute mark. It is proposed that the favourite longshot bias is eliminated in the final 30 minutes because of insider traders. For the two sections of bettors, those that bet before the 30 minute mark and those that bet after the 30 minute mark, the proportion of insider traders is substantially higher for the latter section.

Table 6.3: OLS & WLS Regression Test Results using Closing Win Odds

Description of Grouping	b_0	P-value$_0$	b_1	P-value$_1$	F-statistic
Grouping by rank of favouritism	-0.00313 (0.00076)	0.00070	1.02122 (0.01480)	0.17235	9.53768 [0.00213]
Grouping by 25 even sized odds categories	-0.00080 (0.00129)	0.53952	1.01457 (0.01092)	0.19528	1.04137 [0.36903]
Grouping by 50 even sized odds categories	-0.00087 (0.00121)	0.47905	1.01270 (0.01024)	0.22115	0.81428 [0.47905]
Grouping by even odd range categories	-0.02040 (0.00813)	0.01826	1.14068 (0.02840)	0.00003	15.13926 [0.00000]

Note. Regression results are OLS & WLS estimates for the function, $OP_k = b_o + b_1(SP_k) + e_k$, where e_k is a random error term. The P-value$_0$ is from the t test $b_0=0$, and P-value$_1$ is from the t test $b_1=1$. The F-statistic is the joint test $b_o=0$, $b_1=1$, with the appropriate P-value below it in brackets. Standard errors are shown in parentheses.

The WLS procedure is used for the grouping by rank of favouritism method.

It is believed that the inexperienced bettors create the favourite longshot bias up until the 30 minute mark, and the insider traders, who have a greater understanding of a horse's true winning potential, eliminate the bias in the last 30 minutes.

Another possible explanation is that bettors up until the 30 minute mark distort the odds too much. Because the favourite longshot bias is so prevalent 30 minutes out from the close of the tote, it must be that the odds for favourites are exceptionally high and the odds for longshots exceptionally low, given their true chances of winning. For the favourite longshot bias to become almost eliminated in the final 30 minutes, the punters must be betting proportionately more on the favourites than longshots. This may be because punters recognise these big deviations bought about because of the high favourite longshot bias presence at the 30 minute mark.

6.2.3 Favourite Longshot Bias throughout the Raceday

The idea that the favourite longshot bias becomes more prevalent throughout the raceday is based on the assumption that a proportion of gamblers must back a winning longshot at the end of the day in order to recoup the losses already sustained in previous races. Hence, if this is true, longshots should be even more overbet on the last few races than the first few races. This assumption seems to be valid but whether the results substantiate the theory depends on the method used to test it.

6.2.3.1 Separating Races 1, 2, 3, …, 10

Tables 5.14 and 5.15 present the results from examining how the favourite longshot bias changes when dividing the data into separate race groups. In only two races were the b_{1m} coefficients significantly greater than 1 at the 10 percent level. This occurred for races 4 and 8. These results by themselves were insufficient to conclude whether or not the favourite longshot bias becomes more prevalent throughout the raceday. Table 5.15 gives the ordinary least squares regression results of estimating the equation $b_{1m} = \alpha + \beta(Race_m) + \varepsilon_i$. To accept the hypothesis that the favourite longshot bias is becoming more pronounced, we require the coefficient β to be significantly positive. Contrary to the theory, the β coefficient was –0.00124 and insignificant. For this reason, it would be incorrect to conclude that the favourite longshot bias becomes more prevalent during the raceday, using this method.

It was found that this method led to the data becoming too segregated. This meant that the objective probability variances were relatively high and hence produced relatively few significant z values. In an attempt to avoid this and to still test how the favourite longshot bias changes throughout the raceday, another method was adopted. This method was to split the results into two groups; races 1-7 and races 8-10. The results of this method are explained below.

6.2.3.2 Separating Races 1-7, 8-10.

The results from this method are presented in Table 5.18. Both b_1 coefficients were greater than 1, indicating that there could be a favourite longshot bias present however, neither of these coefficients were significant. For the favourite longshot bias to become more prevalent, we require the b_1 coefficient for races 8-10 to be greater than the b_1 coefficient for races 1-7. What's more, we require the difference to be significant. To estimate this difference, three tests were calculated. These were; the difference between δ_1[18] and δ_2, the difference between δ_0 and δ_1 and the difference between δ_0 and δ_2. The results of these tests are presented in Table 6.4 below

Table 6.4: Testing how the Favourite Longshot Bias changes throughout the Raceday

Test	Test Statistic	df	P-value
Difference between δ_1 and δ_2	-1.53333	17	0.07180
Difference between δ_0 and δ_1	0.85541	34	0.19916
Difference between δ_0 and δ_2	-1.55338	34	0.06480

The first test is the most important since it tests whether the b_1 coefficient for races 8-10 is significantly different from the b_1 coefficient for races 1-7. The other two tests simply test whether the b_1 coefficients for races 1-7 and 8-10 are significantly different from the coefficient for all races. If the theory is correct, we would expect that the b_1 coefficient for races 1-10 to fall within the b_1 coefficients for races 1-7 and races 8-10.

Looking at Table 6.4, because we would expect δ_2 to be greater than δ_1 in 90 out of 100 samples, we can conclude that using this method, there is evidence to suggest that the favourite longshot bias is more prevalent for races 8-10 than for races 1-7. The other two tests are not unexpected. The results from the 2[nd] and 3[rd] tests show that δ_1 is

[18] Recall δ_0 is the b_1 coefficient for races 1-10, δ_1 is the b_1 coefficient for races 1-7 and δ_2 is the b_1 coefficient for races 8-10.

statistically different from δ_0 at the 20 percent level and δ_2 is significantly different from δ_0 at the 10 percent level, which also provides evidence to suggest that the favourite longshot bias is becoming more prevalent throughout the raceday.

Overall, it is believed that the second method is more appropriate to test how the favourite longshot bias changes throughout the raceday, since the first method requires the data to be too segregated.

It must also be noted that bettors may be switching towards trifectas and quinellas later in the day rather than longshots, since these also yield high returns. This effect is not covered in this thesis but may explain why the first method of separating horses into races 1, 2, 3, ...,10 was not conclusive.

6.3 Henery's Test

Results so far have shown that the average return to punters from a bet of given subjective probability odds decreases nearly linearly with the odds. It was Henery's intention to explain why punters would accept long odds when they are clearly unfavourable. It was shown previously in the data set of Asch, Malkiel & Quandt (1982) that while the 9^{th} favourites had average odds of $45.25, the odds that truly reflected these horse's true winning potential were closer to $135.80. Henery is asking the question 'why do punters accept odds of $45.25 when this is clearly unfavourable?'. Henery believed punters discount a constant fraction (1-f) of their losses as 'not typical'. The estimation of f is from the equation $Q_k=fq_k$, where Q_k is the subjective lose probability and q_k is the empirical lose probability.

The computation of the Q_k, q_k pairings for the New Zealand data over all 8 groupings (4 at opening odds and 4 at closing odds) is presented in Tables 5.19 and B.1-B.7. The ordinary least squares estimation of f is presented in Tables 5.20 and 5.21.

The estimation of f is dependant upon the type of grouping method adopted. It was as low as 0.95468 for the grouping by even subjective probability categories, and as high as 0.99475 for the grouping by rank of favouritism. Interestingly, each of the f estimates for the closing odds were higher (although not significantly higher) than the f estimates for the opening odds, except for the grouping by rank of favouritism method. This could be due to the fact that when the more experienced bettors enter the market in the final 30 minutes, they increase the estimate of f because collectively, they discount less of their losses as 'not typical'. It seems reasonable to assume that bettors with a greater understanding of the market will also have a greater understanding of their losses.

It is believed that the estimate of f for the grouping by even subjective probability categories was unrealistically low. While an estimate of this size would explain a great proportion of the differences between the objective probability and the subjective probability, this method continually gave results that were quite inconsistent with the other three methods adopted in this thesis. For instance, the estimation of f for this method was significantly different from the estimates of f in the other 3 methods at the 5 percent level.

The estimates of f for the grouping by 25 and 50 even sized subjective probability categories were very similar to each other and to Henery's results of 0.974 (1975 flat season) and 0.978 (1979-80 flat season). Hence, it seems reasonable to conclude that punters discount approximately 2% of their losses as 'not typical'. To understand the implications of this model, the table below presents the true odds and the acceptable odds for various objective probabilities. The f used in the table is 0.98177 (grouping by 50 even sized odds groups using closing odds).

Table 6.5: Implications of Henery's Test

Objective Probability	True odds	Acceptable odds
0.01	$100.00	$35.65
0.05	$20.00	$14.85
0.1	$10.00	$8.59
0.15	$6.67	$6.04
0.2	$5.00	$4.66
0.25	$4.00	$3.79
0.3	$3.33	$3.20
0.35	$2.86	$2.76
0.4	$2.50	$2.43
0.45	$2.22	$2.17
0.5	$2.00	$1.96
0.55	$1.82	$1.79
0.6	$1.67	$1.65
0.65	$1.54	$1.52
0.7	$1.43	$1.42
0.75	$1.33	$1.33
0.8	$1.25	$1.24
0.85	$1.18	$1.17
0.9	$1.11	$1.11
0.95	$1.05	$1.05

From this table, one can see that this model explains the large variations in the actual and acceptable odds for longshots, although it does not account for very much variation in the actual and acceptable odds for favourites. This is because favourites have very small odds so there is little room for deviations between the actual and acceptable odds.

Overall, because every estimate of f was significantly different from 1 at the 5 percent level, we can conclude with confidence that New Zealand bettors discount a small proportion of their losses as 'not typical'.

6.4 Testing for Skewness

Most economic and finance models assume that investors are risk averse or risk neutral. In order to test economic and financial models on racetrack bettors, we require these investors to also be risk averse. A risk averse investor considers a higher expected return as desirable but dislikes the variance (risk). The third moment, skewness, indicates the extent to which the probability distribution is skewed to the left (negative skewness) or to the right (positive skewness). If a distribution has a right-tailed skew, then the chance of getting a positive return is greater than if the distribution is symmetrical. Hence, risk averse bettors also enjoy positive skewness.

The New Zealand data showed that expected returns were negative and decreased as the objective probability increased. Given that the objective probability can be thought of as a measure of risk, there is a negative trade-off between the risk and return in the New Zealand racing data, whereas the basis of finance models is centred around the risk and return having a positive relationship. In a finance market, an investor is compensated for withstanding risk however, in a betting market, the investor is punished.

The first utility estimation is presented in Tables 5.23-5.24. This power utility function developed by Ali (1977) and repeated by Golec & Tamarkin (1998) indicated a bettor's risk preference through the size of β. It was found that using the New Zealand data, this β value was significantly greater than 1 at the 5 percent level (except the grouping by rank of favouritism which was at the 10 percent level) at both the opening and closing odds. This implied $u'>0$, $u''>0$, and $u'''<0$, which indicated that bettors prefer mean and variance but dislike skewness and hence could be considered risk lovers. For this reason, it is possible to conclude that, with this method, New Zealand bettors have a risk loving preference, both at opening and closing odds.

Golec & Tamarkin (1998) developed the utility estimation of bettors further by approximating the utility function by a Taylor series expansion and truncating it to three terms. This way it was possible to estimate a coefficient for each of the three moments separately. So how do the three moments move in the New Zealand data? Tables 5.22 and C.1-C.7 show that as the objective probability decreased and we move from favourites to longshots, expected returns fell (risk loving attribute), variances increased (risk loving attribute) and the skewness increased (risk averse attribute). It is hoped that the size of the positive skewness will be large enough to offset the mean and variance movements which exhibit risk loving attributes. If this is so, then the assumption that bettors are risk averse may still hold. The results of Golec & Tamarkin's (1998) cubic utility model for the New Zealand data is presented in Tables 5.25-5.26. This utility model showed that at both opening and closing odds, bettors preferred the mean, since all 8 mean coefficients were positive, with 6 being significant at the 1 percent level. It also seems plausible to conclude that bettors prefer variance, which is a risk loving attribute. This is because 6 of the 8 variance coefficients were positive with 3 of these being significant. There seemed to be a preference for positive skewness at opening odds, since all 4 coefficients were positive and 3 of these were significant. However, at closing odds, few conclusions could be drawn as to the preference of skewness, since 2 methods produced positive coefficients while 2 methods produced negative coefficients.

It is believed that it is impossible to label bettors, whether they enter the market before the 30 minute mark or after, as risk averse or risk lovers with this method of utility estimation. Perhaps the results do show that bettors who enter the market up until the 30 minute mark are more risk averse than those bettors that enter after the 30 minute mark.

7. General Conclusions

The calculation of expected returns for each category over all 8 grouping methods showed that the New Zealand horse racing industry was not weak form efficient. Of the betting strategies tested in this thesis, it was found that the expected returns were all negative, apart from two outliers that were slightly, but not significantly, positive. It was found that, unlike the standard positive risk-return trade-off commonly found in the financial market, the racing industry exhibited a negative risk-return trade-off.

The subjective probabilities were calculated as an average of the proportion of the total money bet on a horse. Objective probabilities were calculated as the proportion of the total winners per category. The comparison of these found that at opening odds, horses with a high objective probability of winning were underbet, while horses with a low objective probability of winning were overbet, exhibiting the phenomenon known as the favourite longshot bias. After the 30 minute mark, the inclusion of more experienced bettors tended to eliminate this bias. When the tote closed, it was difficult to draw any conclusions as to the existence of the favourite longshot bias.

Two separate methods of examining how the favourite longshot bias changed over the raceday were adopted. The first method of splitting the data into races and then grouping by rank of favouritism tended to segregate the data too much and no concrete conclusions could be made as to how the favourite longshot bias changed over time. The second method of splitting the data into two groups, races 1-7 and races 8-10, resulted in providing evidence that the favourite longshot bias became more prevalent for later races.

The empirical (q_k) and subjective (Q_k) lose probabilities were calculated for all categories in each group. The estimation of the equation $Q_k=fq_k$ allowed the estimation of f, where (1-f) is the proportion of losses bettors discount as 'not typical'. It was found in the New Zealand data that punters tended to discount approximately 2% of their losses. This 2% proved to explain a great deal of the deviations between the actual and acceptable odds for longshots, and small deviations for favourites.

The estimation of Ali's (1977) power utility function showed conclusively that $u'>0$, $u''>0$, and $u'''<0$ for both the opening and closing odds hence, using this method, we could conclude that bettors are risk lovers. By taking the utility function, approximating it by a Taylor series expansion and truncating it to three moments as Golec & Tamarkin (1998) did, it proved that with this method, it was not possible to label bettors as risk averse or risk lovers. The only conclusion that could be made was that bettors entering the market up until the 30 minute mark seemed to be more risk averse than those entering after the 30 minute mark.

8. Areas of Further Research

The testing of weak form efficiency was performed with four simple betting strategies. These could be extended to include more elaborate strategies such as, bet on a horse if its odds fall by more than 20% in the final 30 minutes, or bet on a horse to win if it has the lowest odds for a place bet. As well as these, several other strategies could be examined using the place pool.

One of the most striking results in this thesis was how bettors in the final 30 minutes of the tote being open were able to eliminate what was very strong evidence of the favourite longshot bias at the opening win odds. It was tentatively concluded that this could be due to the fact that there may be a significant proportion of insider traders who enter the market late and, because they can more accurately predict the horses objective probability, they eliminate the favourite longshot bias. A formal study on the presence of insider traders in the New Zealand racing industry would no doubt reveal some interesting findings.

As for the favourite longshot bias, this could be adapted to include the place odds to answer questions such as 'is the favourite longshot bias more prevalent in win odds or place odds?', or 'how does the favourite longshot bias change throughout the betting period when looking at place odds?'. Also further tests could be done on examining how the favourite longshot bias changes throughout the raceday. Although there was not sufficient evidence to conclude that the favourite longshot bias became more prevalent with the win odds across both methods adopted, it may show a different story with the place odds.

The utility estimation section could be extended to include different types of utility functions. Different utility functions may explain more of the variation in the data than the types estimated in this thesis.

References

Ali, M.M. (1977). Probability and Utility Estimates for Racetrack Bettors. Journal of Political Economy, 85 803-815.

Ali, M.M. (1979). Some Evidence of the Efficiency of a Speculative Market. Econometrica, 47 387-392.

Arditti, F.D. (1967). Risk and the Required Return on Equity. The Journal of Finance, 22 19-36.

Arditti, F.D., & Levy, H. (1975). Portfolio Efficiency Analysis in Three Moments: The Multiperiod Case. The Journal of Finance, 30 797-809.

Asch, P., & Quandt, R.E. (1986). Racetrack Betting: The Professors' Guide to Strategies. Dover, Massachusetts: Auburn House Publishing Company.

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

Asch, P., Malkiel, B.G., & Quandt, R.E. (1984). Market Efficiency in Racetrack Betting. Journal of Business, 57 165-175.

Bird, R., McCrae, M., & Beggs, J. (1987). Are Gamblers really Risk Takers? Australian Economic Papers, 26 237-253.

Busche, K., & Hall, C.D. (1988). An Exception to the Risk Preference Anomaly. Journal of Business, 61 337-346.

Dowie, J. (1976). On the Efficiency and Equity of Betting Markets. Economica, 43 139-150.

Fabricant, B.F. (1965). Horse Sense. New York: David McKay Co.

Fama, F.E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25 383-423.

Figlewski, S. (1979). Subjective Information and Market Efficiency in a Betting Market. Journal of Political Economy, 87 75-88.

Friedman, M. & Savage, L.J. (1948). The Utility Analysis of Choices Involving Risk. Journal of Political Economy, 56, 279-304.

Gandar, J.M., Zuber, R., O'Brien, T., & Russo, B. (1998). Testing Rationality in the Point Spread Betting Market. The Journal of Finance, 43 995-1008.

Gandar, J.M., Zuber, R.A., & Russo, B. (1993). Testing Efficiency in Gambling Markets: A Comment. Applied Economics, 25 937-943.

Golec, J., & Tamarkin, M. (1998). Bettors Love Skewness, Not Risk, at the Horse Track. Journal of Political Economy, 106 205-225.

Griffith, R.M. (1949). Odds Adjustments by American Horse Race Bettors. American Journal of Psychology, 62 290-294.

Gruen, A. (1976). An Inquiry into the Economics of Race-Track Gambling. Journal of Political Economy, 84 169-177.

Hausch, D.B., & Ziemba, W.T. (1985). Transaction Costs, Extent of Inefficiencies, Entries and Multiple Wages in a Racetrack Betting Model. Management Science, 31 381-394.

Hausch, D.B., Ziemba, W.T., & Rubinstein, M. (1981). Efficiency of the Market for Racetrack Betting. Management Science, 27 1435-1452.

Henery, R.J. (1984). An Extreme-Value Model for Predicting the Results of Horse Races. Applied Statistics, 33 125-133.

Henery, R.J. (1985). On the Average Probability of Losing Bets on Horses with given Starting Price Odds. Royal Statistical Society, 148 342-349.

Jullien, B., & Salanie, B. (1994). Measuring the Incidence of Insider Trading: A Comment on Shin. The Economic Journal, 104 1418-1419.

Kanto, A.J., Rosenquist, G., & Suvas, A. (1992). On Utility Estimation of Racetrack Bettors. Journal of Economic Psychology, 13 491-498.

Loistl, O. (1976). The Erroneous Approximation of Expected Utility by means of a Taylor's Series Expansion: Analytical and Computational Results. American Economic Review, 66 904-910.

McEnally, R.W. (1974). A Note on the Return Behavior of High Risk Common Stocks. The Journal of Finance, 29 199-202.

McGlothin, W.H. (1956). Stability of Choices Among Uncertain Alternatives. American Journal of Psychology, 69 604-615.

Preston, M.G., & Baratta, P. (1948). An Experimental Study of the Auction-Value of an Uncertain Outcome. American Journal of Psychology, 61 183-193.

Quandt, R.E. (1986). Betting and Equilibrium. The Quarterly Journal of Economics, 101 201-207.

Rosett, R.H. (1965). Gambling and Rationality. Journal of Political Economics, 75 595-607.

Runyon, D. (1976). An Inquiry into the Economics of Race-Track Gambling. Journal of Political Economy, 84 169-177.

Russo, B., Gandar, J.M., & Zuber, R.A. (1989). Market Rationality Tests Based on Cross-Equation Restrictions. Journal of Monetary Economics, 23 455-470

Seligman, D. (1975). A Thinking Mans Guide to Losing at the Track. Fortune, 92 81-87.

Sharpe, U.F. (1970). Portfolio Theory and Capital Markets. New York: McGraw-Hill Book Co.

Shaw, D.C. (1971). The Performance of Primary Common Stock Offerings: A Canadian Comparison. The Journal of Finance, 26 1101-1113.

Shin, H.S. (1991). Optimal Betting Odds Against Insider Traders. The Economic Journal, 101 1179-1185.

Shin, H.S. (1992). Prices of State Contingent Claims with Insider Traders, and the Favorite-Longshot Bias. The Economic Journal, 102 426-435.

Shin, H.S. (1993). Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims. The Economic Journal, 103 1141-1153.

Snyder, W.W. (1978). Horse Racing: Testing the Efficient Markets Model. The Journal of Finance, 33 1109-1118.

Swindler, S., & Shaw, R. (1995). Racetrack Wagering and the "Uninformed" Bettor: A Study of Market Efficiency. The Quarterly Review of Economics and Finance, 35 305-314.

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

Turnball, P. (1997). Testing for Market Efficiency in the New Zealand Pari-mutuel Racetrack Betting Market. Unpublished honours report, Massey University, Palmerston North, New Zealand.

Weitzman, M. (1965). Utility Analysis and Group Behavior: An Empirical Study. Journal of Political Economics, 73 18-26.

Williams, L.V., & Paton, D. (1997). Why is there a Favorite-Longshot Bias in British Racetrack Betting Markets. The Economic Journal, 106 150-158.

Zuber, R.A., Gandar, J.M., & Bowers, B.D. (1995). Beating the spread: Testing the Efficiency of the Gambling Market for the National Football League Games. Journal of Political Economy, 93 800-806.

Appendices

Appendix A: Favourite Longshot Bias Analysis on Races 2-10

Table A.1: Race 2: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.14	1050	285	0.27143	0.28567	-1.03767
2	$5.16	1050	184	0.17524	0.16964	0.47751
3	$6.87	1050	126	0.12000	0.12626	-0.62446
4	$8.88	1050	123	0.11714	0.09738	1.99165**
5	$11.23	1047	91	0.08691	0.07753	1.07804
6	$14.06	1037	64	0.06172	0.06301	-0.17294
7	$17.72	1015	58	0.05714	0.05078	0.87298
8	$21.97	966	38	0.03934	0.04172	-0.38049
9	$26.79	896	34	0.03795	0.03470	0.50808
10	$32.93	801	22	0.02747	0.02851	-0.18067
11	$38.90	688	6	0.00872	0.02406	-4.32729***
12	$46.15	561	8	0.01426	0.02035	-1.21611
13	$53.65	445	6	0.01348	0.01743	-0.72256
14	$61.69	334	6	0.01796	0.01512	0.39185
15	$67.81	170	1	0.00588	0.01355	-1.30805
16	$73.87	93	1	0.01075	0.01241	-0.15482
17/18	$77.75	50	0	0.00000	0.01177	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.2: Race 3: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.11	1045	304	0.29091	0.28904	0.13322
2	$5.11	1045	172	0.16459	0.17098	-0.55637
3	$6.82	1045	117	0.11196	0.12693	-1.53455
4	$8.69	1044	116	0.11111	0.09934	1.21039
5	$10.91	1040	94	0.09038	0.07969	1.20293
6	$14.08	1030	64	0.06214	0.06301	-0.11646
7	$17.92	997	62	0.06219	0.05044	1.53547
8	$22.31	948	43	0.04536	0.04114	0.62445
9	$27.60	868	28	0.03226	0.03374	-0.24645
10	$33.34	753	20	0.02656	0.02817	-0.27459
11	$38.94	620	8	0.01290	0.02407	-2.46484**
12	$45.67	511	5	0.00978	0.02055	-2.47122**
13	$52.11	385	4	0.01039	0.01784	-1.44209
14	$61.48	281	6	0.02135	0.01525	0.70812
15	$67.02	157	1	0.00637	0.01377	-1.16620
16	$72.69	89	2	0.02247	0.01246	0.63756
17/18	$76.34	44	1	0.02273	0.01186	0.48380

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.3: Race 4: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking (1)	Average Odds (2)	N_k (3)	W_k (4)	OP_k (5)	SP_k (6)	z_k (7)
1	$3.13	1050	312	0.29714	0.28542	0.83108
2	$5.13	1050	162	0.15429	0.16946	-1.36122
3	$6.82	1050	137	0.13048	0.12691	0.34269
4	$8.63	1049	96	0.09152	0.10034	-0.99148
5	$10.91	1048	97	0.09256	0.07988	1.41584
6	$13.78	1039	66	0.06352	0.06389	-0.04904
7	$17.57	1020	61	0.05980	0.05156	1.11093
8	$21.70	960	42	0.04375	0.04182	0.29218
9	$27.19	865	31	0.03584	0.03406	0.28189
10	$32.72	756	15	0.01984	0.02831	-1.67037*
11	$39.19	638	15	0.02351	0.02368	-0.02820
12	$45.24	495	10	0.02020	0.02047	-0.04206
13	$52.06	374	7	0.01872	0.01782	0.12816
14	$58.94	275	2	0.00727	0.01564	-1.63276
15	$64.90	164	1	0.00610	0.01413	-1.32112
16	$74.10	97	1	0.01031	0.01220	-0.18423
17/18	$76.50	38	0	0.00000	0.01182	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.4: Race 5: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.12	1043	284	0.27229	0.28570	-0.97264
2	$5.11	1043	199	0.19080	0.17059	1.66075*
3	$6.77	1043	136	0.13039	0.12791	0.23786
4	$8.67	1043	108	0.10355	0.09982	0.39550
5	$10.97	1037	76	0.07329	0.07970	-0.79207
6	$13.68	1030	61	0.05922	0.06433	-0.69414
7	$17.34	1002	46	0.04591	0.05161	-0.86171
8	$21.73	941	49	0.05207	0.04183	1.41352
9	$27.20	864	30	0.03472	0.03406	0.10684
10	$32.81	733	21	0.02865	0.02842	0.03715
11	$39.44	609	14	0.02299	0.02394	-0.15669
12	$45.03	471	10	0.02123	0.02071	0.07794
13	$51.52	362	5	0.01381	0.01806	-0.69255
14	$59.17	272	4	0.01471	0.01572	-0.13953
15	$63.10	161	1	0.00621	0.01447	-1.33377
16	$72.19	100	0	0.00000	0.01239	-
17/18	$79.93	52	1	0.01923	0.01132	0.41557

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.5: Race 6: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.17	1042	275	0.26392	0.28043	-1.20965
2	$5.14	1042	169	0.16219	0.16981	-0.66711
3	$6.76	1042	148	0.14203	0.12787	1.30993
4	$8.54	1042	118	0.11324	0.10148	1.19869
5	$10.62	1038	84	0.08092	0.08163	-0.08386
6	$13.47	1029	77	0.07483	0.06554	1.13240
7	$17.14	1001	55	0.05495	0.05257	0.33029
8	$21.68	940	37	0.03936	0.04240	-0.47977
9	$26.48	855	24	0.02807	0.03483	-1.19711
10	$31.52	718	24	0.03343	0.02930	0.61516
11	$38.10	595	9	0.01513	0.02445	-1.86335*
12	$44.04	473	13	0.02748	0.02099	0.86345
13	$50.81	354	5	0.01412	0.01793	-0.60701
14	$59.33	272	3	0.01103	0.01541	-0.69193
15	$65.06	148	2	0.01351	0.01389	-0.03963
16	$77.88	86	1	0.01163	0.01158	0.00417
17/18	$85.88	32	0	0.00000	0.01027	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.6: Race 7: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.26	1042	271	0.26008	0.27519	-1.11211
2	$5.19	1042	172	0.16507	0.16739	-0.20217
3	$6.78	1042	139	0.13340	0.12731	0.57766
4	$8.52	1042	110	0.10557	0.10091	0.48869
5	$10.62	1039	98	0.09432	0.08147	1.41775
6	$13.23	1029	77	0.07483	0.06630	1.04037
7	$16.48	1006	50	0.04970	0.05394	-0.61857
8	$20.37	951	44	0.04627	0.04451	0.25845
9	$25.48	863	31	0.03592	0.03631	-0.06086
10	$30.94	732	16	0.02186	0.03007	-1.51996
11	$36.35	585	14	0.02393	0.02567	-0.27442
12	$42.17	464	8	0.01724	0.02206	-0.79717
13	$49.21	354	5	0.01412	0.01905	-0.78610
14	$57.59	255	4	0.01569	0.01636	-0.08662
15	$63.40	154	1	0.00649	0.01456	-1.24706
16	$70.24	101	2	0.01980	0.01296	0.49352
17/18	$83.06	46	1	0.02174	0.01087	0.50536

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.7: Race 8: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.21	1042	291	0.27927	0.27820	0.07737
2	$5.18	1042	183	0.17562	0.16746	0.69295
3	$6.75	1042	137	0.13148	0.12787	0.34502
4	$8.50	1042	128	0.12284	0.10169	2.08039**
5	$10.62	1041	82	0.07877	0.08182	-0.36507
6	$13.23	1038	58	0.05588	0.06639	-1.47408
7	$16.61	1011	56	0.05539	0.05366	0.24078
8	$20.72	952	36	0.03782	0.04388	-0.98137
9	$25.50	831	29	0.03490	0.03629	-0.21830
10	$31.03	692	19	0.02746	0.03010	-0.42484
11	$37.12	576	15	0.02604	0.02521	0.12543
12	$42.54	438	7	0.01598	0.02169	-0.95223
13	$48.80	323	0	0.00000	0.01900	-
14	$56.82	238	3	0.01261	0.01598	-0.46738
15	$64.12	146	3	0.02055	0.01417	0.54317
16	$69.50	82	0	0.00000	0.01279	-
17/18	$84.81	51	0	0.00000	0.01059	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.8: Race 9: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking	Average Odds	N_k	W_k	OP_k	SP_k	z_k
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	$3.24	1021	267	0.26151	0.27589	-1.04589
2	$5.11	1021	173	0.16944	0.16956	-0.01019
3	$6.62	1021	139	0.13614	0.13004	0.56803
4	$8.38	1021	103	0.10088	0.10306	-0.23121
5	$10.56	1018	92	0.09037	0.08203	0.92850
6	$13.20	1009	60	0.05946	0.06641	-0.93307
7	$16.83	978	58	0.05930	0.05359	0.75647
8	$21.15	917	49	0.05344	0.04305	1.39851
9	$26.24	806	27	0.03350	0.03564	-0.33802
10	$31.24	676	23	0.03402	0.03005	0.56950
11	$36.47	542	13	0.02399	0.02557	-0.24133
12	$42.73	438	9	0.02055	0.02183	-0.18946
13	$49.89	323	3	0.00929	0.01904	-1.82655
14	$58.58	237	5	0.02110	0.01632	0.51178
15	$63.95	132	1	0.00758	0.01461	-0.93174
16	$70.54	71	2	0.02817	0.01299	0.77320
17/18	$82.34	41	0	0.00000	0.01112	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Table A.9: Race 10: Favourite Longshot Bias Analysis by Rank of Favouritism using Closing Win Odds

Favourite Ranking (1)	Average Odds (2)	N_k (3)	W_k (4)	OP_k (5)	SP_k (6)	z_k (7)
1	$3.48	975	251	0.25744	0.25441	0.21587
2	$5.42	975	156	0.16000	0.15926	0.06310
3	$6.95	975	123	0.12615	0.12379	0.22231
4	$8.59	975	100	0.10256	0.09976	0.28837
5	$10.71	974	75	0.07700	0.08009	-0.36181
6	$12.91	966	73	0.07557	0.06674	1.03784
7	$15.82	956	64	0.06695	0.05525	1.44726
8	$19.14	928	42	0.04526	0.04601	-0.11050
9	$23.30	882	31	0.03515	0.03839	-0.52314
10	$28.14	807	21	0.02602	0.03219	-1.09986
11	$33.98	710	13	0.01831	0.02695	-1.71754*
12	$39.58	605	14	0.02314	0.02315	-0.00196
13	$46.45	503	5	0.00994	0.01971	-2.20894**
14	$55.70	391	5	0.01279	0.01677	-0.70008
15	$62.69	236	4	0.01695	0.01490	0.24345
16	$70.11	143	2	0.01399	0.01304	0.09615
17/18	$73.67	70	0	0.00000	0.01250	-

*** significant at 1 percent level
** significant at 5 percent level
* significant at 10 percent level

Appendix B: Henery's Test

Table B.1: Henery's Test by 25 Even Sized Subjective Probability Categories using Opening Win Odds

Lower-Upper	Average Odds	N_k	W_k	q_k	Q_k
(1)	(2)	(3)	(4)	(5)	(6)
0-0.016754	$67.99	4725	36	0.99238	0.98529
0.016754-0.021614	$43.07	4724	64	0.98645	0.97678
0.021614-0.025826	$34.91	4724	77	0.98370	0.97136
0.025826-0.029891	$29.71	4724	97	0.97947	0.96635
0.029891-0.033777	$26.04	4724	124	0.97375	0.96159
0.033777-0.037677	$23.18	4724	123	0.97396	0.95686
0.037677-0.041647	$20.90	4724	128	0.97290	0.95215
0.041647-0.045719	$18.97	4724	173	0.96338	0.94727
0.045719-0.04999	$17.30	4724	193	0.95914	0.94220
0.04999-0.054378	$15.87	4724	225	0.95237	0.93700
0.054378-0.05909	$14.63	4724	277	0.94136	0.93164
0.05909-0.063975	$13.48	4724	284	0.93988	0.92579
0.063975-0.069294	$12.45	4724	319	0.93247	0.91967
0.069294-0.075138	$11.48	4724	368	0.92210	0.91289
0.075138-0.081437	$10.59	4724	387	0.91808	0.90558
0.081437-0.088176	$9.79	4724	419	0.91130	0.89783
0.088176-0.095942	$9.01	4724	398	0.91575	0.88900
0.095942-0.104701	$8.27	4724	489	0.89649	0.87910
0.104701-0.114484	$7.57	4724	540	0.88569	0.86798
0.114484-0.126345	$6.89	4724	604	0.87214	0.85495
0.126345-0.142006	$6.19	4724	608	0.87130	0.83849
0.142006-0.162878	$5.47	4724	798	0.83108	0.81708
0.162878-0.192736	$4.70	4724	887	0.81224	0.78740
0.192736-0.248828	$3.84	4724	1155	0.75550	0.73926
0.248828-1	$2.59	4724	1617	0.65771	0.61372

Table B.2: Henery's Test by 50 Even Sized Subjective Probability Categories using Opening Win Odds

Lower-Upper (1)	Average Odds (2)	N_k (3)	W_k (4)	q_k (5)	Q_k (6)
0-0.013394	$81.20	2363	12	0.99492	0.98768
0.013394-0.016754	$54.78	2362	24	0.98984	0.98175
0.016754-0.01939	$45.80	2362	29	0.98772	0.97816
0.01939-0.021614	$40.34	2362	35	0.98518	0.97521
0.021614-0.023815	$36.47	2362	35	0.98518	0.97258
0.023815-0.025826	$33.35	2362	42	0.98222	0.97002
0.025826-0.027908	$30.79	2362	46	0.98052	0.96752
0.027908-0.029891	$28.64	2362	51	0.97841	0.96509
0.029891-0.031847	$26.82	2362	68	0.97121	0.96271
0.031847-0.033777	$25.26	2362	56	0.97629	0.96041
0.033777-0.035756	$23.83	2362	64	0.97290	0.95804
0.035756-0.037677	$22.53	2362	59	0.97502	0.95562
0.037677-0.039649	$21.42	2362	49	0.97925	0.95332
0.039649-0.041647	$20.38	2362	79	0.96655	0.95092
0.041647-0.043723	$19.40	2362	82	0.96528	0.94846
0.043723-0.045719	$18.53	2362	91	0.96147	0.94603
0.045719-0.04792	$17.68	2362	92	0.96105	0.94345
0.04792-0.04999	$16.92	2362	101	0.95724	0.94090
0.04999-0.052186	$16.20	2362	105	0.95555	0.93828
0.052186-0.054378	$15.54	2362	120	0.94920	0.93567
0.054378-0.056686	$14.93	2362	123	0.94793	0.93300
0.056686-0.05909	$14.33	2362	154	0.93480	0.93021
0.05909-0.061499	$13.74	2362	138	0.94157	0.92724
0.061499-0.063975	$13.21	2362	146	0.93819	0.92427
0.063975-0.066558	$12.70	2362	168	0.92887	0.92124
0.066558-0.069294	$12.20	2362	151	0.93607	0.91803
0.069294-0.072212	$11.72	2362	174	0.92633	0.91468
0.072212-0.075138	$11.24	2362	194	0.91787	0.91102
0.075138-0.07814	$10.80	2362	188	0.92041	0.90742
0.07814-0.081437	$10.38	2362	199	0.91575	0.90366
0.081437-0.084693	$9.98	2362	194	0.91787	0.89983
0.084693-0.088176	$9.59	2362	225	0.90474	0.89574

0.088176-0.091907	$9.20	2362	199	0.91575	0.89125
0.091907-0.095942	$8.82	2362	199	0.91575	0.88664
0.095942-0.100064	$8.45	2362	237	0.89966	0.88172
0.100064-0.104701	$8.09	2362	252	0.89331	0.87636
0.104701-0.109343	$7.74	2362	261	0.88950	0.87086
0.109343-0.114484	$7.41	2362	279	0.88188	0.86498
0.114484-0.119978	$7.07	2362	317	0.86579	0.85856
0.119978-0.126345	$6.72	2362	287	0.87849	0.85116
0.126345-0.133768	$6.37	2362	295	0.87511	0.84302
0.133768-0.142006	$6.01	2362	313	0.86749	0.83370
0.142006-0.151341	$5.66	2362	390	0.83489	0.82326
0.151341-0.162878	$5.28	2362	408	0.82727	0.81044
0.162878-0.175907	$4.91	2362	426	0.81964	0.79616
0.175907-0.192736	$4.50	2362	461	0.80483	0.77785
0.192736-0.215233	$4.07	2362	584	0.75275	0.75433
0.215233-0.248828	$3.60	2362	571	0.75826	0.72224
0.248828-0.310036	$3.02	2362	675	0.71423	0.66898
0.310036-1	$2.16	2362	942	0.60119	0.53636

Table B.3: Henery's Test by Even Subjective Probability Categories using Opening Win Odds

Lower-Upper (1)	Average Odds (2)	N_k (3)	W_k (4)	q_k (5)	Q_k (6)
0-0.0125	$86.15	1836	8	0.99564	0.98839
0.0125-0.025	$43.97	11370	155	0.98637	0.97726
0.025-0.0375	$26.78	14927	356	0.97615	0.96266
0.0375-0.05	$19.10	14401	497	0.96549	0.94763
0.05-0.0625	$14.83	12744	689	0.94594	0.93256
0.0625-0.075	$12.12	10784	781	0.92758	0.91748
0.075-0.0875	$10.23	9145	774	0.91536	0.90221
0.0875-0.1	$8.86	7436	668	0.91017	0.88710
0.1-0.1125	$7.81	6227	700	0.88759	0.87198
0.1125-0.125	$7.00	5058	623	0.87683	0.85712
0.125-0.1375	$6.33	3992	499	0.87500	0.84210
0.1375-0.15	$5.78	3356	525	0.84356	0.82688
0.15-0.1625	$5.31	2541	427	0.83196	0.81151
0.1625-0.175	$4.93	2334	425	0.81791	0.79723
0.175-0.1875	$4.58	1770	342	0.80678	0.78147
0.1875-0.2	$4.27	1588	339	0.78652	0.76606
0.2-0.225	$3.91	2363	598	0.74693	0.74413
0.225-0.25	$3.50	1569	388	0.75271	0.71434
0.25-0.275	$3.15	1228	342	0.72150	0.68279
0.275-0.3	$2.88	877	249	0.71608	0.65271
0.3-0.325	$2.65	593	184	0.68971	0.62197
0.325-0.35	$2.46	484	160	0.66942	0.59391
0.35-0.375	$2.28	353	140	0.60340	0.56182
0.375-0.4	$2.14	268	97	0.63806	0.53302
0.4-0.45	$1.95	365	161	0.55890	0.48829
0.45-0.5	$1.75	195	96	0.50769	0.42740
0.5-0.55	$1.58	124	55	0.55645	0.36605
0.55-0.6	$1.45	78	43	0.44872	0.30882
0.6-1	$1.23	95	69	0.27368	0.18420

Table B.4: Henery's Test by Rank of Favouritism using Closing Win Odds

Ranking (1)	Average Odds (2)	N_k (3)	W_k (4)	q_k (5)	Q_k (6)
1	$3.19	10361	2852	0.72474	0.68691
2	$5.15	10361	1763	0.82984	0.80599
3	$6.77	10361	1349	0.86980	0.85240
4	$8.60	10359	1108	0.89304	0.88370
5	$10.82	10333	871	0.91571	0.90755
6	$13.56	10246	654	0.93617	0.92623
7	$17.15	10006	549	0.94513	0.94169
8	$21.33	9448	417	0.95586	0.95312
9	$26.39	8585	296	0.96552	0.96211
10	$31.86	7394	190	0.97430	0.96862
11	$37.81	6151	120	0.98049	0.97355
12	$43.94	4920	101	0.97947	0.97724
13	$50.59	3767	47	0.98752	0.98023
14	$58.80	2809	41	0.98540	0.98299
15	$64.84	1601	18	0.98876	0.98458
16	$72.58	942	11	0.98832	0.98622
17/18	$80.22	457	3	0.99344	0.98753

Table B.5: Henery's Test by 25 Even Sized Subjective Probability Categories using Closing Win Odds

Lower-Upper	Average Odds	N_k	W_k	q_k	Q_k
(1)	(2)	(3)	(4)	(5)	(6)
0-0.01311	$81.92	4725	26	0.99450	0.98779
0.013111-0.016763	$55.30	4724	41	0.99132	0.98192
0.016764-0.020143	$44.82	4724	80	0.98307	0.97769
0.020144-0.023511	$37.91	4724	89	0.98116	0.97362
0.023512-0.027174	$32.64	4724	100	0.97883	0.96936
0.027175-0.030911	$28.49	4724	103	0.97820	0.96490
0.030912-0.034839	$25.15	4724	144	0.96952	0.96024
0.034840-0.039012	$22.41	4724	181	0.96169	0.95538
0.039013-0.04345	$20.03	4724	227	0.95195	0.95007
0.043451-0.048241	$18.02	4724	235	0.95025	0.94451
0.048242-0.053391	$16.26	4724	226	0.95216	0.93850
0.053392-0.058838	$14.72	4724	276	0.94157	0.93207
0.058839-0.064795	$13.38	4724	282	0.94030	0.92526
0.064796-0.071462	$12.14	4724	355	0.92485	0.91763
0.071463-0.078806	$11.00	4724	372	0.92125	0.90909
0.078807-0.086929	$9.97	4724	412	0.91279	0.89970
0.086930-0.095917	$9.05	4724	444	0.90601	0.88950
0.095918-0.10624	$8.19	4724	509	0.89225	0.87790
0.106241-0.118228	$7.37	4724	518	0.89035	0.86431
0.118229-0.132756	$6.60	4724	580	0.87722	0.84848
0.132757-0.151227	$5.84	4724	693	0.85330	0.82877
0.151228-0.176891	$5.07	4724	779	0.83510	0.80276
0.176892-0.215018	$4.26	4724	952	0.79848	0.76526
0.215019-0.275531	$3.42	4724	1132	0.76037	0.70760
0.275532-1	$2.42	4724	1661	0.64839	0.58678

Table B.6: Henery's Test by 50 Even Sized Subjective Probability Categories using Closing Win Odds

Lower-Upper (1)	Average Odds (2)	N_k (3)	W_k (4)	q_k (5)	Q_k (6)
0-0.010795	$94.69	2363	11	0.99534	0.98944
0.010795-0.01311	$69.16	2362	15	0.99365	0.98554
0.01311-0.01503	$58.69	2362	19	0.99196	0.98296
0.01503-0.016763	$51.91	2362	22	0.99069	0.98073
0.016763-0.018482	$46.90	2362	34	0.98561	0.97868
0.018482-0.020143	$42.74	2362	46	0.98052	0.97660
0.020143-0.021798	$39.38	2362	45	0.98095	0.97461
0.021798-0.023511	$36.44	2362	44	0.98137	0.97256
0.023511-0.025332	$33.83	2362	43	0.98180	0.97044
0.025332-0.027174	$31.46	2362	57	0.97587	0.96821
0.027174-0.028985	$29.42	2362	58	0.97544	0.96601
0.028985-0.030911	$27.57	2362	45	0.98095	0.96373
0.030911-0.032848	$25.94	2362	76	0.96782	0.96144
0.032848-0.034839	$24.36	2362	68	0.97121	0.95896
0.034839-0.036845	$23.03	2362	73	0.96909	0.95658
0.036845-0.039012	$21.78	2362	108	0.95428	0.95410
0.039012-0.041227	$20.56	2362	108	0.95428	0.95137
0.041227-0.04345	$19.50	2362	119	0.94962	0.94872
0.04345-0.045876	$18.48	2362	113	0.95216	0.94589
0.045876-0.048241	$17.56	2362	122	0.94835	0.94305
0.048241-0.050845	$16.67	2362	110	0.95343	0.94003
0.050845-0.053391	$15.84	2362	116	0.95089	0.93689
0.053391-0.056033	$15.08	2362	137	0.94200	0.93367
0.056033-0.058838	$14.37	2362	139	0.94115	0.93040
0.058838-0.061699	$13.70	2362	148	0.93734	0.92702
0.061699-0.064795	$13.06	2362	134	0.94327	0.92340
0.064795-0.068008	$12.43	2362	171	0.92760	0.91958
0.068008-0.071462	$11.85	2362	184	0.92210	0.91561
0.071462-0.075083	$11.27	2362	188	0.92041	0.91129
0.075083-0.078806	$10.73	2362	184	0.92210	0.90683
0.078806-0.082783	$10.22	2362	199	0.91575	0.90212
0.082783-0.086929	$9.73	2362	213	0.90982	0.89725

0.086929-0.091264	$9.28	2362	214	0.90940	0.89223
0.091264-0.095917	$8.83	2362	230	0.90262	0.88670
0.095917-0.100908	$8.40	2362	247	0.89543	0.88089
0.100908-0.10624	$7.98	2362	262	0.88908	0.87466
0.10624-0.111957	$7.58	2362	274	0.88400	0.86802
0.111957-0.118228	$7.17	2362	244	0.89670	0.86056
0.118228-0.12505	$6.80	2362	287	0.87849	0.85284
0.12505-0.132756	$6.41	2362	293	0.87595	0.84406
0.132756-0.141188	$6.04	2362	322	0.86367	0.83438
0.141188-0.151227	$5.64	2362	371	0.84293	0.82274
0.151227-0.162937	$5.27	2362	378	0.83997	0.81017
0.162937-0.176891	$4.87	2362	401	0.83023	0.79458
0.176891-0.194337	$4.47	2362	452	0.80864	0.77609
0.194337-0.215018	$4.05	2362	473	0.79975	0.75310
0.215018-0.240894	$3.64	2362	545	0.76926	0.72521
0.240894-0.275532	$3.21	2362	587	0.75148	0.68836
0.275532-0.329301	$2.75	2362	688	0.70872	0.63645
0.329301-1	$2.10	2362	973	0.58806	0.52337

Table B.7: Henery's Test by Even Subjective Probability Categories using Closing Win Odds

Lower-Upper (1)	Average Odds (2)	N_k (3)	W_k (4)	q_k (5)	Q_k (6)
0-0.0125	$84.70	4073	22	0.99460	0.98819
0.0125-0.025	$45.35	16765	248	0.98521	0.97795
0.025-0.0375	$26.89	15351	423	0.97244	0.96281
0.0375-0.05	$19.07	12647	607	0.95200	0.94756
0.05-0.0625	$14.79	10794	606	0.94386	0.93237
0.0625-0.075	$12.08	8825	644	0.92703	0.91724
0.075-0.0875	$10.20	7464	620	0.91693	0.90193
0.0875-0.1	$8.84	6366	643	0.89899	0.88686
0.1-0.1125	$7.79	5306	575	0.89163	0.87168
0.1125-0.125	$6.97	4508	514	0.88598	0.85655
0.125-0.1375	$6.31	3709	479	0.87085	0.84151
0.1375-0.15	$5.75	3127	465	0.85130	0.82617
0.15-0.1625	$5.30	2519	406	0.83882	0.81115
0.1625-0.175	$4.90	2186	370	0.83074	0.79610
0.175-0.1875	$4.56	1814	332	0.81698	0.78085
0.1875-0.2	$4.26	1642	324	0.80268	0.76535
0.2-0.225	$3.89	2630	559	0.78745	0.74269
0.225-0.25	$3.48	2008	471	0.76544	0.71249
0.25-0.275	$3.15	1613	415	0.74272	0.68212
0.275-0.3	$2.87	1326	391	0.70513	0.65111
0.3-0.325	$2.63	912	263	0.71162	0.62002
0.325-0.35	$2.45	689	214	0.68940	0.59103
0.35-0.375	$2.28	499	188	0.62325	0.56060
0.375-0.4	$2.13	370	149	0.59730	0.53001
0.4-0.45	$1.95	516	215	0.58333	0.48698
0.45-0.5	$1.75	259	131	0.49421	0.42983
0.5-0.55	$1.58	114	66	0.42105	0.36807
0.55-0.6	$1.44	47	33	0.29787	0.30319
0.6-1	$1.26	22	17	0.22727	0.20578

Appendix C: Mean, Variance and Skewness of Unit Bet Returns

Table C.1: Mean, Variance and Skewness of Unit Bet Returns by 25 Even Sized Subjective Probability Categories using Opening Win Odds

Lower-Upper	N_k	OP_k	Mean	Variance	Skewness
(1)	(2)	(3)	(4)	(5)	(6)
0-0.016754	4725	0.00762	-0.51062	31.1937	1973.067
0.016754-0.021614	4724	0.01497	-0.35601	27.28072	1138.104
0.021614-0.025826	4724	0.01802	-0.37055	21.59592	727.3444
0.025826-0.029891	4724	0.02270	-0.32441	19.65439	558.5107
0.029891-0.033777	4724	0.02901	-0.24322	19.16769	470.9704
0.033777-0.037677	4724	0.02878	-0.33123	15.09404	330.574
0.037677-0.041647	4724	0.02995	-0.37267	12.74718	251.0218
0.041647-0.045719	4724	0.04048	-0.23051	14.03627	245.2343
0.045719-0.04999	4724	0.04516	-0.21694	12.96587	204.5347
0.04999-0.054378	4724	0.05264	-0.16206	12.63549	179.9454
0.054378-0.05909	4724	0.06481	-0.05077	13.00154	165.7404
0.05909-0.063975	4724	0.06645	-0.10269	11.31205	132.4561
0.063975-0.069294	4724	0.07464	-0.06886	10.7494	114.0858
0.069294-0.075138	4724	0.08610	-0.00936	10.41643	99.20784
0.075138-0.081437	4724	0.09055	-0.0383	9.289208	80.79293
0.081437-0.088176	4724	0.09803	-0.0387	8.502083	67.02271
0.088176-0.095942	4724	0.09312	-0.15905	6.887117	50.61187
0.095942-0.104701	4724	0.11441	-0.05136	6.96555	44.53823
0.104701-0.114484	4724	0.12635	-0.04101	6.359346	36.07195
0.114484-0.126345	4724	0.14132	-0.0238	5.790333	28.693
0.126345-0.142006	4724	0.14226	-0.11797	4.690889	20.81001
0.142006-0.162878	4724	0.18671	0.02075	4.538536	15.54687
0.162878-0.192736	4724	0.20753	-0.02423	3.635709	9.998947
0.192736-0.248828	4724	0.27024	0.033282	2.883178	5.065823
0.248828-1	4724	0.37833	-0.06737	1.429232	0.857307

Table C.2: Mean, Variance and Skewness of Unit Bet Returns by 50 Even Sized Subjective Probability Categories using Opening Win Odds

Lower-Upper (1)	N_k (2)	OP_k (3)	Mean (4)	Variance (5)	Skewness (6)
0-0.013394	2363	0.00508	-0.60568	30.46286	2341.366
0.013394-0.016754	2362	0.01016	-0.44355	30.16361	1618.305
0.016754-0.01939	2362	0.01228	-0.43692	25.50707	1141.082
0.01939-0.021614	2362	0.01482	-0.40009	23.9275	940.0014
0.021614-0.023815	2362	0.01482	-0.45807	19.52601	692.9514
0.023815-0.025826	2362	0.01778	-0.40531	19.53557	630.1215
0.025826-0.027908	2362	0.01948	-0.39837	18.22389	541.0522
0.027908-0.029891	2362	0.02159	-0.37982	17.42856	478.9782
0.029891-0.031847	2362	0.02879	-0.22523	20.25025	513.5936
0.031847-0.033777	2362	0.02371	-0.39999	14.82479	357.3896
0.033777-0.035756	2362	0.02710	-0.35246	15.05593	340.3136
0.035756-0.037677	2362	0.02498	-0.43511	12.45559	267.6064
0.037677-0.039649	2362	0.02075	-0.55407	9.386507	193.3956
0.039649-0.041647	2362	0.03345	-0.31683	13.48767	257.0695
0.041647-0.043723	2362	0.03472	-0.32437	12.69234	229.8614
0.043723-0.045719	2362	0.03853	-0.28419	12.78716	219.2747
0.045719-0.04792	2362	0.03895	-0.30929	11.7714	192.4833
0.04792-0.04999	2362	0.04276	-0.27455	11.78124	182.7803
0.04999-0.052186	2362	0.04445	-0.27751	11.22026	166.1446
0.052186-0.054378	2362	0.05080	-0.20771	11.72801	164.3135
0.054378-0.056686	2362	0.05207	-0.22121	11.04053	147.918
0.056686-0.05909	2362	0.06520	-0.06443	12.5497	156.5995
0.05909-0.061499	2362	0.05843	-0.19515	10.43973	127.0109
0.061499-0.063975	2362	0.06181	-0.18147	10.16912	118.0141
0.063975-0.066558	2362	0.07113	-0.09503	10.69541	116.7246
0.066558-0.069294	2362	0.06393	-0.21773	8.960285	95.62404
0.069294-0.072212	2362	0.07367	-0.13483	9.412454	94.2574
0.072212-0.075138	2362	0.08213	-0.07384	9.585891	90.33694
0.075138-0.07814	2362	0.07959	-0.13711	8.610263	78.48653
0.07814-0.081437	2362	0.08425	-0.12301	8.359771	72.35662
0.081437-0.084693	2362	0.08213	-0.17831	7.545258	63.08524
0.084693-0.088176	2362	0.09526	-0.08413	7.966992	62.00632
0.088176-0.091907	2362	0.08425	-0.22269	6.567351	50.38161

0.091907-0.095942	2362	0.08425	-0.25497	6.033206	44.36178
0.095942-0.100064	2362	0.10034	-0.14958	6.484468	43.92987
0.100064-0.104701	2362	0.10669	-0.13421	6.276329	40.06478
0.104701-0.109343	2362	0.11050	-0.14201	5.925839	35.84341
0.109343-0.114484	2362	0.11812	-0.12314	5.740472	32.54702
0.114484-0.119978	2362	0.13421	-0.04888	5.835835	30.25672
0.119978-0.126345	2362	0.12151	-0.18136	4.845342	24.71178
0.126345-0.133768	2362	0.12489	-0.20241	4.457406	21.35531
0.133768-0.142006	2362	0.13251	-0.20158	4.173112	18.47977
0.142006-0.151341	2362	0.16511	-0.06504	4.420063	16.76345
0.151341-0.162878	2362	0.17273	-0.08719	3.990524	13.80263
0.162878-0.175907	2362	0.18036	-0.11414	3.566327	11.19828
0.175907-0.192736	2362	0.19517	-0.11995	3.193741	8.779513
0.192736-0.215233	2362	0.24725	0.008106	3.09408	6.377189
0.215233-0.248828	2362	0.24174	-0.12977	2.375334	4.416528
0.248828-0.310036	2362	0.28577	-0.13684	1.862063	2.409699
0.310036-1	2362	0.39881	-0.16951	1.039704	0.43815

Table C.3: Mean, Variance, and Skewness of Unit Bet Returns by Even Subjective Probability Categories using Opening Win Odds

Lower-Upper (1)	N_k (2)	OP_k (3)	Mean (4)	Variance (5)	Skewness (6)
0-0.0125	1836	0.00436	-0.6399	29.62951	2427.309
0.0125-0.025	11370	0.01363	-0.41906	24.41935	1012.257
0.025-0.0375	14927	0.02385	-0.36776	16.36058	413.0241
0.0375-0.05	14401	0.03451	-0.34352	12.05665	213.5126
0.05-0.0625	12744	0.05406	-0.19954	11.21043	148.0294
0.0625-0.075	10784	0.07242	-0.12234	9.86575	102.2421
0.075-0.0875	9145	0.08464	-0.13361	8.118173	69.03518
0.0875-0.1	7436	0.08983	-0.20325	6.431718	46.79527
0.1-0.1125	6227	0.11241	-0.12021	6.111541	37.07743
0.1125-0.125	5058	0.12317	-0.13635	5.309799	28.05945
0.125-0.1375	3992	0.12500	-0.20721	4.39959	20.92765
0.1375-0.15	3356	0.15644	-0.09542	4.41236	17.53138
0.15-0.1625	2541	0.16804	-0.10659	3.951657	13.94818
0.1625-0.175	2334	0.18209	-0.10197	3.622434	11.35893
0.175-0.1875	1770	0.19322	-0.11386	3.278761	9.226083
0.1875-0.2	1588	0.21348	-0.08437	3.088895	7.592154
0.2-0.225	2363	0.25307	-0.00953	2.895527	5.596793
0.225-0.25	1569	0.24729	-0.13438	2.280744	4.035034
0.25-0.275	1228	0.27850	-0.11869	2.012174	2.820768
0.275-0.3	877	0.28392	-0.1804	1.6942	2.113522
0.3-0.325	593	0.31029	-0.17742	1.504065	1.512906
0.325-0.35	484	0.33058	-0.18618	1.341163	1.118751
0.35-0.375	353	0.39660	-0.09062	1.258165	0.596589
0.375-0.4	268	0.36194	-0.2235	1.062941	0.629668
0.4-0.45	365	0.44110	-0.1344	0.949366	0.219478
0.45-0.5	195	0.49231	-0.13664	0.768681	0.020739
0.5-0.55	124	0.44355	-0.29431	0.624761	0.112226
0.55-0.6	78	0.55128	-0.19806	0.523465	-0.0781
0.6-1	95	0.72632	-0.10925	0.298974	-0.16596

Table C.4: Mean, Variance and Skewness of Unit Bet Returns by Rank of Favouritism using Closing Win Odds

Ranking	N_k	OP_k	Mean	Variance	Skewness
(1)	(2)	(3)	(4)	(5)	(6)
1	10361	0.27526	-0.1831	1.756994	2.343667
2	10361	0.17016	-0.16289	3.417546	11.09137
3	10361	0.13020	-0.15213	4.802502	23.13042
4	10359	0.10696	-0.11561	6.530317	42.44452
5	10333	0.08429	-0.12815	8.25754	71.00995
6	10246	0.06383	-0.18316	9.785901	109.2441
7	10006	0.05487	-0.12925	13.06089	184.5341
8	9448	0.04414	-0.14175	15.95263	282.8248
9	8585	0.03448	-0.18468	18.61528	409.8428
10	7394	0.02570	-0.27167	20.113	540.7759
11	6151	0.01951	-0.34513	21.55345	695.2656
12	4920	0.02053	-0.19677	30.78352	1155.038
13	3767	0.01248	-0.43637	25.14376	1107.506
14	2809	0.01460	-0.23486	39.52414	2011.425
15	1601	0.01124	-0.34376	37.87294	2160.883
16	942	0.01168	-0.22605	50.6976	3281.692
17/18	457	0.00656	-0.52046	34.80076	2508.829

Table C.5: Mean, Variance and Skewness of Unit Bet Returns by 25 Even Sized Subjective Probability Categories using Closing Win Odds

Lower-Upper	N_k	OP_k	Mean	Variance	Skewness
(1)	(2)	(3)	(4)	(5)	(6)
0-0.01311	4725	0.00550	-0.56359	34.42021	2699.77
0.013111-0.016763	4724	0.00868	-0.51936	26.38674	1435.918
0.016764-0.020143	4724	0.01693	-0.23841	33.67026	1462.932
0.020144-0.023511	4724	0.01884	-0.28268	26.79722	981.8479
0.023512-0.027174	4724	0.02117	-0.30568	22.29138	700.1951
0.027175-0.030911	4724	0.02180	-0.37584	17.47813	478.5223
0.030912-0.034839	4724	0.03048	-0.2295	18.88196	448.1756
0.034840-0.039012	4724	0.03831	-0.13681	18.70143	389.0336
0.039013-0.04345	4724	0.04805	-0.03155	18.58036	338.4812
0.043451-0.048241	4724	0.04975	-0.09877	15.515	253.1143
0.048242-0.053391	4724	0.04784	-0.21836	12.15978	179.6618
0.053392-0.058838	4724	0.05843	-0.13462	12.06908	157.8769
0.058839-0.064795	4724	0.05970	-0.19727	10.15009	120.1944
0.064796-0.071462	4724	0.07515	-0.08305	10.34781	107.2864
0.071463-0.078806	4724	0.07875	-0.12918	8.871678	82.65627
0.078807-0.086929	4724	0.08721	-0.12512	8.010862	66.34306
0.086930-0.095917	4724	0.09399	-0.14546	7.039258	51.97034
0.095918-0.10624	4724	0.10775	-0.11333	6.510365	42.02965
0.106241-0.118228	4724	0.10965	-0.18735	5.362246	31.02498
0.118229-0.132756	4724	0.12278	-0.18538	4.741376	23.73397
0.132757-0.151227	4724	0.14670	-0.13992	4.302842	17.8257
0.151228-0.176891	4724	0.16490	-0.16127	3.562506	12.14373
0.176892-0.215018	4724	0.20152	-0.13969	2.932549	7.473304
0.215019-0.275531	4724	0.23963	-0.17866	2.140621	3.820791
0.275532-1	4724	0.35161	-0.16664	1.280681	0.900847

Table C.6: Mean, Variance and Skewness of Unit Bet Returns by 50 Even Sized Subjective Probability Categories using Closing Win Odds

Lower-Upper	N_k	OP_k	Mean	Variance	Skewness
(1)	(2)	(3)	(4)	(5)	(6)
0-0.010795	2363	0.00466	-0.5685	39.81127	3655.909
0.010795-0.01311	2362	0.00635	-0.55992	30.30232	2073.195
0.01311-0.01503	2362	0.00804	-0.52571	27.7405	1609.328
0.01503-0.016763	2362	0.00931	-0.51373	25.15028	1288.571
0.016763-0.018482	2362	0.01439	-0.32118	31.55113	1445.057
0.018482-0.020143	2362	0.01948	-0.16285	35.28486	1457.673
0.020143-0.021798	2362	0.01905	-0.24548	29.31232	1116.643
0.021798-0.023511	2362	0.01863	-0.31713	24.56648	867.0055
0.023511-0.025332	2362	0.01820	-0.38046	20.69988	678.7962
0.025332-0.027174	2362	0.02413	-0.23611	23.59696	710.8984
0.027174-0.028985	2362	0.02456	-0.27354	20.96397	589.7458
0.028985-0.030911	2362	0.01905	-0.47168	14.37162	383.352
0.030911-0.032848	2362	0.03218	-0.16128	21.15923	516.0564
0.032848-0.034839	2362	0.02879	-0.29367	16.83072	389.16
0.034839-0.036845	2362	0.03091	-0.28359	16.09341	349.9915
0.036845-0.039012	2362	0.04572	0.001984	20.95328	417.1756
0.039012-0.041227	2362	0.04572	-0.05283	18.7233	352.3827
0.041227-0.04345	2362	0.05038	-0.01137	18.4225	325.0792
0.04345-0.045876	2362	0.04784	-0.1101	15.76118	265.1248
0.045876-0.048241	2362	0.05165	-0.088	15.27126	241.7878
0.048241-0.050845	2362	0.04657	-0.21939	12.47513	189.6298
0.050845-0.053391	2362	0.04911	-0.21735	11.86014	170.4436
0.053391-0.056033	2362	0.05800	-0.1196	12.5884	168.9122
0.056033-0.058838	2362	0.05885	-0.14891	11.58458	147.8227
0.058838-0.061699	2362	0.06266	-0.13641	11.15665	134.4966
0.061699-0.064795	2362	0.05673	-0.25525	9.22222	107.3297
0.064795-0.068008	2362	0.07240	-0.09467	10.50172	112.3112
0.068008-0.071462	2362	0.07790	-0.07196	10.19458	102.5277
0.071462-0.075083	2362	0.07959	-0.09778	9.412896	89.71302
0.075083-0.078806	2362	0.07790	-0.15907	8.370582	76.28178
0.078806-0.082783	2362	0.08425	-0.13381	8.155163	69.71651
0.082783-0.086929	2362	0.09018	-0.11686	7.868913	63.16397
0.086929-0.091264	2362	0.09060	-0.15538	7.16052	54.65744

0.091264-0.095917	2362	0.09738	-0.13601	6.919558	49.4391
0.095917-0.100908	2362	0.10457	-0.11692	6.677515	44.59596
0.100908-0.10624	2362	0.11092	-0.10954	6.355456	39.70135
0.10624-0.111957	2362	0.11600	-0.11651	5.948211	34.79175
0.111957-0.118228	2362	0.10330	-0.25445	4.824961	27.62819
0.118228-0.12505	2362	0.12151	-0.17017	4.978675	25.73879
0.12505-0.132756	2362	0.12405	-0.19972	4.522441	21.93756
0.132756-0.141188	2362	0.13633	-0.17338	4.328999	19.09244
0.141188-0.151227	2362	0.15707	-0.10858	4.2644	16.5989
0.151227-0.162937	2362	0.16003	-0.15259	3.769116	13.57026
0.162937-0.176891	2362	0.16977	-0.16928	3.374785	10.90643
0.176891-0.194337	2362	0.19136	-0.14069	3.120269	8.648875
0.194337-0.215018	2362	0.20025	-0.18532	2.650612	6.464515
0.215018-0.240894	2362	0.23074	-0.15577	2.376166	4.681947
0.240894-0.275532	2362	0.24852	-0.1988	1.94108	3.147482
0.275532-0.329301	2362	0.29128	-0.19478	1.577579	1.820505
0.329301-1	2362	0.41194	-0.14543	1.04251	0.380896

Table C.7: Mean, Variance and Skewness of Unit Bet Returns by Even Subjective Probability Categories using Closing Win Odds

Lower-Upper	N_k	OP_k	Mean	Variance	Skewness
(1)	(2)	(3)	(4)	(5)	(6)
0-0.0125	4073	0.00540	-0.55579	36.33507	2955.927
0.0125-0.025	16765	0.01479	-0.34929	28.20027	1203.786
0.025-0.0375	15351	0.02756	-0.26446	19.09281	481.5612
0.0375-0.05	12647	0.04800	-0.08505	16.60473	286.1543
0.05-0.0625	10794	0.05614	-0.16803	11.63668	153.0802
0.0625-0.075	8825	0.07297	-0.11562	9.935756	102.838
0.075-0.0875	7464	0.08307	-0.14928	7.989038	68.22766
0.0875-0.1	6366	0.10101	-0.10386	7.14773	50.6056
0.1-0.1125	5306	0.10837	-0.15151	5.923455	36.32686
0.1125-0.125	4508	0.11402	-0.20141	4.955582	26.79388
0.125-0.1375	3709	0.12915	-0.18121	4.520808	21.25914
0.1375-0.15	3127	0.14870	-0.14014	4.232643	17.19559
0.15-0.1625	2519	0.16118	-0.14161	3.834818	13.84006
0.1625-0.175	2186	0.16926	-0.16591	3.414564	11.13044
0.175-0.1875	1814	0.18302	-0.16016	3.148502	9.159267
0.1875-0.2	1642	0.19732	-0.15423	2.909867	7.550329
0.2-0.225	2630	0.21255	-0.17022	2.550915	5.725322
0.225-0.25	2008	0.23456	-0.1792	2.198528	4.084206
0.25-0.275	1613	0.25728	-0.18679	1.909045	2.929096
0.275-0.3	1326	0.29487	-0.14901	1.731741	2.050355
0.3-0.325	912	0.28838	-0.23556	1.44202	1.617872
0.325-0.35	689	0.31060	-0.23488	1.299383	1.212531
0.35-0.375	499	0.37675	-0.13573	1.235675	0.698718
0.375-0.4	370	0.40270	-0.13571	1.107952	0.462727
0.4-0.45	516	0.41667	-0.18006	0.941227	0.308701
0.45-0.5	259	0.50579	-0.10803	0.777383	-0.01588
0.5-0.55	114	0.57895	-0.07546	0.621656	-0.15675
0.55-0.6	47	0.70213	0.020327	0.441665	-0.25946
0.6-1	22	0.77273	-0.00859	0.289085	-0.20231