AN EXAMINATION OF INFORMATION EFFICIENCY IN FINANCIAL MARKETS, WITH SPECIAL REFERENCE TO BRITISH RACETRACK BETTING MARKETS

Leighton Vaughan Williams

A thesis submitted in partial fulfilment of the requirements of The Nottingham Trent University for the degree of Doctor of Philosophy

March 1997

PhD SLC
97/WIL Ref.

ACKNOWLEDGEMENTS

It seems a long time since I first set out on the long and winding road that would lead to presentation of this thesis. In that time, I have benefited from the support, encouragement and advice of numerous people. I would like to take this opportunity to thank all of them for their help and kindness. Without them, there is little doubt that this presentation would have floundered or at the very least been severely delayed. There is in fact no-one in the Department of Economics and Politics at Nottingham Trent University who has not helped me in some way over the years, directly or indirectly, and so contributed to the final product. I must single out some, however, for special thanks. In particular, I owe a profound debt of gratitude to my Director of Studies, Andy Cooke, whose limitless patience and attention to detail are legendary. I would also like to thank my other supervisors; Vaughan Galt, for unstinting advice, encouragement and guidance, and John Stancer, for his support from the outset. Others in the Department to whom I would like to offer special thanks are Tony Dnes, Mike Pearson, Stuart Fraser, Steve Heasell, Dave Pierce and Phil Quinn, for always being there when I needed help, assistance, or just reassurance. The secretarial support I have received from the Department has been quite simply the best anyone could have asked for. The efficiency and speed with which my requests for assistance have been met has never ceased to amaze me, and continues to do so. I would like to thank in particular Sandra Odell, Emma Graham, Linda Osborne and Claire Prendergast. Angela Sampson, however, deserves a special mention, not only for her total professional competence, but also for a generosity of assistance which has regularly far surpassed the normal call of duty.

David Paton may be comparatively new to the joys of betting markets, but his enthusiasm for the subject reflects the zeal of the convert. I owe him a debt of gratitude, and thank him also for his steadfast friendship even on those occasions when my advice has cost him dearly in the local betting shop. In particular, I would like to acknowledge his work in converting my original database into a form which was compatible with the Stata econometrics package, and his contribution in devising and operationalizing some of the statistical tests which have been incorporated into chapters 7, 8 and 9 of this thesis. Specifically, this contribution is reflected in three published papers which are based on work in this thesis, i.e. Vaughan Williams and Paton (1996, 1997a, 1997b). The responsibility for all other elements, in particular the theory, original database construction, formulation of hypotheses, and interpretation of results, rests with me alone.

I must also mention Alistair Bruce, of Nottingham University. He was one of the first people I spoke to about my idea of undertaking a doctoral programme on a subject involving the economics of betting. His support and obvious enthusiasm for the idea was a great motivation.

Finally, I would like to thank my family. To my parents - Thank you! Without them this thesis would never have been written, in more senses than one. My mother has for years reminded me that she was told when I was a baby that I would grow up to be a doctor. She may now see the prediction come true. To my wife, Julie - my love and my thanks. I owe her more than she will ever realize.

And finally, I would like to mention my great uncle, Austin Donelon, who first introduced me as a young boy in Merthyr Tydfil to the joys of horse racing. It was he who transported my first bet to the bookies'. It was 6d. each way on Grey of Falloden. I remember listening to the radio, as it stormed through from the back (or so it seemed) to land the spoils at 100 to 8. After that, I was hooked. If the Grey had lost, no doubt this thesis would never have been written.

A veteran of Dunkirk, Uncle Austin died some years ago. I remember him, and the rest of my family from those carefree years in the Welsh valleys, with deep affection.

PUBLISHED WORK

The following articles have been published in academic journals, and are based on work included in this thesis.

1. Vaughan Williams, Leighton and Paton, David (1996), 'Risk, Return and Adverse Selection: A Study of Optimal Behaviour Under Asymmetric Information', *Rivista di Politica Economica*, no. XI-XII, November/December, pp. 63-81.

2. Vaughan Williams, Leighton and Paton, David (1997a), 'Why is there a Favourite-Longshot Bias in British Racetrack Betting Markets?', *Economic Journal*, vol.107, pp. 150-158.

3. Vaughan Williams, Leighton and Paton, David (1997b), 'Does Information Efficiency Require a Perception of Information Inefficiency?', *Applied Economics Letters* (forthcoming, October).

I would also like to thank David Paton for assistance in the preparation of Chapter 9, a form of which has been accepted for presentation at the Royal Economics Society conference at Stoke, on Tuesday, March 25th.

In addition, I have been commissioned by the *Bulletin of Economic Research* to write a survey article under the title: 'Information Efficiency in Betting Markets'. This survey is based on work in this thesis, and was submitted in February, 1997.

AN EXAMINATION OF INFORMATION EFFICIENCY IN FINANCIAL MARKETS, WITH SPECIAL REFERENCE TO BRITISH RACETRACK BETTING MARKETS

Leighton Vaughan Williams

Submitted for the degree of Doctor of Philosophy

The issue of economic efficiency has, for many years, been at the heart of a great deal of economic theory and analysis. An aspect of economic efficiency which has occupied a growing place in the literature on financial markets is information efficiency, i.e. the extent to which markets incorporate available relevant information. This thesis traces the development of the idea of informationally efficient markets, and identifies the various precise definitions and variations of the concept extant in the literature on financial markets. The theoretical background is clarified, and empirical tests of information efficiency are reviewed and evaluated.

Information efficiency is also central to much of the literature in the field of betting markets. This research programme aims, therefore, within the context of an examination of information efficiency generally, to examine specifically the conduct and performance of those active in British horse race betting markets.

The first stage of this analysis involves an extensive survey of the existing literature on information efficiency as it applies to financial markets, betting markets and most specifically British betting markets.

To test and build upon the conclusions of earlier research, a large new database of information relating to British betting markets has been collected and collated. The data sets have been broken down into sub-periods and include extensive information on a large number of characteristics relevant to these markets.

A selection of the recommendations of professional forecasting services have also been collected and collated, and the performance of these has been monitored and assessed.
Tests performed in other studies on previous data are, where appropriate, adapted and tested on the current data set, and a number of new tests are proposed and applied.

The aim of this research, therefore, is to add to our understanding of betting markets in particular and financial markets more generally. This requires an examination of the availability of information to market participants and where appropriate, the relative availability of such information to different subsets of those active in the market. As a result it is possible to assess the degree to which such markets are or can be informationally efficient as variously defined.

CONTENTS

LIST OF TABLES

Chapter 9

Appendix 2

CHAPTER ONE

AN EXAMINATION OF INFORMATION EFFICIENCY IN FINANCIAL MARKETS, WITH SPECIAL REFERENCE TO BRITISH RACETRACK BETTING MARKETS : AN INTRODUCTION

1.1 Introduction

In the field of economics the concept of a market has a special meaning. The market is the medium in which exchanges take place between buyers and sellers. The market mechanism is the process whereby the independent decisions of these buyers and sellers are co-ordinated. Although early analyses of the workings of the market mechanism bypassed consideration of the precise influence of changes in price on market demand, subsequent work sought to formalize the way in which the interplay of supply, demand and price could combine to produce optimum results. Later studies introduced uncertainty into the process, seminal research in this area being identified with the work of Allais (1953), Arrow (1953), Debreu (1959) and Borch (1962). It is models such as these which served as the basis for developments in the analysis of financial markets.

A central issue in the analysis of markets is the degree to which they are efficient. Although efficiency has a variety of meanings in different contexts, it has a particular meaning in neo-classical equilibrium models of supply and demand. A situation may be termed efficient in this framework if it is not possible to increase the well-being (utility) of any one person without reducing the utility of another. This is usually referred to as Pareto efficiency[1]. An implication of Pareto efficiency is productive efficiency, a situation which exists when it is not possible to increase the quantity produced of any one good without reducing the quantity produced of another. In the analysis of financial markets, however, the examination of efficiency assumes an informational dimension, the existence of which may well be related to that of Pareto or productive efficiency, but the meaning of which is quite distinct. Basically, a market can be termed efficient with respect to some information set if prices of all assets within the market incorporate that information set. It is this form of efficiency which is the subject of investigation in this study. This thesis traces the development of this idea of informationally efficient markets, and identifies the various precise definitions and variations of the concept extant in the literature on financial markets. The theoretical background is clarified, and empirical tests of information efficiency are reviewed and evaluated.

While an examination of the issue of information efficiency is at the heart of numerous analyses of financial markets, it is also central to much of the literature on betting markets. Indeed, there are a number of special features of betting markets which warrant particular attention and make them of unique relevance to a study of market efficiency. In particular, these markets not only possess many of the usual attributes of financial markets, notably a large number of investors (or bettors) with potential access to widely available rich information sets, but also the important additional property that each asset (or bet) possesses a well-defined end point at which its value becomes certain[2]. This contrasts with most financial markets, where the value of an asset in the present is dependent both on the present value of future cash flows and also on the uncertain price at which it can be sold at some future point in time. The defined termination point of betting markets is of particular appeal, therefore, in that it allows researchers employing empirical techniques to avoid many of the difficulties associated with indefinite expected future outcomes. Moreover, by enabling a more productive and clearer learning process, a delineated end-point might be expected in particular to promote information efficiency. Evidence of inefficiency in such markets is therefore of special significance. The possibility of insider information and consequent opportunities for insider trading in horse race betting markets is also somewhat analogous to the operation of conventional financial markets, but in some respects easier to measure and assess. For these reasons, the information provided by an examination of horse race betting markets is a convenient and useful perspective from which to consider the evidence and interpretations of consumer/investor behaviour observed in conventional financial markets, as well as the operation of these markets.

1.2 Aims and Objectives

This research programme aims, within the context of an examination of information efficiency, to assess behaviour within various financial markets, with specific reference to the conduct and performance of those active in British horse race betting markets. The first stage of this analysis involves an extensive survey of the existing literature on information efficiency as it applies to financial markets, betting markets and most specifically, British betting markets.

To test and build upon the conclusions of earlier research, a large new database of information relating to British betting markets has been collected.[3] The database is a collation of evidence on 58 characteristics relevant to racetrack betting markets. The data sets are broken down into sub-periods and include extensive information on starting prices[4], on-course opening prices[5], forecast prices[6] and early bookmaker prices[7]. Other information includes the position of the race in the day's running order, totalisator ('tote')[8] - or parimutuel - returns, and other pertinent data.

A selection of the recommendations of professional forecasting services are also collected and collated, and the performance of these is monitored and assessed.

Tests performed in other studies on previous data are, where appropriate, adapted and tested on the current data set, and a number of new tests are proposed and applied.

1.3 Purpose and scope of the study

The research adopts a three-tiered structure: financial markets, racetrack betting markets generally, and British racetrack betting markets specifically. An examination is undertaken of the availability of information to market participants and where appropriate, the relative availability of such information to different subsets of those active in the market. As a result, it is possible to assess the degree to which such markets are or can be efficient with respect to information, as variously defined. This thesis surveys and assesses the literature to date.

Most research on British racetrack betting markets has addressed the bookmaker/bettor market rather than the British version of the parimutuel, i.e. the totalisator ('tote'). An additional source of interest in these markets is the significance of any information made available by movements in the odds, and the possibility for profitable arbitrage at prices different from the starting price or through differences in the tote return and the return offered by bookmakers. In this respect, exclusive research into parimutuel market behaviour does not address adequately the particular issues relating to information efficiency posed by the sequential price dynamics of a bookmaking system. Moreover, the possibilities for arbitrage in such frameworks are limited and usually confined to the context of cross-track betting on horses exhibiting different odds in different pools.

By examining data from a recent period in British horse race betting, this research programme seeks to build upon and refine the work to date, in the following three ways:

1. by assessing how far the findings of earlier studies of British racetrack efficiency can be reproduced in the new data set;
2. by analyzing where appropriate how far the findings of studies of racetrack efficiency in non-British markets can be reproduced using this new, specifically British data;
3. by proposing and running new tests of the British data collected for this research programme. On the basis of the findings to draw implications for the existence and extent of information efficiency in these markets.

1.4 Methodology

The methodology is essentially quantitative, applying tests of various hypotheses about market behaviour to a database of information about market odds, movements in these odds, forecast odds, race forecasts and a large number of other characteristics.

1.5 Hypotheses to be tested

This analysis will test three hypotheses with respect to the degree to which British horse race betting markets are efficient with respect to information, using concepts and definitions of weak, semi-strong and strong form efficiency, which were popularized by Fama (1970), and are traceable to Roberts (1959, 1967). The specific hypotheses are defined below.

1. That British horse race betting markets are weak form efficient, i.e.

efficient with respect to the information set of historical prices. In a 'weakly efficient' market no patterns can be identified which would allow future price movements to be predicted from past price movements, and no trading 'rule' will produce consistent above-average or abnormal returns except by chance. Prices are influenced solely by new economic events and new information.

2. That British horse race betting markets are semi-strong form efficient, i.e. efficient with respect to all publicly available information. In a market which is 'semi-strong form efficient', new information impacts on security prices instantaneously and in an unbiased fashion. Such prices, therefore, most faithfully reflect publicly available information.

3. That British horse race betting markets are strong form efficient, i.e. efficient with respect to all information, including privately and monopolistically held information. In a market which is 'strong form efficient', prices reflect all information, including that not publicly available.

1.6 Structure of the thesis: Reviewing the literature: Chapters 2 to 4.

In chapter 2, a review is undertaken of the academic literature which has investigated the issue of information efficiency in financial markets. The development of the idea of an informationally efficient market is explored, and the various classifications of this issue are identified. Empirical tests of information efficiency in financial markets are assessed and evaluated. The chapter is entitled: "Information Efficiency in Financial Markets:

Concepts, Definitions and Tests."

In chapter 3, a review is undertaken of the academic literature which has investigated the issue of weak form information efficiency in racetrack betting markets. The various empirical tests which have been applied in this area are assessed and evaluated. The chapter is entitled:

"Weak Form Efficiency in Racetrack Betting Markets: Concepts, Definitions and Tests."

Chapter 4 follows the same approach as chapter 3, but this time with respect to semi-strong and strong form efficiency. The chapter is entitled: "Semi-Strong and Strong Form Information Efficiency in Racetrack Betting Markets: Concepts, Definitions and Tests."

1.7. Structure of the Thesis: Applying new ideas and tests to investigate the existence of information efficiency in British racetrack betting markets: Chapters 5 to 9; Summary and conclusions: Chapter 10.

In order to contribute something original to the literature, the following structure and approach will be adopted.

In chapter 5 (and in Appendix 1), the large new data set assembled for this project is described and explained in detail. The chapter is entitled: "Constructing a New Data Set to Test for Information Efficiency in British Racetrack Betting Markets."

In chapter 6, initial use is made of the new data set to assess, in the context of particular identified earlier studies, the potential to earn above-average or abnormal returns on the basis of information contained in past and present patterns of odds. Implications are drawn from this for the existence of weak form efficiency in betting markets. The chapter is entitled: "Employing the New Data Set to Test for Weak Form

Efficiency in British Racetrack Betting Markets."

In chapter 7, an analysis is presented of the results of an extensive investigation into the performance of forecasts (tips) sold privately by well-publicized professional tipping organisations. In the same chapter, tests are conducted of the hypothesis that the expected return to identical bets placed in different arenas are also identical. Specifically, tote and starting price returns are compared in recent British racetrack betting markets, and the tote Dual Forecast is compared with an equivalent bookmaker-generated Computer Straight Forecast bet. Evidence is also presented on an observed relationship between the existence of significant tote/starting price differentials and movements in bookmakers' odds during the development of the market. An interpretation of this evidence is proposed which asks whether the realization of information efficiency in a market actually requires a perception of information inefficiency. Three separate tests for the existence of market 'anomalies' are also proposed and applied to the same data set. The first test is for a 'calendar effect' in British racetrack betting markets, i.e. significant differences in expected returns on different days of the week; the second test is for a 'gambler's fallacy' in these same markets, i.e. an over-reaction to recent events; the third test is of whether horses starting at given odds offer a different expected return depending on their rank order of favouritism. The chapter is entitled: "Employing a New Data Set to test for Semi-Strong Form Efficiency in British Racetrack Betting Markets."

Chapter 8 investigates the incidence of insider trading in British racetrack betting markets. To do this, some earlier studies of this issue are re-evaluated using the new data set, and new tests are proposed which seek to isolate the effect of insider trading on the favourite-longshot bias observed in British racetrack betting markets. A new model of bookmaker behaviour is also proposed and developed, which seeks to explain the behaviour of bookmakers operating in the context of insiders who may possess information superior to their own. Empirical tests are applied to the new model. Implications are drawn from these results for the existence of valuable inside information in recent British racetrack betting markets, and hence for the existence of strong form efficiency in these same markets. The chapter is entitled: "Employing a New Data Set to test for Strong Form Efficiency in British Racetrack Betting Markets."

In Chapter 9, a general theory is proposed which attempts to explain the different findings reported in a range of studies which have investigated the direction and extent of a favourite-longshot bias in identified fixed-odds and parimutuel betting markets. The chapter is entitled: "Explaining Positive, Negative and Zero Favourite-Longshot Biases."

In Chapter 10, a summary is presented of the main findings and conclusions of this investigation into:-

"Information Efficiency in Financial Markets, with Special Reference to British Racetrack Betting Markets."

ENDNOTES

1. After Vilfredo Pareto, who first developed this idea in his Manuel d'Economie Politique (1909), and whose contributions to economics are discussed in J.A.Schumpeter, "Vilfredo Pareto, 1848-1923," *Quarterly Journal of Economics*, 1949 (see Blaug, 1968, p.612).

2. This is a point raised in Thaler and Ziemba (1988), Gabriel and Marsden (1990), Cain, Law and Peel (1992), and Vaughan Williams and Paton (1996, 1997a).

3. This is derived from information supplied in the racing trade newspaper, the *Sporting Life*.

4. The starting price of a horse is the officially determined independent assessment of the general price at which a sizeable bet could be placed with bookmakers on the course at the start of the race.

5. The opening price of a horse is the independently determined assessment of the general price at which a sizeable bet could be placed with bookmakers at the opening of the on-course betting market for the race in question.

6. The forecast price is the forecast provided in the *Sporting Life* as to the likely starting price of most of the horses running that day. Where a race is characterized by a large number of runners a common (known as 'bar') price is frequently allocated to the least favoured runners.

7. Early prices are prices offered by bookmakers on selected races prior to the opening of the on-course market. For the present purposes the prices offered by the three major U.K. bookmakers in terms of turnover (Ladbrokes, William Hill and Corals) in *The Sporting Life* on the day of the relevant races are considered.

8. The 'tote' (or totalisator) is the British form of the parimutuel system of betting. The parimutuel is a pool system, where the total sum of bets taken is divided among the winners, subject to deductions for taxes, operating costs, profits, etc. Basically, the parimutuel odds about a given outcome are calculated by dividing the total money wagered on all the other outcomes on offer by the total amount wagered on the outcome under consideration. For win bets, the parimutuel win probability for horse h, π_h, can be written as :

$$\pi_h = w_h / W,$$

where w_h = total amount bet on horse h,

W = size of the win-bet pool.

Fabricand (1965, pp.22-32) offers a more precise analysis of how the returns are calculated.

CHAPTER TWO

INFORMATION EFFICIENCY IN FINANCIAL MARKETS:

CONCEPTS, DEFINITIONS AND TESTS

2.1 Introduction

This chapter examines some of the basic issues relating to the theory of information efficiency in financial markets and in particular, some of the definitions and distinctions which have influenced the academic literature to date. Various empirical tests of information efficiency are then reviewed and assessed.

Section 2.2 outlines the concept of information efficiency and traces the development of the terms, definitions and meanings associated with this idea. Sections 2.3, 2.4 and 2.5 review the methods which have been applied to test for the existence of information efficiency, as variously defined, in financial markets.

2.2 The 'Efficient Markets Hypothesis'

In this section, a review is undertaken of the literature which has investigated the concept and existence of information efficiency in financial markets, and in particular the role and relevance of the 'Efficient Markets Hypothesis' in our understanding of the operation of these markets.

2.2.1 The Efficient Markets Hypothesis: Reviewing the development of an idea

The concept of information efficiency in a market is contained in the so-called 'Efficient Markets Hypothesis', a standard definition of which can be found in Fama (1991),

"I take the market efficiency hypothesis to be the simple statement that security prices fully reflect all available information." (p.1575)

The origin of the ideas central to this hypothesis can be traced back to pioneering work undertaken by Bachelier (1900) into the dynamics of stock price behaviour. His examination of the behaviour of securities prices on La Bourse (the Paris Stock Exchange) led him to conclude that the price changes were identically and independently distributed, so that the next movement in a particular time series could not be predicted from an examination of previous movements. In particular, the stochastic process employed by Bachelier to describe such stock price changes has the characteristic that increments in the process are the result of independent random variables, are normally distributed with a zero mean, and possess a variance increasing in proportion to time elapsed. The implication is that stock prices have no memory and having no systematic tendencies, cannot be exploited by arbitrage. This proposition that stock price movements observe a normal distribution, and that the price changes follow a 'random walk,' laid the basis of much subsequent work into what has come to be known as efficient markets theory.

Kendall (1953), for example, analyzed serial correlations in the behaviour of weekly changes in spot prices for wheat, cotton and nineteen indices of U.K. industrial share prices. His conclusion was that the series appeared "wandering", "... almost as if once a week the Demon of Chance drew a random number from a symmetrical population of fixed dispersion and added it to the current price to determine the next week's price." (p.13)

A serious challenge to this orthodoxy can be traced to Mandelbrot (1963), whose analysis of the actual distribution of price changes disclosed evidence of high tail distributions without a finite variance. This work served to cast doubt on the value of the existing standard statistical techniques such as serial correlation analyses to test for dependence, and generated a whole new literature proposing and applying new techniques to test for such dependence.

The other important development in the literature since the late 1950s has been the clarification of hitherto implicit distinctions. In particular, the concepts of a random walk, a 'fair game', and the various 'martingale'[1] specifications are now clearly contrasted. Basically, if prices follow a stochastic process, then this can be identified as a martingale if the best forecast of tomorrow's price that can be made, based on present information, is today's price. Likewise, the stochastic process is identified as a fair game if the expected gain from forecasting tomorrow's price based on present information is zero, i.e. there is no systematic difference between actual and expected returns. The implication of the above is that if a variable in an investor's information set can be used to predict future returns, then the martingale model is violated, and returns cannot follow a fair game. The stochastic process is identified as a random walk if it satisfies the martingale conditions and also that there exists no dependence involving the higher conditional moments of future prices. The random walk specification is, therefore, more restrictive than the martingale. These issues are addressed in more detail in Section 2.2.2.

The possibility that market inefficiency can exist independently of price dependence, however categorized, can be traced to definitions originally associated with Roberts (1959, 1967), and popularized by Fama (1970), i.e. "weak form", "semi-strong form" and "strong form" efficiency. The idea is that the existence of market efficiency may best be examined in terms of three distinct types of test, each subjecting the Efficient Markets Hypothesis to different levels of strictness.

Fama (1970) discussed the tests in terms of the information subset relevant to

X

changes in security prices. First, weak form tests which are concerned with the information set of historical prices. Second, semi-strong form tests, concerned with "information that is obviously publicly available" (p.383). Third, strong form tests "concerned with whether given investors or groups have monopolistic access to any information relevant for price formation" (p.383). The three tests seek to identify which subset of information is relevant in the formation of expectations, and thereby security prices. Weak form information is limited to the price history of the relevant security; semi-strong information is limited to publicly available information; strong form information includes all known relevant information, including private information. These issues are explored in greater detail in Section 2.2.3.

2.2.2 Random walks, fair games and martingales

The idea that the absence of a random walk by financial variables is sufficient in itself to reject the existence of information efficiency in the relevant financial markets was challenged by Fama (1965). He produced findings that larger than average daily stock price changes in his data set tended to be followed by larger than average daily price changes. However, the signs of the successor changes appeared random. He concluded that although this represented a contradiction of a random walk by these variables, it did not contradict the existence of information efficiency in the markets exhibiting these characteristics.

This distinction was developed in Fama (1970), where he differentiated between a random walk and a fair game, arguing that a fair game assumption is sufficient for information efficiency, but that a fair game formulation is not sufficient in itself to lead to a random walk. In so doing he echoed Alexander's (1961) contention that assuming a "fair game" would take one "...well on the way to picturing the behavior of speculative prices as a random walk." (p.200)

LeRoy (1989) offered a clear presentation of these sort of distinctions. Specifically, he identified a stochastic process x_t as a martingale[2] with respect to a sequence of information sets I_t, if x_t has the property

$$E(x_{t+1} \text{ given } I_t) = x_t.$$

where E (n) represents the expected value of n.

So, in assuming that x_t is in I_t, then if x_t is a martingale, the best forecast of x_{t+1} based on current information I_t would be x_t. If the process is a fair game, then the expected gain from forecasting x_{t+1} based on current information I_t is zero.

The implication of the above is that if a variable in an investor's information set can be used to predict future returns the martingale model is violated, and returns cannot follow a fair game. A stochastic process is identified as a random walk if it satisfies the martingale conditions and also that there exists no dependence involving the higher conditional moments of x_{t+1}. If, for instance, we model security price behaviour in such a way that successive conditional variances of such prices are positively autocorrelated (though not their levels), then this satisfies the martingale conditions, but not the random walk. The existence of risk neutrality, in which investors are unconcerned about the higher moments of their return distributions, points therefore to a martingale formulation but not a random walk, since investors in such a scenario are not led to bid away serial dependence in these higher conditional moments. The presence of risk aversion, on the other hand, runs contrary to a martingale and a fair game modelling. The reason stems from the fact that risk-averse investors will only hold more risky assets if they are compensated in terms of higher expected returns. As a consequence, knowledge of the riskiness of the current information set implies some knowledge about the level of expected returns. The idea of a sub-martingale, originally proposed by Fama (1976) in reply to LeRoy (1976), is that expected rates of return (ignoring dividends), conditional on currently available information, are non-negative, i.e.

$$E\ (p_{t+1} \text{ given } I_t) \geq p_t\ ,$$

implying that no trading rule based on the current information set can outperform a strategy of buy-and-hold.

Granger (1992) pointed out that if stock prices were not a martingale, then ignoring transactions costs "... price changes would be consistently forecastable and so a money machine is created and indefinite wealth is created." (p.3). He took care to differentiate, therefore, between a martingale process and the various interpretations identified with the Efficient Markets Hypothesis, expressing his own preference for Jensen's (1978) definition, i.e. that a market is efficient with respect to a given information set if it is impossible to make economic profits[3] by trading on the basis of this information.[4]

Support for Jensen's definition is offered by Fama (1991), in a follow-up to his original 1970 survey of the literature on efficient capital markets. Fama (1991) noted Grossman and Stiglitz's (1980) finding that for security prices to fully reflect all available information then information and trading costs must be zero. Finding this implausible, he preferred Jensen's "... weaker and more sensible version of the efficiency hypothesis [which] says that prices reflect information to the point where the marginal benefits of acting on information (the profits to be made) do not exceed the marginal costs." (p.1575)

A related issue is raised by Keane (1993), who highlighted a distinction between rationality and exploitability as aspects of pricing efficiency. For Keane (1993), the market is rational if prices and market movements reflect the best estimates of intrinsic values. It is fair-game efficient or inexploitable if systematic abnormal returns cannot be earned through an analysis of price behaviour. The distinction is made clear in a situation where the market in aggregate is subject to excessive movements that are difficult to identify or are unpredictable in behaviour. In such a situation, irrational market behaviour can co-exist with fair-game efficiency or inexploitability.

The essential issues can, however, be categorized into two parts. First, is there evidence in financial markets of price change dependence as variously defined? Second, can any such evidence be used to secure systematic abnormal returns?

2.2.3 Weak, semi-strong and strong form efficiency: classifications of information efficiency

The weak form of the Efficient Markets Hypothesis holds that current security prices fully and instantaneously reflect all weak form information, and similarly for the semi-strong and strong forms of the hypothesis. In a weak form market it follows that no patterns can be identified which would allow future price movements to be predicted from past price movements, and no trading rule will produce consistent above-average or abnormal returns except by chance. Prices are influenced solely by new economic events and new information. Fama (1991) has proposed extending the categorization of research in this area to include such variables as dividend yields, interest rates, earnings/price ratios, and other term-structure variables. He identifies these as tests for return predictability, a more general category which includes weak form tests. In a semi-strong form market, new public information impacts on security prices instantaneously and in an unbiased fashion. Such prices, therefore, most faithfully reflect the available published information. In a strong form market, share prices reflect all information, including that not publicly available.[5]

Dowie (1976) made a basic distinction between the strong form of inefficiency (as hitherto defined) and the other forms of inefficiency (weak and semi-strong). The former tells us about access to and the availability of information, whereas the latter is concerned with how well the market responds to information. Although related, these are quite separate issues. Since strong inefficiency implies the existence of subsets of investors who possess monopolistic access to information (which can be exploited to earn above-average returns), he uses the term "equitable" to describe markets which pass the strong test, and "efficient" to describe those that pass the weak and semi-strong tests.

Keane (1987) also made a clear distinction. Whereas the weak and semi-strong classifications apply to the stock market itself, strong efficiency, he argues, is about a broader concept of capital markets. Specifically, whereas "... semi-strong efficiency is concerned with how well the market processes the information disclosed to it... strong efficiency is concerned primarily with the adequacy of the information disclosure process." (p.6). In this sense, it might be considered misleading to view strong efficiency as a progression from the weak and semi-strong forms, since this confuses the ability of the market to respond to and interpret information with the failure of the market to supply information (what we might call the information production function).

It can be seen that the development of research into information efficiency in recent years has sought to clarify the nature of the distribution of stock price changes, and in this context to develop statistical tests which offer the possibility of testing for dependence between successive price changes. The type and degree of dependence under examination has been clarified, and the concept of information efficiency itself has been broadened and made more explicit.

2.2.4 The Efficient Markets Hypothesis: Summary

An informationally efficient market can in essence be defined as a market which incorporates all information. This is a stringent requirement, and so studies of financial markets have also addressed the issue with respect to subsets of the totality of information. The three principal (though not exclusive) levels at which studies of information efficiency have been undertaken are with respect to weak, semi-strong and strong information.

Weak form information is information contained in the set of historical prices. A market is weakly efficient (with respect to information), therefore, if this is fully and (in the strictest form) instantaneously incorporated in present prices. In such a market, present prices reflect all information available in patterns of historical prices, and so future price movements cannot be derived from an examination of past prices.

Semi-strong information is that contained in the set of all public information. A market is semi-strong efficient if this is fully and (in the strictest form) instantaneously incorporated into present prices. In such a market, present prices reflect all available public information, and so future price movements reflect future (and as yet unknown) revelations of publicly available information.

Strong information is that contained in the set of all information, including that privately and monopolistically held. A market is strongly efficient if all information is fully and (in the strictest form) instantaneously incorporated into present prices. In such a market, present prices reflect all information, and so future price movements reflect future (and as yet unknown) revelations of information.

All these definitions of information efficiency require the incorporation of relevant information. In less strict formulations, it is sufficient for efficiency to exist that it is not possible to trade upon this information so as to earn greater than normal profits.

2.3 Empirical tests of weak form information efficiency in financial markets

This section reviews the empirical tests which have been proposed and applied in the literature to investigate the existence of weak information efficiency in financial markets.

It has already been shown that in a financial market characterized by strict weak form efficiency, no patterns can be identified from the history of price data which would allow one to predict the future pattern of price changes. In a market which is weakly inefficient as so defined, the pattern of incremental prices is well approximated by a random walk specification. Sub-sections one to four of this section review tests which have been applied to examine for the existence of such a specification. Part one (2.3.1) assesses serial correlation techniques of price dependence, part two (2.3.2) variance ratio tests, part three (2.3.3) cointegration approaches, and part four (2.3.4) looks briefly at how rescaled range analysis and chaos theory have been applied to the theory of financial markets.

A less strict form of weak efficiency holds that no information can be gathered from such price data which would allow one to make abnormal returns except by chance. In sub-section five (2.3.5), a review is undertaken of attempts in the literature to use price dependence as a means of earning abnormal returns. Finally, sub-section six (2.3.6) surveys work which assesses the possibility for predicting returns using a range of indicators, such as dividend yields. Tests of this possibility are usually termed 'tests of return predictability.'

2.3.1 Testing for price change dependence using serial correlation techniques

In examining securities markets, the most obvious test of strict weak form efficiency is to test for price change dependence. The idea here is that there should be zero correlation between increments of a random walk (cumulated series of probabilistically independent shocks).

Although Working (1934) contended that random walks generated patterns that appeared similar to those frequently imputed by market analysts to stock prices, the first rigorous empirical backing for Bachelier's (1900) idea of a 'random walk' in share prices (or "random wander" as Rowley [1987, p.131] terms it) was provided by Kendall's (1953) serial correlation analyses of weekly changes in commodity spot prices and U.K. industrial share prices. This was extended by Roberts (1959), whose work emphasised the implications of Kendall's findings for financial analysis and stock market research. He compared movements in a variable generated from a random walk process with movements in the Dow Jones (stock market) industrial average over a 52-week period between 30 December 1955 and 28 December 1956, using actual changes in Friday to Friday closing levels. He concluded that the patterns produced by both were so similar as to suggest that there is a random distribution of index changes. He noted that the random walk model, implied by the instantaneous adjustment of prices to new information, would be just what would be expected in an ideal market composed of rational investors.

Another early test of dependence can be found in Moore (1962, 1964), who examined changes in the prices of common U.S. stocks measured over weekly intervals from 1951 to 1958. He reported an average serial correlation coefficient which was not different from zero at any conventional level of statistical significance. Serial correlation and runs tests of dependence by Brealey (1970) for the British equity market, and serial correlation tests by Cunningham (1973) similarly were unable to detect evidence of dependence. Hagerman and Richmond (1973) also established no evidence of substantial direct dependence between lagged price changes in securities which were traded over-the-counter (which might therefore be smaller and less well analyzed than typical securities), a result supported by Solnik (1973) in a separate analysis of European stock prices. Cooper (1982) examined world stock markets, calculating the correlation coefficients between successive monthly changes in the stock market indices of 36 countries. He reported coefficients ranging from zero in the cases of Lebanon and Mexico to 0.40 in Ireland, concluding that overall there was no evidence of any significant relationship between successive market movements.

In conclusion, serial correlation techniques, each employing different data sets, and conducted in different periods, fail to provide convincing evidence of the existence of any systematic pattern of security price change dependence through time. It is not possible on this basis, therefore, to reject a null hypothesis of weak form information efficiency in the markets studied.

2.3.2 Variance ratio tests

Tests of information efficiency in financial markets based on an examination of the existence of any serial correlation in changes in stock market returns characterize much of the efficient markets literature, and especially the early literature. A significant trend in later research papers has been the application of variance ratios tests to the data. Such tests are based on the idea that the variance of a sample of stock price returns should, if these returns are generated by a random walk, increase in proportion to time elapsed. For example, the variance over six months should be six times as great as the variance over one month. Work by Lo and MacKinlay (1988, 1989) and Cochrane (1988) is indicative of the seminal literature in this field. The test is a widespread method of testing for mean-reversion[6] in stock prices, the idea behind such tests being that non-mean reverting stock prices implies non-predictability in the long run.[7] Evidence of eventual mean-reversion is offered by Cutler, Poterba and Summers (1990, 1991), who found negative serial correlation in returns at a three to five-year time horizon (although positive correlations at a shorter horizon), and Chopra, Lakonishok and Ritter (1992), who reported negative serial correlation in the returns of individual stocks and various portfolios over intervals of three to ten years. French and Roll (1986) and Lehmann (1990) also found negative serial correlation in weekly and daily returns of individual securities, while Rosenberg, Reid and Lanstein (1985) reported predictable return reversals on a monthly basis at the level of individual securities. Fama and French (1988a), Cochrane (1991), and Jegadeesh (1990) also found evidence of mean reversion in their data sets. This was confirmed by Kim, Nelson and Startz (1991), although they argued that the assumption of normally distributed returns tended to overstate the extent of mean-reversion in previous studies.

McQueen (1992) incorporated the findings of Kim, Nelson and Startz to re-examine the existence of mean-reversion in stock returns. For the period from 1926 to 1987, he concluded that the random walk hypothesis could not be rejected for value- or equally-weighted real returns at any of ten return horizons or by joint tests over all ten horizons simultaneously. The conclusions were unaffected by extending the study period to cover the years 1871 to 1987.

A potential difficulty in interpreting the existence of return autocorrelations generally was raised by Lo and Mackinlay (1988) who, using variance ratio tests, identified positive serial correlation in short-horizon stock returns (of the order of thirty per cent for weekly and monthly stock returns). The particular problem they highlighted was their finding that autocorrelation was stronger for the portfolios of small stocks, indicating the possible influence of a nonsynchronous trading effect. This 'nonsynchronicity problem', which can be traced to work by Fisher (1966)[8], and perhaps more generally to Working (1960)[9], arises because multiple time series are assumed to be sampled simultaneously when they may not occur simultaneously.[10] It has obvious implications for tests of the efficient markets hypothesis which rely on testing for the existence of autocorrelation in a series of returns. Conrad and Kaul (1988) tried to eliminate any nonsynchronicity effect by limiting their analysis to the Wednesday-to-Wednesday returns of size-grouped portfolios of stocks that trade on both Wednesdays. However, Fama (1991) demonstrated that their findings of positive autocorrelation was not totally free of such an effect, particularly for small stocks.

Nevertheless, an analysis by Lo and Mackinlay (1990a) indicated that while the nontrading effect may explain some of the time series properties of stock returns, there was "little support for nonsynchronous trading as an important source of spurious correlation in the returns of common stock." (p.203).

The possibility of spurious autocorrelation caused by nonsynchronous trading, and the implications of this for random walk analysis are emphasised, however, by Ayadi and Pyun (1994), in particular for the stock markets of the developing world. Applying a variance ratio test to the Korean Stock Exchange they were unable to reject the random walk hypothesis for time horizons of a week or longer. Wider tests of the random walk hypothesis are found in Frennberg and Hansson (1993) and in Huang (1995), both of whom applied variance ratio tests to specified national stock markets. Frennberg and Hansson rejected a random walk formulation for Swedish stock prices, and Huang (1995), in an analysis of a number of Asian stock markets, concluded that the random walk hypothesis was rejected for Korea and Malaysia for all holding periods. Positive serial correlation was also evident in some periods in the Hong Kong, Singapore and Thai markets.

2.3.3 Cointegration tests of the efficient markets hypothesis

Cointegration studies as a research method are traceable to the seminal work of Granger (1986), Engle and Granger (1987) and Johansen (1988). Basically, they employ the idea that since asset prices in an efficient market cannot be related in the long run, then the absence of common stochastic trends in a system of stock prices implies the existence of efficient markets. Tests of cointegration are employed to check for such trends.[11]

In the first paper to apply a cointegration methodology to the examination of stock market efficiency, MacDonald and Power (1991) tested for the existence of market efficiency in the weekly share prices of forty U.K. companies, over an eight year period. The prices were grouped into the five major industrial classifications, but no cointegrating relationships could be identified. MacDonald and Power concluded that the U.K. stock market was a rational processor of information.

Chelley-Steeley and Pentecost (1994) extended the work of MacDonald and Power by using a longer time period in order to improve the reliability of the tests. They also classified the stock prices by firm size. They found no significant evidence of cointegration in the share prices of large firms, conclusions consistent with the existence of stock market efficiency for large firms. They did find, however, considerable evidence of cointegration in the share prices of small firms, suggesting the existence of static inefficiency in the data relating to small firms.

Choudhry (1994) also employed cointegration tests to look for any evidence of common stochastic trends in a system of stock indices drawn from seven OECD[12] countries between 1953 and 1989. Being unable to identify any such trends, the findings were offered as evidence in support of the Efficient Markets Hypothesis.

The application of this approach to test for efficiency in a range of commodity markets was undertaken by Beck (1994), at an eight and 24 week horizon. His results indicated that while all five of the commodity markets he studied were inefficient sometimes, none was inefficient at all times.

Further tests, which sought to examine the existence of both short-run and long-run efficiency using cointegration (among other) techniques, were developed and applied to the FTSE-100[13] stock index futures contract by Antoniou and Holmes (1996). Their results showed that while this market was efficient over short periods (one or two months), this was not the case for longer periods. They concluded that there are consequent opportunities for consistent speculative profits to be made.

Cointegration analysis has also been employed extensively to test for efficiency in the foreign exchange markets. Ukpolo (1995), for example, used such techniques to test for efficiency in the Japanese foreign exchange market, as did Alexander and Johnson (1992) for exchange markets more generally, using London daily closing rates for the six major currencies. The conclusions of these studies were inconsistent with the efficiency hypothesis, a result reproduced by Diamandis and Kouretas (1995) in a time series analysis of the Greek Drachma.

2.3.4 The application of rescaled range analysis and chaos theory to efficient markets analysis

In addition to the serial correlation, variance ratio and cointegration approaches outlined above, some authors have applied rescaled range analysis and chaos theoretic analyses to the problem of stock return dynamics.

The idea of rescaled range analysis was first proposed by Hurst (1951), as a result of his observations of natural phenomena. The statistic, since refined by Mandelbrot (1972, 1975), Mandelbrot and Taqqu (1979) and Lo (1991), can be used to test for long-term dependence. Essentially it is a method of measuring how the path of a time series varies over various time scales. Specifically, the rescaled-range statistic is the range (i.e. high minus low) of partial cumulative sums of deviations of a time series from its mean, rescaled by its standard deviation. A convenient way of viewing its application is through an examination of the so-called Hurst exponent. Named after H.E.Hurst, who first developed its use in studies of the Nile river dam project, it is a measure of correlative persistence.

The correlation can be derived from the following equation:

$C = 2^{(2H-1)} - 1,$

where C is the measure of correlation and H is the Hurst exponent.

Thus, if the Hurst exponent equals 0.5, C=0, and the probability that a move in one direction will be followed by a move in the same direction (e.g.positive followed by positive) is fifty per cent. If the Hurst exponent is less than 0.5 the system can be characterized as mean-reverting (sometimes termed antipersistent or ergodic), if greater than 0.5 it is correlative or persistent (also sometimes termed trend-reinforcing). The period of time over which H is greater than 0.5 is a measure of the memory cycle of the system, and so measures the time period over which information can be used predictively.[14] This approach is particularly useful in the context of non-normal distributions. The reason is that the variance of such distributions may not exist (i.e. the expected value of the variance may be infinite). The essential intuition behind this is that the tails of the distribution decay too slowly. In these cases, variance ratio tests are inappropriate. The only requirement of rescaled range tests is that the mean exists (i.e. the expected value of the mean is less than infinity). In other words, these tests are less demanding about the existence of moments of the distribution. Its importance in this respect is indicated by a rescaled range analysis undertaken by Peters (1989), and quoted in the *Economist* (23/10/93, pp.1-24), of monthly data of the Standard and Poor (S&P) 500 index from January 1950 through July 1988. This study reported evidence that this data followed not a normal but a highly leptokurtic distribution.[15] He found, for periods greater than twenty and less than 110 days, evidence that the market revealed an over-long trend in one direction compared with what would be expected from a random walk. In particular, for the S&P 500 prices were found to increase with the 0.78 root of time, in contrast to the square root configuration which would be consistent with a random walk. Moreover, a memory length of about four years was identified, the length proving independent of the resolution of the data. Scrambling the data randomly so as to alter the order of the returns, though not the probability distribution, and re-running the analysis ruined the structure of the original series. This is consistent with the presence of a memory effect as proposed. Analysis of other capital markets yielded similar results.

Applying a variation of rescaled range analysis to measure the shape of the probability density function of the markets, Peters (1989) also found evidence that stock markets display consistent statistical characteristics prior to particular phases or developments, e.g. a downturn or a period of trend reinforcement. Another study to apply rescaled range statistics was that of Ambrose, Ancel and Griffiths (1992), whose examination of long-term memory in a number of U.S. asset classes concluded that the returns all displayed tendencies consistent with a random walk process. A recent application of (modified)[16] rescaled range tests was applied by Huang and Yang (1995) to nine Asian stock markets, together with U.K. and U.S. indices. No evidence of long-term memory was found for the Asian markets studies, except for the Philippines. There was evidence, however, of such an effect in the U.K. market for various data frequencies and lags.

The other type of analysis, and one which has been used in particular to explain the frequency of large movements in asset prices compared with what would be expected under a linear modelling or a normal distribution, is chaos theory. Essentially, 'chaos' is a deterministic nonlinear process which appears to be random. In chaotic models, external shocks can cause dynamic processes which follow a nonlinear path, and which by a process of self-generation (feedback) can create large and volatile movements.

Although Baumol and Benhabib (1989) offer a useful general survey of economic models which can produce chaotic behaviour and Hsieh (1991) has examined the role of chaotic processes in the specific context of financial markets, there is no coherent empirical evidence of chaotic behaviour in financial markets. Indeed, evidence produced by Hsieh contradicting a null hypothesis of independent and identical distributions in his data was explained in the study as a consequence, not of chaotic dynamics, but of conditional heteroscedasticity, e.g. predictable variance changes. Nevertheless the ideas behind chaos theory continue to motivate some economists, and coherent and convenient outlines of such can be found in Savit (1988, 1992) and Cunningham (1994).

It has been seen, therefore, that a succession of tests, using various statistical techniques to test for price dependence, and covering a wide array of data sets, fail thus far to reach a simple consensus in terms of validating the Efficient Markets Hypothesis as broadly defined. How far the differences in the conclusions can be attributed to variations in the testing techniques adopted, and how far to variations in the databases under examination awaits further work.

It is important, however, to make the distinction between the existence of stock price dependence per se, and the possibility of using such a configuration, were it to exist, in order to generate abnormal returns. The following section develops this thread in the literature.

2.3.5 Price dependence, abnormal returns and information efficiency

Fama (1970) drew a clear distinction between a definition of market efficiency which sees any existence of statistical dependence in successive price changes as refutation, and a less strict interpretation which identifies market efficiency as existing if no profitable trading rules can be based upon such dependence.

The idea of using the past history of security prices, with the object of formulating rules which would permit the generation of abnormal returns, can be traced in the modern literature to Alexander (1961)[17], who advocated a 'filter' system, whereby securities are bought or sold according to their patterns of falling and rising prices. The filter (sometimes referred to as the 'k per cent filter rule') is the name given to the percentage change in the security price used to initiate a position.

Using filters of from 1 per cent to 50 per cent for daily data on price indices between 1897 and 1959, Alexander (1961) sought to generate a profitable rule through a process of separating out random from non-random movements. Alexander (1964) concluded that, taking account of transactions costs, "... for any reader who is interested only in practical results, and who is not a floor trader" (p.351)... the filter strategy could not out-perform a simple policy of buy-and-hold.

Fama and Blume (1966), like Alexander, found that if one ignored transactions costs it was possible to formulate a trading strategy which would outperform buy-and-hold, in their case for very small filters based on very short-term trading (i.e."at most daily", p.395). Allowing for even minimum trading costs, however, the advantage disappeared. Fama (1970) concluded that although "...the filter tests, like the serial correlations, produce empirically noticeable departures from the strict implications of the efficient markets model" (p.396), "...using a less than completely strict interpretation of market efficiency, this positive dependence does not seem of sufficient importance to warrant rejection of the efficient markets model." (p.414).

Early research into the consequences of employing buy or sell strategies based on deviations from a moving average of their prices over various periods also failed to identify profitable trading rules net of transactions costs. Work by Cootner (1962) and by Van Horne and Parker (1967) typify the literature. Nevertheless, an empirical analysis by Stottner (1990) of a simple downward averaging device appears to produce a significant improvement in returns compared to a buy-and-hold strategy. Such evidence seems at variance with the implications of at least the stricter form of weak efficiency.

Fortune (1991) conducted a time series analysis of daily stock prices in the 1980s (specifically, the daily closing prices of the Standard & Poor 500 between January 2, 1980 and September 21, 1990). He rejected the random walk hypothesis, finding statistically significant coefficients in a moving average model of stock price behaviour. He calculated that a trading strategy based on these findings would not, however, be sufficient to cover retail transactions costs, although it could cover institutional transactions costs.

Brock, Lakonishok and Le Baron (1992) used data from the Dow Jones Index from 1897 to 1986 to test specified trading rules, based on a moving average and a trading range break. In particular, they explored twenty versions of the moving-average rule (i.e. buy when a short-term moving average exceeds a long-term one), and six versions of the trading-range break rule (i.e. buy when the index exceeds its last peak, sell when it falls below its last trough). They found that 'buy' signals consistently outperformed 'sell' signals in terms of returns, and that these returns were less volatile than those following 'sell' signals. Specifically, 'buy' signals produced an average annual return of 12 per cent, whereas sell signals produced an annual average loss of 7 per cent. Gencay (1996) offered further evidence of weak form inefficiency in applying a moving-average rule to the Dow Jones Industrial Average Index between January 1963 and June 1988. By employing the past buy and sell signals of these rules, Gencay (1996) provided convincing evidence of non-linear predictability in these stock market returns.

The profitability of technical trading systems when applied to futures markets was the subject of a study by Lukac and Brorsen (1990). Applying a trading simulation to 23 trading systems on 30 futures markets for eleven years, they found that all but two yielded significant positive gross returns, contrary to the implications of a random walk model. Raj and Thurston (1996) also applied technical trading strategies to a futures market, notably the (Hong Kong) Hang Seng Futures Index. They found that while their moving average strategy failed to produce significant excess returns, the majority of their trading range break rules were able to do so. This is in clear conflict with a weak form efficiency specification for this market.

The existence of a systematic link between trends in past exchange rates and subsequent returns in foreign-exchange markets was proposed by Taylor, quoted in the *Economist* (5/12/92, pp.23-26).[18] Employing data from a ten-year period to December 1991, he demonstrated the availability on this basis of particular trading rules which could produce above-average returns. Using a "double moving average" rule, for example, Taylor found average annual returns of 14.2 per cent, compared with an average annual return on U.S. Treasury bills of 8 per cent. Specifically, this rule entailed the trader using a short and a long moving average, selling when the shorter average falls below the long, and buying when the reverse occurs.

Froot, also quoted in the *Economist* (5/12/92, pp.23-26), devised a trading rule on the basis of his conclusion that short-term interest rates can be used successfully to forecast returns in foreign exchange stock, bonds and commodity markets at the same time. According to the rule, a fall of one per cent in (annualized) short rates is usually associated with an extra three percentage points in (annualized) excess returns to those investors who trade on the basis of that change in the interest rate.

There is, therefore, substantial supporting evidence for the view that trading rules can in certain circumstances produce excess returns. The methodological soundness of such studies should perhaps be placed, however, in the context of an early investigation by Levy (1967) into two hundred stocks listed on the New York Stock Exchange. Levy (1967) claimed significant abnormal returns for a trading rule which bought stocks with substantially higher current prices than their average over the previous 27 weeks. The problem came when trying to replicate these findings. A basic criticism of Levy's work, highlighted by Jensen and Bennington (1970) was that it produced a rule from existing data, rather than seeking to test such a rule against new data. In particular, Jensen and Bennington noted that Levy arrived at the successful rule only after a separate examination of 68 other possible trading rules which failed. In consequence, Levy's (1967) findings were, they suggested, due to a form of selection bias.

A more general discussion of the problems for empirical work of data-instigated pre-test biases was presented by Leamer (1978), and this issue of 'data snooping' has since been taken up by Lakonishok and Smidt (1988), Lo and MacKinlay (1990b) and Brock, Lakonishok and LeBaron (1992). Merton (1987) presented the same issue in the context of cognitive psychology, noting unintended selection biases resulting from what he proposes is a natural predilection for individuals to focus, often disproportionately, on the unusual.

The *Economist* (5/12/92, pp.23-26) highlighted a similar point made by Black and Scholes[19] in an attack on "data-mining" (the idea that data is mined by researchers until an apparent trend is found). Criticizing the claim from data analysis that shares of smaller firms out-perform those of larger firms, Black claims that "...it sounds like people searched over thousands of rules till they found one that worked in the past. Then they reported it, as if past performance were indicative of future performance. As we might expect, in real life the rule did not work any more." (p.22)[20]. Nelson and Kim (1990) made a related point, showing that overly encouraging results can result from small-sample in-sample biases of coefficients.

Early work, then, commonly suggesting that no profitable trading rules could be devised so as to generate abnormal returns, particularly net of transactions costs, has been challenged by more recent strategies of greater sophistication. The possibility modern computer power offers to generate large numbers of rules has, however, led some writers to challenge the predictive as opposed to the descriptive value of these findings. More work is needed to test proposed new trading rules on future data, and to allow for the possibility of various potential selection biases.

2.3.6 Tests of return predictability

A growing field of literature in recent years has concentrated on forecasting returns using variables such as dividend yields, interest rates, earnings/price ratios and other term-structure variables. Fama (1991) names this area of research "tests for return predictability", identifying it as a more general category which includes the weak form tests identified above. Work in this field can be traced to Bodie (1976), Nelson (1976), Jaffe and Mandelkar (1976) and Fama (1981) on the negative relationship between monthly stock returns and expected inflation, and also (Fama and Schwert, 1977) of a similar relationship between monthly stock returns and the level of short-term interest rates. Later work by Shiller (1984) and by Rozeff (1984) found that dividend yields could be used to forecast short-horizon stock returns.

Fama and French (1988b) used dividend yields to forecast the portfolio returns of New York Stock Exchange (NYSE) stocks for horizons from one month to five years, finding that such yields served to explain small fractions of monthly and quarterly return variances. Other evidence of the forecasting power of the aggregate equity market dividend yield for U.S. equity returns has been provided by Campbell and Hamao (1989), Attanasio and Wadhwani (1990) and Shah and Wadhwani (1990), although Shah and

Wadhwani called into question the general applicability of these results for other countries. Later work by Clare and Thomas (1992), using U.S., U.K., German and Japanese equity and government bond markets in the 1980s, found clear evidence of the forecasting power of assorted yield spreads. Other findings for the U.S. are offered by Campbell and Shiller (1988), who reported evidence of reliable forecasting power by earnings/price ratios, which increased with the return horizon; and by Campbell (1987) and Keim and Stambaugh (1986) who found that a common set of stock market and term-structure variables could be used to predict stock and bond returns. Harvey (1991) reported that the returns on portfolios of foreign common stocks could be forecast from U.S. term-structure variables and from the dividend yield on the Standard & Poor 500 portfolio. Evidence that such financial variables as the short rate, changes in the short rate and the term structure of interest rates can in some measure predict U.S. equity returns is offered by Fama and Schwert (1977), Keim and Stambaugh (1986), Campbell (1987), Campbell and Hamao (1989), Fama and French (1989), Attanasio and Wadhwani (1990), and Shah and Wadhwani (1990).

Tests for a relationship between average return and specific market variables are found in Banz (1981), i.e. a strong negative relationship between average return and firm size; Bhandari (1988), i.e. a positive relationship between average return and leverage; in Basu (1983), i.e. a positive relationship between average return and the Earnings/Price ratio; in Stattman (1980) and Rosenberg, Reid and Lanstein (1985), i.e. a positive relationship between return and book-to-market equity for U.S. stocks; and Chan, Hamao and Lakonishok (1991), i.e. a strong predictive power for BE/ME (Book to Market Equity) in explaining average returns on Japanese stocks.

A study which seeks to address the issue of robustness in the predictability of U.S.

stock returns in terms of a range of economic factors was presented by Pesaran and Timmerman (1995). They found that the strength of any such relationships can be linked to the volatility of the markets. In this context they identified clear evidence of past predictability, predictability which was sufficient at certain times to yield excess returns.

Thus, tests for weak efficiency have taken the form of tests of price dependence through time; of tests which seek to determine the predictability of prices in terms of identifiable trading rules or economic variables; and also of the possibilities for exploiting any such predictability in order to earn abnormal returns. Tests of price dependence have progressed form simple serial correlation tests of short-term dependence to variance ratio and mean-reversion tests of long-run patterns in the data. Cointegration tests have also begun to play an increasing role in the literature. Trading rules based on postulated strategies linked to price movements and/or to the performance of individual economic variables have and are being tested for evidence of systematic predictive validity. Any such patterns provide indicative evidence of weak efficiency. At another level, weak form inefficiency is assessed in terms of the possibilities such patterns provide for earning abnormal returns.

Although many studies have failed to identify evidence of weak form inefficiency, there are others which provide strong evidence of dependence and predictability, at least in past data sets and over specified time periods. There is much less compelling evidence, however, that these 'inefficiencies' are sufficient to provide investors with abnormal returns after the fact of their identification.

2.3.7 Empirical tests of weak form information efficiency in financial markets:

Summary

A standard test for the existence of weak information efficiency in financial markets is to test for the existence of unpredictability in the movements of security prices. A standard specification of unpredictability is the random walk. In a random walk process, price changes are identically and independently distributed, so that the next movement in a particular time series cannot be predicted from an examination of previous movements. Less strict formulations, such as a fair game or a martingale process, require only that there is no dependence between the means of the series, so that the best forecast of tomorrow's price is today's. A stochastic process is identified as a random walk if it satisfies these conditions and also that there exists no dependence through time involving the higher conditional moments of the distributions. Tests of a less strict formulation of weak form efficiency investigate the existence of trading rules, based on information contained in historical prices, which can be used to earn abnormal profits.

Early analysis of price dependence concentrated on testing for serial correlation in the behaviour of security price movements through time. Most of the empirical work was unable to reject the hypothesis of independence, although there was some evidence of correlation in the size (but not the sign) of price changes. A general consensus of such studies is that no clear trading rule can be identified, based on historical prices, which can be employed to produce abnormal profits. There was also broad support for the idea of a martingale specification (sometimes adjusted for a gradual upward drift). Evidence for a random walk was more mixed, but still substantial. Later empirical work applied variance ratio tests to the question of price dependence. The idea of these tests is to check for evidence of mean-reversion in the whole data set. If a system is mean-reverting, there is the implication that if it has moved up for a number of observations,

it is more likely to move down than up over subsequent observations, and vice versa. Such a specification is contrary to a random walk model, and implies some sort of predictability in the data. Since variance ratio tests are applied to the whole data set, they are very useful in detecting any long-term trends in the data which might otherwise have been missed. The basic idea behind these tests is that in a random walk, the variance of the returns is proportional to time elapsed. If the actual variance is less than this, evidence exists of mean-reversion. Although most studies have found evidence of mean-reversion in their data sets, some investigators have explained this as an effect of the test specification rather than as genuine inefficiency. The sample size, the sample period, the assumption of normality in the distribution, and the possibility for spurious autocorrelation (arising from nonsynchronous trading) have been cited as reasons. A number of studies have now addressed these problems, but there still remains some evidence of apparently genuine mean-reversion in a number of data sets. A potential weakness of variance ratio tests is the requirement of a finite variance in the distribution of price changes. Some evidence exists, however, that stock returns may follow a leptokurtic (fat-tailed) distribution, and typically a Cauchy distribution, characterized by an infinite variance. Analysis of such distributions require a different type of approach. Rescaled-range analysis is an approach which can be employed to test for long-term dependence in these cases. Specifically, the rescaled-range statistic is the range of partial cumulative sums of deviations of a time series from its mean, rescaled by its standard deviation. The only requirement of rescaled range tests is that the mean exists (i.e. the expected value of the mean is less than infinity). There are a very limited number of studies which have applied this analysis to financial markets, some markets displaying results consistent with a random walk formulation, others demonstrating evidence of a memory effect in the data.

Another technique comparatively new to the literature is the application of cointegration methodology to stock market analysis. The basic idea behind these tests is to check for a long-run relationship between the prices of various assets. Any evidence of such cointegration is evidence of predictability and information inefficiency. On the basis of a small number of studies, there is some evidence of cointegration in share prices, particularly for small firms and over longer periods, which might be employable in order to make consistent speculative profits. Chaos theory offers another potential avenue for future research, the essential idea being that nonlinear processes can cause large, volatile movements in asset prices which appear random but are in fact deterministic. There is, however, to date no coherent empirical evidence of chaotic behaviour in financial markets, and the practical value of this approach has yet to be demonstrated.

A succession of tests, therefore, using various statistical techniques, over a wide range of data sets, have failed to reach a consensus on the validity of the efficient markets hypothesis. There does seem substantial evidence, however, that some markets do diverge, at least in defined circumstances, from a random walk specification. Whether any price dependence through time can be used to secure abnormal returns depends on the reason for the dependence, and the costs of implementing a trading strategy based on evidence of return predictability.

A number of such strategies have been advocated. In 'filter' systems, securities are bought or sold according to their patterns of falling and rising prices. The filter is the name given to the percentage change in the security price used to initiate a position. A number of studies of filter strategies produced empirically discernible deviations from what would be expected in markets which were strictly weak form efficient. There was a lack of evidence, however, that these deviations could be used to make above-average returns net of transactions costs. In this sense, the efficient markets hypothesis was not rejected. Early research into the consequences of employing buy or sell strategies based on deviations from a moving average of their prices over various periods similarly failed to identify profitable trading rules net of transactions costs. Some more recent studies of these 'moving-average' rules, and also of the 'trading-range break' rule (buy when the index exceeds its last peak, sell when it falls below its last trough), have demonstrated some support for the view that they can be effective. Similar success has been claimed by authors of other specified technical trading rules. The problem of data-mining and other testing biases has to be considered, however, in evaluating the conclusions of these studies, as well as making allowance for all the risks and costs of implementation.

There is, therefore, some evidence of predictability and dependence in stock price movements through time, evidence which constitutes prima facie evidence of information inefficiency as strictly defined. There is less compelling evidence that this information can be utilized in the market so as to earn abnormal returns. To this extent, the case for the existence of information inefficiency is less strong.

2.4 Empirical tests of semi-strong form information efficiency in financial markets

This section reviews the empirical tests which have been proposed and applied in the literature to investigate the existence of semi-strong information efficiency in financial markets.

In a financial market characterized by strict semi-strong efficiency, prices reflect all publicly available information as soon as it becomes available. In a less strict form of semi-strong efficiency, it is not possible to make above-average or abnormal returns from any divergences between actual security prices and the prices which would obtain if all publicly available information were incorporated into the prices instantaneously and in an unbiased fashion.

Two main approaches have been adopted in the literature as a means of evaluating the extent of semi-strong form efficiency in financial markets. These are addressed in two sub-sections. In sub-section 2.4.1 the impact of new public information on prices is assessed, whereas sub-section 2.4.2 investigates opportunities for identifying particular conditions which might systematically produce the possibility of earning above-average or abnormal returns (so-called market 'anomalies'). Sub-section 2.4.3 presents a summary and conclusions.

2.4.1 The impact of new public information on security prices

Early contributions to the literature, designed to establish whether share prices fully reflect all obviously available public information, concentrated on the phenomenon of 'stock splits' (otherwise termed stock dividends, scrip issues or [in the U.K] capitalization issues).

A pioneering study of this type was undertaken by Fama, Fisher, Jensen and Roll (1969), who examined the market's reaction to 940 such stock splits, and found no evidence which could be used to yield a profitable trading strategy by or consequent upon the time of the stock split. Studies by Pettit (1972) and Charest (1978) of the market reaction to dividend announcements also found generally quick adjustment to the new information. Pettit (1972) examined abnormal daily price behaviour in 135 stocks on the New York Stock Exchange during the days surrounding a dividend announcement, concluding that it would not have been possible to make an abnormal profit by buying or selling after the announcement date. Expanding the sample to include eighteen thousand announcements between 1964 and 1968 (in order to examine the relative performance of stocks during the months surrounding dividend announcements) gave broad, though not total support to the hypothesis that abnormal profits could not be made by buying or selling after the announcement month.

Asquith (1983), however, produced evidence that although the stock prices of acquiring firms in a merger barely move in response to the announcement, they subsequently exhibit evidence of a slow drift downwards. Fama (1991) identified three distinct explanations already extant in the literature for these findings. First, that acquiring firms pay too much for target firms, but that an inefficient market responds to this information rather slowly (Roll, 1986); second, that there is a measurement bias in calculating the abnormal returns (Franks, Harris and Titman, 1991); and third, that Asquith's (1983) findings are sample-specific (Mitchell and Lehn, 1990).

Residual analysis of the effect of various other items of information on share prices provides a broad consensus in favour of the semi-strong form of market efficiency. Research along these lines includes work by Kraus and Stoll (1972) on block trading in the New York Stock Exchange; by Foster (1973) on estimates of earnings per share by company officials; by Waud (1970) on Federal Reserve Discount Rate changes; by Scholes (1972) on secondary market issues; by Brown and Kennelly (1972) on quarterly earnings announcements; by Firth (1976) on earnings of similar-type companies in the U.K.; by Foster (1973) on earnings by similar-type companies in the U.S.; and by Kaplan and Roll (1972), Ball (1972) and Sunder on various changes in accounting procedures.

These early studies taken as a whole thus seem to suggest that market prices broadly reflect and adjust to new and existing published information, and that there is little evidence that a trading rule based on the available public information can be devised so as to provide superior profit performance.

Subsequent investigations also indicate that share prices move rapidly to a new equilibrium value consequent upon the announcement of new information (Patell and Wolfson, 1984; Brown and Warner, 1985)[21]. A survey of daily data studies by Fama (1991) suggests that stock prices seem to adjust within a day to specific event announcements, although the dispersion of returns increases around information events. Direct tests by Colling and Irwin (1990) of the Efficient Markets Hypothesis, undertaken in the U.S. using market survey data in the live hog futures market, provided evidence that futures prices reacted to new information in the way which would be expected if the market were efficient. Thus, the weight of evidence seems to suggest that markets succeed in reflecting most or all obviously available public information, and do so quite quickly. Evidence to the contrary is limited, and of reduced significance if account is taken of the transactions costs involved in any attempt to profit from it. However, the findings are not unanimous. In particular, Bernard and Thomas (1990) produced evidence consistent with a failure of stock prices to reflect fully the implications of current earnings for future earnings. It is as if, they argue, stock prices fail to reflect the extent to which the earnings series of each firm differs from that of a seasonal random walk; specifically, that the market fails to understand the autocorrelation of quarterly earnings, and is, therefore, inefficient.

A somewhat different means of addressing this issue is to link regularities in financial markets with the frequency at which news is reported. Studies by Atkins and

Basu (1991) and by Berry and Howe (1994) adopt this method. A parallel approach is to relate market volatility to the timing of the release of public announcements, such as macroeconomic data and government policy declarations. This is a technique employed by Harvey and Huang (1991), and by Ederington and Lee (1993). Similar work was undertaken by Mitchell and Mulherin (1994), who related aggregate measures of securities market activity, such as trading volume and market returns, to the news announcements of Dow Jones and Company. At the same time they attempted to tackle potential estimation problems identified in earlier research, such as the variation in the importance of news and the endogeneity of news reporting. They found a direct relation between market activity and the number of Dow Jones announcements. Furthermore, their results appear robust to the addition of other factors previously identified as influential on financial markets, e.g. day-of-the-week dummy variables. While there is evidence from these studies that some identifiable relationships existed between public announcements and subsequent market indicators, these relationships did not, however, appear to be very strong. In other cases the relationship did not appear to exist at all. Mitchell and Mulherin (1994), for example, found difficulty in confirming any link between volume and volatility and observed measures of information.

In summary, while evidence does exist to suggest that the market does not always adjust fully and instantaneously to new public information, it is less clear how this can be exploited so as to earn abnormal returns.

2.4.2 Reconciling market 'anomalies' with information efficiency

This section examines whether it is possible to identify particular conditions in a financial market which might systematically offer the opportunity of earning above-average or abnormal returns. Such possibilities are generally termed 'market anomalies'.

Some studies appear to suggest that stocks perform better at particular times of the year, e.g. Bonin and Moses (1974), Rozeff and Kinney (1976), Tinic and West (1984), Keim (1983), Reinganum (1983); or at particular times of the week, e.g. Cross (1973), French (1980), Gibbons and Hess (1981), Rogalski (1984); or that the shares of smaller companies seem to earn a greater amount on average than those of larger companies, even allowing for differences in their risk profiles, e.g. Reinganum (1982), Ibbotson (1990). However, there is less evidence that such information can be turned into profitable trading rules, or where there is evidence, it tends to suggest that the possibility soon disappears. Even so, a review of the field by Fortune (1991) concludes that empirical analysis provides an "overwhelming case against the efficient market hypothesis" (p.34). He cites as evidence such "well-established" anomalies as the 'small firm effect', the 'January' and 'weekend' effects, the 'loser's blessing and 'winner's curse', and the 'closed-end fund puzzle' (see below). These constitute a mere selection of the many 'anomalies' proposed in the academic literature. An interesting example of the others is a 'local weather effect', identified by Saunders (1993), who reported significance at the 0.0001 level for a correlation between the local weather and listed stock prices in New York City!

The more standard 'anomalies' discussed by Fortune (1991) are considered below in greater detail.

a. The small firm effect

The 'small firm effect' refers to the tendency displayed by the common stocks of small-capitalization companies to show unusually high rates of return for much of this century.

Early evidence of this association was reported by Banz (1981), who identified a negative correlation between the average returns to stocks and the market value of the stocks. Controlling for risk, small firms exhibited greater returns than was consistent with their riskiness. In particular, the statistical association between the size of the firm and the average stock return was comparable to that identified by Fama and MacBeth (1973) between average return and risk. Supporting evidence for the existence of this small firm effect was offered by Fortune (1991). Using data in Ibbotson (1990), he calculated and compared the accumulated values of two investments notionally made in January 1926, the first in a portfolio represented by the Standard & Poor 500 and the second in a portfolio of small-firm stocks. He reported that the latter portfolio, in the years since the Great Depression, significantly out-performed the former. Hulbert too, quoted in *Euromoney* (1992b)[22], concluded on the basis of documented research into 'small-cap' stocks (defined as the 20 per cent of companies with the lowest market capitalizations) that such companies showed evidence of outperforming companies with larger capitalizations - the 'small cap effect'. Yet any possibility of outperforming the market is, he argues, limited in this small-cap sector to a small portfolio turnover, owing to the relatively high transactions costs associated with these kind of stocks.

b. The January effect

Another 'anomaly', arguably related to the size effect, is the 'January effect', i.e. that stock performance improves or is unusually good in January. The literature on this can be traced to work on the seasonality of returns by Bonin and Moses (1974) and Rozeff and Kinney (1976). Rozeff and Kinney (1976) reported a 3.5 per cent stock return average in January, compared with 0.5 per cent in other months, a configuration incompatible with a martingale specification. Supporting evidence was offered by Keim (1983), who calculated an average risk-adjusted return to a portfolio of stocks of small firms in various months, concluding that it was significantly larger in January than the rest of the year. Specifically, about one half of the size effect occurred in January, about one quarter of the annual size effect occurring during the first five trading days of that month. Similar evidence is reported by Guletkin and Guletkin (1987) and Lakonishok and Smidt (1988).

On the basis of his own findings, Keim (1983) argued that the January and the small firm effect might be one and the same thing, the January effect appearing only in samples which weighted small and large firms equally, rather than weighting firms in terms of their value. Reinganum (1983) also noted that the January effect seemed to occur predominantly in smaller firms - and moreover, that much of the small firm effect occurs in January. Keim and Stambaugh (1986), Rogalski and Tinic (1986), and Tinic and West (1984, 1986) all linked the January returns to seasonality in the risk-return relationship. In particular, Rogalski and Tinic (1986) showed that the risk, as measured by Capital Asset Pricing Model betas, associated with small firms is greater in January than in any other month. In consequence, "the 'abnormal' return on these stocks may not, after all, be abnormal" (Rogalski and Tinic, 1986, p.63). Tinic and West (1984), re-examining Fama and MacBeth's (1973) findings that riskier stock earn higher average returns for monthly data, concluded that the trade-off was limited to January. Incidentally, Tinic and West (1984) found that this U.S. phenomenon translated into an "April effect" when applied to U.K. data. The idea that this small firm effect can essentially be redefined as a "losing firm" effect is a conclusion supported by De Bondt and Thaler (1985), whose results suggest that 'losers' earn exceptionally large January returns while 'winners' do not. Keim (1989) found an average return of 7.46 in Januarys for a portfolio which used the highest earnings/price ratios and the smallest size, but only 1.39 in other months. Moreover, the bottom twenty per cent of companies in terms of the market value of their equity out-performed the Standard & Poor index by 5.5 per cent for Januarys from 1926 through 1986, under-performing in only seven of these years. Similar findings are reported in Ikenberry and Lakonishok (1989). A test of the 'January effect', quoted in the *Wall Street Journal* (1992)[23], lends support to the hypothesis. In that report, an examination of the industrial average in the eleven years from 1980 to 1990 (inclusive) indicated an improvement in seven of them, an effect particularly marked in the case of the U.S. Nasdaq Composite Index of smaller stocks, these outperforming the Dow industrials in five of the seven "up" Januaries. In the years in which the Dow industrials showed a January downturn (1981, 1982, 1984), the small stocks declined further. Taken in aggregate over the eleven-year period studied, the industrials showed a 27 per cent rise, and the Nasdaq Composite Index a 38 per cent improvement.

One explanation of the 'January effect', common in the literature, centres on the existence of tax-loss selling at year end, e.g. Branch (1977), Dyl (1977), Reinganum (1983), Roll (1983), Rozeff (1985), Chan (1986), Griffiths and White (1993). The idea is that some investors will sell securities at year-end in order to institute short-term capital losses for income tax purposes.

A related 'anomaly' is a 'turn-of-the-year effect', in which lower capitalized small stocks outperform higher capitalized large stocks. Ziemba (1994) has examined this effect in some detail, and recommends that investors employ a futures spread which is long in small stocks and short in large stocks. Such a strategy has the advantage of low costs of implementation. Moreover, Ziemba claims to have made consecutive profits for each of the eleven years between 1983 and 1993 by using this approach with money in his own mutual fund.

c. The weekend effect

The original 'weekend effect', traceable to findings by Cross (1973), is the proposition that large stock market decreases tend to occur between the Friday close and the Monday close. Since Cross's seminal findings, French (1980), Lakonishok and Levi (1982), Keim and Stambaugh (1984), Rogalski (1984), Jaffe and Westerfield (1985), and Harris (1986) have all found evidence that U.S. stock returns are, on average, negative from the close of Friday trading to the opening of Monday trading.[24]

Although the January and weekend effects, and the turn-of-the-year effect are the best documented of the calendar effects, they are by no means the only ones. Other well-documented 'calendar anomalies' are a 'turn-of-the-month' effect (see Ariel, 1987) and a 'holiday effect' (see Lakonishok and Smidt, 1988; Ariel, 1990; Liano, Marchand and Huang, 1992; Wilson and Jones, 1993; Liano and White, 1994).[25]

d. The winner's blessing and loser's curse

The so-called winner's blessing and loser's curse 'anomalies' were first developed by De Bondt and Thaler (1985, 1987). The basis of these ideas can be found in seminal work by Kahneman and Tversky (1973), which reported that individuals, in revising their beliefs, tend to overweight fresh information and underweight prior data. To test this hypothesis for those active in financial markets, De Bondt and Thaler (1985) used monthly return data for New York Stock Exchange (NYSE) common stocks for the period between January 1926 and December 1982. They concluded that a stock selection strategy based on this hypothesis could yield large abnormal returns. Specifically, they found that 36 months after portfolio formation, portfolios of prior 'losers' (i.e. stocks that have experienced a recent reduction in their price/earning ratios) had earned about 25 per cent more than those of prior 'winners'. Over five-year test periods the portfolio of losers outperformed the portfolios of winners by an average of 31.9 per cent. Moreover, since the strategy is based on past returns only, it also contradicts the weak form of market efficiency.

Support for this hypothesis can be found in an analysis by Dark and Kato (1986) of the Japanese stock market for the years 1964 to 1980, which revealed that the three year returns for portfolios of extreme previous losers out-performed those of extreme previous winners by an average of seventy per cent. Dyl and Maxfield (1987), using a random sample of two hundred trading days between January 1974 and January 1984 for NYSE (New York Stock Exchange) and AMEX (American Stock Exchange) stocks, offer further support. In particular, they found that the three biggest losers in any particular day experienced an average risk-adjusted return of plus 3.6 per cent over the following ten days, whereas the three biggest gainers experienced an average loss of 1.8 per cent over a comparable period.

Using data from the German stock market for a period from 1973 to 1989, Stock (1990) also reported evidence of long-term over-reaction, although he noted that in the short run the extreme stocks in his data set showed a strong tendency to continue their initial performances. Jegadeesh (1990) and Lehmann (1990) both found significant abnormal returns resulting from contrarian strategies which selected stocks on the basis of returns in the previous week or month. Brown and Harlow (1988) also identified a 'magnitude effect', i.e. a tendency for the most extreme initial winners and losers to show the most extreme ensuing price reversals.

An explanation of the winner-loser 'anomaly', offered by Vermaelen and Verstringe (1986), along with Chan (1986, 1987), is that it is no more than a rational response to changes in risk. In particular, Chan (1986, 1987) argues that a fall in stock prices leads to an increase in debt-equity ratios and risk (as measured by Capital Asset Pricing model betas), and vice-versa. In consequence, a higher return is required to compensate for the higher risk incurred in buying losers compared with winners.

e. The closed-end fund puzzle

Closed-end mutual funds are distinguished from the more common open-end funds in that a fixed number of shares are issued, trading in which is between investors who already have shares, i.e. a shareholder must sell his shares to someone else in order to liquidate a holding. Closed-end funds thus trade in secondary markets. The net asset value (NAV) of these funds is the market value of the securities portfolio, net of liabilities. Unlike open-end mutual funds, which sell and redeem shares on the basis of this prevailing net asset value, both the market value of these funds' assets and the market price of its shares are observable and can differ from one another. In an 'efficient market,' one might expect this market price to equal the net asset value. In fact, empirical evidence indicates otherwise, suggesting that short of jettisoning the efficient markets hypothesis we require an explanation in terms of characteristics unique to the fund. Otherwise the opportunity exists for arbitrage which would eliminate the difference. The term commonly used to describe this 'anomaly', first coined by Lee, Shleifer and Thaler (1990, 1991), is the "closed-end fund puzzle." Recent evidence of a closed-end fund effect can be found in Pontiff (1995).

Leonard and Shull (1996) used evidence of a January effect in the returns of closed-end funds (as well as small firms) to suggest that the 'anomaly' has its origins in tax motivations possessed by individual investors who operate in this market. Arak and Taylor (1996) addressed the separate issue of the gain to be made from a strategy of switching from equities into closed-funds at times of a large discount and reversing the process when the discount reduces. Using Monte Carlo simulations, they found that the returns to this strategy were large, even after allowing for the systematic risk of closed-end country funds. This abnormal return, they conclude, is further to that which may be earned (by holding the fund) on the stock portfolio itself.

Thus it would appear that a wide body of prima facie evidence exists to support a contention that assets perform (in the sense of offering a return to a given outlay), or at least have at some time performed, systematically better at particular clearly defined times; and that certain types of particular clearly defined assets perform systematically better than others. Some evidence suggests that profitable trading rules can, or once could, be constructed on the basis of such information so as to yield systematically abnormal returns. Others, while less convinced or unconvinced about this possibility, offer the 'existence' of the 'anomalies' as evidence of market inefficiency.

The relevance of these findings for the efficiency debate turns ultimately on three questions, i.e. are the findings statistically significant, are they statistically valid, and do they simply compensate for systematic biases elsewhere?

The breadth and depth of the evidence would seem to offer overwhelming support for its statistical significance. The justification for its statistical validity is less clear-cut, involving judgements of whether, for example, sufficient explanatory variables have been used, and whether any selection bias has occurred. Yet the most wide-ranging criticisms of the orthodox anti-efficiency case centre on the last issue. Does a systematically higher return by certain stocks or at certain times simply compensate, for instance, for greater risk, or less information, or the necessity for making more complex or more time-consuming decisions? Apart from the difficulties of identifying and measuring these factors, there is the problem of assessing the appropriate balance to be applied to the trade-offs.

2.4.3 A summary of empirical tests of semi-strong form information efficiency in financial markets

Two main approaches have been adopted in the literature as a means of evaluating the extent of semi-strong form efficiency in financial markets. One is to assess the impact of new public information on prices. The other is an exploration of opportunities for earning systematic abnormal returns on the basis of identifiable circumstances (so-called 'market anomalies').

Semi-strong form tests, designed to establish whether share prices fully reflect all obviously available public information, have broadly concentrated on market reaction to fresh public announcements, of new issues or dividends or earnings for example.

While evidence does exist that the market does not always adjust fully and instantaneously to new public information, there is very limited evidence that it is possible to exploit this so as to earn abnormal returns. This is particularly true without a capacity to react very quickly in real time to the publication of such information, and especially so net of transactions costs.

The other method of testing for semi-strong efficiency is to examine whether it is possible to identify particular conditions in a financial market which might systematically offer the opportunity of earning above-average or abnormal returns. Such possibilities are generally termed 'market anomalies'.

For instance, do stocks perform systematically better at particular times of the year, or at particular times of the week, or is the market return linked in any systematic way to the size of the company. A number of these 'anomalies' have been identified, some of which have not been replicated, some of which have not been confirmed out-of-sample, and some of which are found in a number of studies but rejected in others. There remains a general consensus that there exists or has existed a 'January effect' (shares perform systematically better on average in January) and a 'small firm effect' (the shares of smaller companies perform systematically better on average than do those of large firms). Other interesting 'anomalies', both of them supported by a weight of empirical evidence, are the 'winner's blessing/loser's curse', and the 'closed-end fund puzzle.' The basis of the 'winner's blessing/loser's curse' anomaly lies in the idea that individuals, in revising their beliefs, tend to overweight fresh information and underweight prior data, so that investors who act in a contrary fashion to this prevailing psychology can on average earn an above-average return. Closed-end mutual funds are distinguished from the more common open-end funds in that a fixed number of shares are issued, trading in which is between investors. Both the market value of these funds' assets and the market price of its shares are observable, and in an efficient market one might expect these to converge. The 'closed-end fund puzzle' is the finding that these not only can but do differ systematically from one another.

The issue has developed into the question of whether the identified possibility of earning above-average returns from systematic exploitation of genuine 'anomalies' implies the opportunity to earn abnormal returns. Ultimately, the question is not simply empirical, but turns on the weight attached to the costs of purchasing and trading in one set of stocks compared to another. Market efficiency would imply that any empirically observed systematic divergences between the returns to stocks or portfolios of stocks is fair recompense for additional observed or unobserved costs of trading in these particular assets compared to others. Until and unless a general method of evaluating these costs can be agreed and implemented, the issue is ultimately unresolvable.

2.5 Empirical tests of strong form information efficiency in financial markets

This section reviews the empirical tests which have been proposed and applied in the literature to investigate the existence of strong information efficiency in financial markets.

In a financial market characterized by strict strong form efficiency, security prices reflect ALL available information, public and private, as soon as it becomes available. In a less strict form, it is not possible to make above-average or abnormal returns from any divergences between actual security prices and the prices which would obtain if all available information were incorporated into the prices instantaneously and in an unbiased fashion.

Two main approaches have been adopted in the literature in an attempt to measure the extent of strong form efficiency in financial markets. The first is to assess the impact of identifiable monopolistic access to information and assess the impact of this insider knowledge on profitability. Because of the legal implications of overt trading on the basis of insider information, there are of course inherent difficulties in identifying such trading for purposes of evaluation. These studies will be examined in the first part of this section (2.5.1). Following on from this, a second approach will be considered (in sub-section 2.5.2) which involves assessing the performance of individuals and organisations, in particular professional forecasting services, in order to assess whether they have access to private information not reflected in stock prices.

2.5.1 The effect of inside information on expected returns

Niederhoffer and Osborne (1966) offered seminal findings that specialists on major security exchanges in the U.S. have monopolistic access to information which could be used to derive profitable trading rules, a state of affairs which they explain in terms of the market structure of the New York Stock Exchange. Any use of this informational monopoly power to make a systematic abnormal return is evidence of a strong form market inefficiency.

Scholes (1972) presented evidence of monopolistic access to information by corporate insiders about their firms, information which was reportedly not reflected in prices. Similar work was conducted by Finnerty (1976). Collins (1975) looked at SEC (Securities and Exchange Commission) product-line reporting, concluding that information available privately, though not publicly, on historical segment prices in his data could be used to yield above-average returns.

An examination by Lorie and Niederhoffer (1968) of "insider trades" reported in the official summary of stock reports of the U.S. Securities and Exchange Commission also produced evidence of performance superior to the market average, conclusions in line with work by Pratt and de Vere (1968), and by Jaffe (1974). Jaffe found evidence that outsiders can also profit from information publicly available about insider behaviour, for up to eight months after the information becomes public. If true, this would constitute evidence of a semi-strong form inefficiency. Seyhun (1986), however, offered an explanation of Jaffe's findings in terms of a size effect, smaller firms (exhibiting relatively higher returns - see Banz, 1981) displaying a bias to "insider buying" compared to larger firms (which favour "insider selling"). Rozeff and Zaman (1988) suggested that these studies are in any case flawed, by failing to take account of biases in the calculations which could be caused by size and earning/price ratio effects. Controlling for these variables, they found, on the basis of a data sample of New York Stock Exchange (NYSE) issues between 1973 and 1982, a zero or negative return to outsiders (net of transactions costs), and only a small return to insiders in the aggregate (three per cent per annum, after deducting an assumed two per cent transactions cost). An examination of OTC/NASDAQ securities by Lin and Howe (1990) found that transactions costs could reduce or even eliminate the possibility for insiders to earn excess risk-adjusted returns from an active trading strategy, while Jarrell and Poulsen (1989) suggested that much of the evidence adduced to demonstrate insider trading is in reality simply evidence of the legitimate influence on the market of pre-bid media rumours. Other work has asked how far insiders possess valuable information, unknown to outsiders, which would cause any of their actions, where publicly known, to signal information about their company. Grammatikos and Saunders (1990) is an example of this approach, reporting a positive effect in the banking sector. These findings are supported, albeit subject to identified qualifications, by Madura and Wiant (1995).

"...insider buy transactions appear to contain favourable information for banks that did not recently experience adverse valuation effects." (p.227).

The use of such information early so as to earn abnormal profits has been reported in Damodaran and Liu (1993) who found, in a study of real estate investment trusts, that insiders traded on internal appraisal information in the time between the appraisal and its public announcement, sufficient to elicit significant abnormal returns. Sivakumar and Wagmire (1994) also offered evidence, based on an identification of the relationship between insider trades and quarterly earnings announcements, of abnormal profits to insider trading. Lustgarten and Mande (1995) produced results indicating increased insider purchases (sales) prior to the announcement of good (bad) earnings news. This is consistent with Pettit and Venkatesh's (1995) finding of anticipatory insider trading in equity markets, and Allen and Ramanan's (1995) study of reportable insider trading and annual earnings figures for a large sample of firms (1978-87). Detta and Iskandardatton (1996), by investigating price reactions in response to the publication by the *Wall Street Journal* of insider transactions, also document significant information content in trading by registered corporate insiders, for the bond and stock markets.

In summary, there is some evidence that insiders can earn abnormal returns, but these returns may be somewhat limited net of operating costs. Similarly, the information which can be gleaned by outsiders from the public behaviour of insiders may be somewhat restricted, both in scope and value. Even so, none of these studies serve to contradict the presumption that individual cases of trading on the basis of inside information can and do yield large abnormal returns to such individuals, nor that some returns may have escaped analysis.

"Some insider trades may be hidden from the SEC and not reported. Some profits may go undetected- namely, those from trading in shares of other companies or those garnered through arrangements in which inside information is shared with others. In other words, it is possible that, even though corporate insider trading profits net of transactions and other costs appear to be zero based on the reported trades of corporate insiders, they are really not zero." (Rozeff and Zaman, 1988, pp.39,40).

2.5.2 The performance of professional forecasting services

The idea of using trading rules in order to generate abnormal returns can be traced in the academic literature to work by Wyckoff (1910), Schabacker (1930), Gartley (1930) and Neill (1931).[26]

The first major study to analyze the performance of professional forecasters per se, however, was published by Cowles (1933), who concluded that the recommendations of major brokerage houses failed to outperform the market. Much more extensive studies by Ambachtsheer (1972, 1974) of the forecasting ability of market analysts concentrated on comparing rankings of stock in terms of prospective performance by professional analysts with the actual outcome some time later. In each case they found evidence of a significant, albeit small degree of forecasting ability. A later study by Ambachtsheer and Farrell (1979) into the performance of investment advisory services produced similar results.

An important contribution to the literature is associated with Dimson and Marsh (1984), who in an analysis of brokers' and analysts' unpublished forecasts of U.K. stock returns, undertook a survey of thirty published studies, involving 47053 "investment tips" by over two hundred advisory firms over a fifty year period (1933 to 1984). They found an average gain of 1.5 per cent by the day after publication, the longer term abnormal gain averaging 0.6 per cent (over periods typically between a quarter and a year). Although statistically significant, thereby providing evidence of at least a small degree of forecasting skill, most of the information content appeared to be impounded into share prices by the end of the publication day. Moreover, such returns did not take account of dealing costs, these returns working out to be less, they calculated, than the round trip costs incurred.[27]

Dimson and Fraletti (1986) focused on the value of verbal recommendations made prior to, or in the absence of publication. Specifically, they investigated the profitability of following the telephone recommendations given daily during 1983 by a leading U.K. stockbroker (1649 recommendations for 90 different companies). They concluded that these verbal recommendations were of similar value to written advice.

"Neither the freedom given to the brokers to choose their own time to favour a stock, nor the focus on unpublished advice, led to any proof of marked outperformance" (p. 157).

A report by *Insurance Age* (1988)[28] is also unencouraging, concluding that only one in three U.K. equity fund managers were able to match the return of the FT (Financial Times) All-Share Index over a five-year period. Rugg (1986), however, reported evidence that active selection among managed funds can be profitable. In particular he recommended, subject to stipulated caveats, investing in aggressive-growth equity funds which are top-ranking performers in the most recent phase (one to six months) of a bull market, while a systematic study of the equity markets over a number of decades by Pratten, quoted in the Financial Times (8/11/93, p.22), led him to conclude that some institutions had outperformed the market for quite long periods. Investors like Warren Buffet, quoted in the *Economist* (5/12/92, pp.21-23)[29] have also reportedly made consistent profits from "value investing", i.e. buying shares with a below-average ratio of market price to book value.

Studies such as those of Grinblatt and Titman (1992), Brown, Draper and

Mackenzie (1993), and Goetzmann and Ibbotson (1994) also identify consistency in the performance of managed funds. Indeed, Hendricks, Patel and Zeckhauser (1993) concluded that some American mutual fund managers were able to successfully out-perform the market (possessing what the authors term "hot hands"), at least for a limited period. Specifically, they examined quarterly returns between 1974 and 1988 inclusive, from a sample of open-end, no load, growth-oriented equity funds. They found that those funds which performed relatively well (compared to other similar funds) in the most recent years, also exhibited superior performance in the near term thereafter, the "near term" being identified as one to eight quarters. Moreover, those funds which under-performed in this regard continued to under-perform both in the short-term and in the longer-term. They concluded that the "icy hand" is more sustained than the "hot hand." To counter the possibility that the results are due to known anomalies in the sample, they constructed a sample specifically to avoid problems of survivorship bias [a potential problem identified by Brown, Goetzmann, Ibbotson and Ross (1992)], and tested for other possible influences. They concluded that "superior performance is also achieved relative to an eight-portfolio benchmark that accounts for effects of firm size, dividend yields, and reversion in returns." (p.122). Their finding of "hot hands" benefits confirmed results they reported for a different sample covering mutual fund performance for the eight quarters during 1989 and 1990.[30] Malkiel (1995), however, conducted an examination of all equity mutual funds between 1971 and 1991 and found substantial evidence of hitherto unreported survivorship bias. He claimed that statistical evidence that mutual fund returns are predictable on a period-to-period basis are not robust. Moreover, they do not outperform the market.

Lakonishok, Shleifer and Vishny (1992) offered ambivalent backing for the idea of persistent over-performance. They found that in the period from 1983 to 1989, managers running equity portfolios for a sample of 769 American pension funds underperformed the Standard and Poor 500 index, even before subtracting their fees, by 2.6 per cent after weighting by size (1.3 per cent using unweighted figures). Even so, those pension fund managers who performed relatively well in one three-year period had a better than chance probability of performing relatively well in the next period, although they did not do well enough to score above-average returns after allowing for fees. Restricting the analysis to those managers whose performances over the previous three years placed them in the top quartile of all pension funds, they found a performance 2.1 per cent superior to that by managers in the bottom quartile. These results are broadly comparable with the results of a twelve-year study published in the Hulbert Financial Digest into the performance of investment advisory letters.[31] Letters out-performing the market over one three-year period achieved better results in the following three-year period than those under-performing the market in the same initial term, by an average of 2.8 per cent per year. Applying the test to six-year periods yielded even more significant differences, letters out-performing the market over this expanded time scale making 5.1 per cent more in the subsequent six-year period than those under-performing the market. Overall, of 67 investment advisory letters tracked over the six-year period from mid-1986, fifteen out-performed the market on average.

An extended examination by Sirri and Tufano (1993a, 1993b) of a large sample of equity mutual funds reported evidence that although funds and management houses which performed relatively well did tend to attract more assets, the relationship was asymmetric. In other words, good performance was rewarded to a greater degree than bad performance was punished.

A smaller scale test of forecasting performance was offered by Allen and Taylor (1989, 1990), in a Bank of England study of a select group of twenty foreign exchange chartists between June 1988 and March 1989.[32] The authors found 50 per cent accuracy in predictions of the direction of currency movements over a one-week period, and between 46 and 49 per cent accuracy over a four-week period. In particular, the 'experts' tended to miss turning points. Moreover, a rise in one period led them on average to expect a smaller rise in the next, while a fall led them to expect a smaller fall. They concluded that, taken as a whole, these currency forecasts were no better than a random walk.[33] They noted, however, evidence of variations around this average, concluding that in terms of forecasting the direction of future rates the best among those studied out-performed a wide range of exchange forecasting methods. In terms of forecasting actual rates, however, only one of the chartists consistently out-performed a random walk. Even so, an analysis of 200 questionnaires returned by foreign exchange dealers revealed that 90 per cent claimed to pay attention to chartists in predicting up to one week ahead. Greater than 60 per cent viewed charts as at least as useful as fundamentals.[34]

In a London School of Economics (LSE) laboratory study, conducted by Curcio and Goodhart, quoted in the *Economist* (5/10/91, p.84), sixty LSE students took part in an experiment in which they traded (hypothetically) in one or more of nine assets, ranging from the FTSE-100 index to U.S. bonds. Half the students were in possession of a chartist package from the London firm of Fiamss, the others were limited to information on the price history of each asset. Actual reward was linked to hypothetical returns. There was no evidence of a significant difference in the performance of the two groups, a result repeated when the experiment was re-run using 24 foreign exchange dealers. Moreover, the foreign exchange dealers performed barely any better than the students.

Even so, the authors did find evidence that those who used charts to 'trade' performed more similarly as a group than the rest of the sample.[35] A report on this study, in the *Economist* (5/10/91, p.84), concluded thus:

"So charts may be worthwhile for cautious firms, which do not pine for the profits of a trading ace so long as they never have to suffer the losses of a trading fool. Either that, or chartism thrives because there's one born every minute."

De Bondt and Thaler (1990) tested whether the experimental 'over-reaction' noted by Kahneman and Tversky (1973) for student subjects, applied also to stock market professionals. Using analysts' earnings forecasts taken from the Institutional-Brokers-Estimate-System tapes between 1976 and 1984, they concluded that "...the same pattern of overreaction found in the predictions of naive undergraduates is replicated in the predictions of stock market professionals. Forecasted changes are simply too extreme to be considered rational." (p.57).

Evidence supporting the view that particular identified fund managers can outperform a prima facie random walk interpretation is, however, offered in a report in the *Economist* (8/8/92, pp.67-68)[36], of the performance of the Boston-based fund-management firm, Fidelity. In particular, in an analysis undertaken by Lipper Analytical Services, Fidelity's Magellan equity fund out-performed the Standard and Poor 500 by 27 per cent in the year to June 1978, 21 per cent in 1979, 26 per cent in 1980, 51 per cent in 1981 and 53 per cent in 1983, though its subsequent performance waned, in later years barely out-performing the S and P 500. One explanation offered to explain the earlier success centres on the riskiness of Magellan's investments. An analysis by Rekenthaler, of specialist research firm Morningstar, quoted in the same report, found a significantly higher variation in the month-on-month returns of Magellan investments in the period

1977 to 1983 compared with both the Standard and Poor 500 and other funds possessing a similar profile of small-firm investment.

The excess returns documented by, for example, Bjerring, Lakonishok and Vermaelen (1983), Copeland and Mayers (1982), Stickel (1985), Huberman and Kandel (1987, 1990) for U.S. investment advisory service Value Line seem somewhat more difficult to explain away. The Value Line Investment Survey[37] produces reports on 1700 publicly traded firms, ranking their stocks on a scale from one to five in order of their desirability of purchase (their "timeliness"). Studies dating back to Black (1973) reported that the higher ranked stocks generated significantly higher returns than those lower ranked, a conclusion which if not produced by chance might most obviously be explained in an efficient market by variations in other costs, such as the level of risk associated with different rankings. However, Black found that the mean beta coefficients (i.e. measures of risk) were the same for all rankings, concluding that the service provided genuine predictive value. Holloway (1981) compared the results of active and passive trading strategies based on the Value Line rankings, an active strategy being defined as one of changing stocks before a year end if and when it is downgraded in rank, a passive strategy being one of buy-and-hold. Although the active strategy produced higher returns than the passive approach, the advantage was reversed net of associated transactions costs. Even so, the Value Line Ranking system provided profitable information, even allowing for transactions costs and risk. Stickel (1985) also identified information contained in the Value Line rankings not reflected in prices, although the information in Value Line rank changes was stronger for smaller stocks.

Fama (1991) placed such findings within a more general theoretical perspective, proposing that they are compatible with Grossman and Stiglitz's (1980) "noisy rational expectations" model of competitive equilibrium. He argued that "... because generating information has costs, informed investors are compensated for the costs they incur to ensure that prices adjust to information. The market is then less than fully efficient (there can be private information not fully reflected in prices), but in a way that is consistent with rational behaviour by all investors" (p. 1605).

An explanation offered by Lee and Park (1987) contradicted earlier findings by Black (1973), in reporting evidence that the beta (risk) coefficients of stock were after all inversely related to their Value Line rank, i.e. higher ranking stocks tending to be associated with higher risk profiles. Affleck-Graves and Mendenhall (1992) offered an alternative explanation of the 'abnormal' returns generated by the Value Line ranking system as no more than the post-earnings-announcement drift already documented by, for example, Ball and Brown (1968), Bernard and Thomas (1989), Abarbanell and Bernard (1992). The reason, they suggest, is that Value Line rank changes follow closely upon recent earnings surprises. After controlling for earnings surprises, they found no significant abnormal returns associated with the Value Line rank. Peterson (1995) acknowledged the presence of post-earnings announcement drift, but contended that it is too small to explain the short-term abnormal returns which can be made from the Value Line announcements.

Not all studies, however, have supported the persistence or even existence of a genuine Value Line anomaly. Hulbert (1990), for example, found a weakening of Value Line's Group 1 stocks after 1983. Keane (1991) made the more general assertion that the significance of the "Value Line enigma" has diminished over time as the research methodology applied to it has become increasingly more refined. These conclusions are supported by Chandy, Peavy and Reichenstein (1993), who reported that a significant three-day return which they found to a weekly Value Line "highlights" announcement, was largely reversed over a short subsequent period.

An interesting perspective on this whole issue is associated with the empirical studies of, for example, Stael von Holstein (1972) and Yates, McDaniel and Brown (1991), which suggest that so-called 'experts' are not in fact able to outperform a random dart-throwing approach to stock- picking. The view was given some additional support by a *Wall Street Journal* (1989)[38] analysis of stocks recommended by investment professionals in its column's stock-picking contest over a period of a year. This indicated a worse performance in four of the twelve months examined compared with a random 'dart-tossing' approach. Taken as a whole, however, their results did show a 66.537 per cent annual return for the investment professionals (ignoring dividends, taxes and commissions), compared with 12.383 per cent from a dart-throwing approach, and 27.446 per cent from tracking the Dow Jones Industrial average. Any boost to the professionals' fortunes from a publicity effect is more difficult to assess. In an assessment of investment dartboard columns since new and still current rules were adopted in 1990, and taking account of price changes only (i.e. ignoring dividends), Dorfman (1993) produced evidence that the professionals out-performed the darts on 24 occasions, compared to seventeen successes for the darts. The average six-month gain for the professionals was 8.4 per cent, compared to 3.3 per cent for the darts.

In an analysis of dartboard contests surveyed between January 1990 and December 1992, Metcalf and Malkiel (1994) reported that the experts beat the market eighteen times out of thirty (yielding a total return of 9.5 per cent), while the darts beat the market fifteen times (yielding a total return of 6.9 per cent). The stock chosen by the professionals also out-performed the market, as proxied by the S & P 500 stock index.

Even so, Metcalf and Malkiel (1994) failed to reject the hypothesis that the experts won by chance at conventional levels of significance. The "superior performance" by the professionals is in any case explained by Metcalf and Malkiel (1994) as a consequence of the tendency of the 'experts' to choose riskier, more volatile stock than would a random approach, and to a favourable publicity or announcement effect. The stock chosen by the professionals was in fact 40 per cent more volatile than the market, compared to a 6 per cent greater volatility displayed by the darts. Adjusting for risk, they concluded that the margin of superiority exhibited by the professionals fell to the statistically insignificant figure of 0.4 per cent, and disappeared altogether when allowance is made for any announcement effect. Critics of this approach claim that the professionals are given an unfair task in the rules of the contest, in particular the stipulation that the base line is taken at 4 p.m., when most of the gains from professional tipping will already have been realized. Moreover, the six-month contest period is sufficient time, it is argued, to obviate any persisting announcement effect.

A different though related investigation was offered in an *Investors Chronicle* (1991)[39] examination. This study compared the performance of a 'dartboard' selection of three portfolios of 25 shares each, chosen from the London Service pages of the Financial Times, with the FT All-Share Index. In each case, the FT All-Share index significantly out-performed the random portfolio compilation.

Other evidence was offered by Sundali and Atkins (1994), who produced evidence that the 'experts' in their study sample did outperform both darts and market averages. They also found, however, that no particular class of expert was able to consistently outperform any other.

The return to following the advice of 'experts' has recently been examined by

Zivney, Bertin and Torabzedeh (1996). They found evidence of different reactions to different pieces of professional advice, even on the same page of the *Wall Street Journal*. However, their most interesting conclusion is that the market appears to over-react to "rumours" published in that Journal's "Abreast of the Market" (APTM) column, at least in their data set taken from 1985 to 1988. Indeed, trading on these over-reactions would, they report, have permitted a twenty per cent annual excess return. Other evidence on the issue has been produced by Sant and Zaman (1996) and by Womack (1996). In a study of stocks mentioned in the columns of Business Week magazine, Sant and Zaman (1996) identified short-term positive abnormal returns for stocks tipped by a limited number of analysts (less than twenty, and the less the better), although a subsequent negative bias in the returns was sufficient to offset any positive effect by the end of a six-month period. Womack (1996), however, used a new data source (created by First Call Corporation of Boston) which is able to provide (at a cost) the exact date and approximate time that information is made available to investors. This study concluded that both buy and sell recommendations have a substantial (and, for the period studied, non-reverting) impact on stock prices, an effect which continues to influence prices (at least for sell recommendations) for a period of months.

2.5.3 A summary of empirical tests of strong form information efficiency in financial markets

In a financial market characterized by strict strong form efficiency security prices reflect ALL available information, public and private, as soon as it becomes available. In a less strict form, it is not possible to make abnormal returns from any divergences between actual security prices and the prices which would obtain if all available information were incorporated into the prices instantaneously and in an unbiased fashion.

One way of measuring the extent of strong form efficiency in financial markets is to assess the impact of insider knowledge on profitability. The second approach is to assess the performance of professional forecasters, as a method of indicating whether they have access to private information not reflected in stock prices.

Because genuine insider activity is illegal it is, of course, difficult to measure directly. Most studies examine reportable insider trades, or else adopt an indirect approach, for example by seeking to identify unusual trading patterns prior to the announcement of new information. Although there is evidence that above-average returns can be earned by insiders, there is only limited support for the view that, net of operating costs, these are sufficient to constitute serious abnormal profits. Similarly, the evidence tends to suggest that the information which can be gleaned by outsiders from the public behaviour of insiders is positive but somewhat restricted, both in scope and value. It is, nevertheless, very possible that the most successful insider trades escape detection altogether.

Examination of the performance of professional forecasters have taken the form of empirical tests which are applied to a sample of forecasts. The results are examined for evidence of significant above-average returns, and these findings are interpreted for the light they throw upon the possibility for identified groups or individuals to earn abnormal returns. Early studies of the forecasting ability of market analysts concentrated on comparing rankings of stock, as judged by the analysts, with the performance of the stock, actually ranked, some time later. These studies produced evidence of some, albeit a small, degree of forecasting ability, as did some analyses of the professional buy and sell recommendations. Allowing for risk, transactions costs (including bid-ask spreads), management salaries, etc., other studies revealed evidence of under-performance by "experts" relative to the market. There was a general consensus, in any case, that most of the information content in the recommendations of professional advisors was rapidly impounded into share prices, and any gain was often less than the costs involved. Evidence has been provided, however, of a 'hot hand' effect, i.e. advisors who performed relatively well (compared to other similar funds) in the most recent years, also exhibited superior performance in the near term thereafter, and vice-versa (an 'icy hand').

Smaller scale tests of forecasting performance have also been undertaken at experimental level. In these laboratory studies the forecasts of experts were compared with those of non-specialists, the results showing no significant differences. Indeed, evidence of identical 'anomalous' behaviour, in particular over-reaction to recent data, was found in both groups.

Excess returns have, however, been recorded in some cases on a systematic basis. The most well-documented is the forecasting performance of the U.S. investment advisory service, Value Line. A number of studies have reported that their higher ranked stocks generated over a long period significantly higher returns than those lower ranked. The implications of these findings have to be judged in the context of other studies which ascribe the performance to variations in other costs, such as the level of risk associated with different rankings.

One other challenging and radical approach to the question of forecasting ability has been proposed by some advocates of the efficient markets hypothesis. It is based on the idea that, in a truly informationally efficient market, so-called 'experts' should not be able to outperform a random dart-throwing approach, and has served to popularize the debate. A number of empirical studies and challenges have developed from this, some of which have shown an under-performance by the 'experts' in individual periods, though not overall. This evidence of superior performance by professional forecasters over a sustained period has, however, been rejected by some economists as either not significant, or else simply a fair higher return for the higher risk profiles of the stocks compared to a random approach. There is also the suggestion of a boost to the professionals' performance in the form of an announcement effect.

A compromise position to all such studies of information efficiency is perhaps offered by Womack (1996);

"The [positive] returns I document [to prior recommendations] ... are consistent with the expanded view of market efficiency suggested by Grossman and Stiglitz (1980): that there must be returns to information search costs. These information search costs are often assumed to be zero when considering the efficient market hypothesis. The nontrivial magnitude of the returns reported here challenges the innocence of that assumption." (Womack, p.165).

ENDNOTES

1. "The French word martingale refers to Martigues, a city in Provence. Inhabitants of Martigues were reputed to favour a betting strategy consisting of doubling the stakes after each loss so as to assure a favourable outcome with arbitrarily high probability." (Le Roy, 1989, p.1588).

2. The link between capital market efficiency and martingales can be traced to work by Samuelson (1965). Samuelson held that the martingale model is satisfied if agents have common and constant time preferences, have common probabilities, and are risk-neutral.

3. By economic profits is meant the risk-adjusted returns net of all costs.

4. Le Roy (1989) proposed that a market is efficient if, given transactions costs, no agent is able to earn returns in excess of the opportunity cost.

5. Fama (1991) also proposed replacing the traditional terminology, i.e. "semi-strong" and "strong" form tests of efficiency, with new terms, i.e. "event studies" and "tests for private information."

6. Mean-reversion in stock prices is the tendency for these prices always to revert to some fundamental level. A useful survey of the literature on mean reversion can be found in Forbes (1996).

7. A useful and more detailed review of this issue can be found in Engel and Morris (1991).

8. Fisher (1966) argued that spurious positive serial correlation in portfolio returns, caused by nonsynchronous closing trades for the portfolio securities, was especially significant in portfolios weighted toward small stocks.

9. Working (1960) identified potential problems in using serial correlation analyses on indices. In particular, he demonstrated that in a random chain the correlations of first differences of averages can lead to correlations which do not exist in the original data. A further discussion of issues related to non-synchronous trading can be found in Cohen, Hawawini, Maier, Schwartz and Whitcomb (1980).

10. The idea is that if two stocks X and Y have independent returns, but X trades less often than Y, then the price of Y will react more swiftly to news affecting both, the result of which being that the return to X will seem to react with a time delay to the return to Y. These positive serial correlations measured across a number of stocks will show up as positive autocorrelation in an index of such stocks. Such positive autocorrelation in the index may even mask negative autocorrelation in the returns to an individual stock.

11. Granger (1992) summarizes that if series are cointegrated there is necessarily an error-correction data generating model of the form (ignoring constants):
$\Delta X_t = \alpha_x Z_{t-1}$ + lagged ΔX_t, ΔY_t terms + white noise,

plus a similar equation for ΔY_t, with at least one of α_x, α_y being non-zero.

Cointegration between a pair of stock prices would, therefore, contradict the efficient markets hypothesis inasmuch as such cointegration would imply forecastability of share prices from past information, if such information is exploited rationally and equilibrium expected returns are constant.

12. Organisation for Economic Co-operation and Development.

13. Financial Times Stock Exchange 100

14. A convenient outline of these issues can be found in Cunningham (1994).

15. The Cauchy distribution is an example of a highly leptokurtic distribution. It is characterized by an expected value of the variance of infinity, and an expected value of the mean of less than infinity.

16. This modification was developed by Lo (1991) to allow for short-term dependency.

17. Even so, examples of the use of trading rules can be found as early as the work of Wyckoff (1910), such techniques being commonly accepted as originating in the work of Charles Dow (an editor of the Wall Street Journal) around the turn of the century. Other references are found in Neill (1931) and Schabacker (1930), whilst Gartley (1930) discusses the use of a moving-average trading rule. Coslow and Schultz (1966) provide a useful survey of the early work.

18. *Economist*, 5th.December, 1992, pp.23-26. Beating the Market: Yes, It Can be Done.

19. *Economist*, 5th.December, 1992, pp.23-26. Beating the Market: Yes, It Can be Done.

20. *Economist*, 5th.December, 1992, pp.23-26. Beating the Market: Yes It Can be Done.

21. Froot and Perold (1990) suggest that the spread of stock-index futures and portfolio trading has made the short-term adjustment of share prices to market-wide information very efficient albeit slowing, they suspect, the reflection of stock-specific information.

22. Hulbert, M., editor of the Alexandria, Virginia-based Hulbert Financial Digest, quoted in *Euromoney*: Forbes, 17th August, 1992, p.135.

23. *Wall Street Journal*, March 1992, quoted in Reuters Textline service as "Wall Street Journal C/C 3/92".

24. A negative effect on Monday returns is also found in Canadian, Japanese and Australian equity markets, the latter two countries displaying a similar effect on Tuesdays (see Alexakis and Xanthakis, 1995). Results for European markets have been more mixed
(Jaffe and Westerfield, 1985; Hawawini, 1984; Solnik and Bousquet, 1990). A study of the

Greek stock market by
Alexakis and Xanthakis (1995) also found a Monday effect and something of a Tuesday effect, but only in data since 1988. Prior data revealed no such biases. This may be related to major structural changes in the Greek market over this period as its characteristics have changed in line with that found in most developed countries.

25. Earlier studies of the influence of business cycles on calendar effects can be found in Liano and Gup (1989), Liano, Manakyan and Marchand (1992), Liano (1992) and Liano, Huang and Gup (1993).

26. Coslow and Schultz (1966) survey this work.

27. The estimates of the round trip dealing spread, i.e. the difference between the bid and offer price of the securities, are taken from unpublished research by Dimson and Marsh, quoted in Dimson and Fraletti (1986), p.45.

28. *Insurance Age*, 10/88, pp.10-11.

29. A more detailed analysis of the Buffet "phenomenon", and its implications for efficient market theory, can be found in Lowenstein (1996).

30. In Patel, J., Zeckhauser, R. and Hendricks, D. (1992)

31. Mark Hulbert: The Hulbert Guide to Financial Newsletters: New York Institute of Finance. Quoted in *Euromoney,* Forbes, 17th August, 1992, p.135.

32. See *Financial Times*, 1st.September, 1989, p.10, and *Chartists with a Hot Line on Currency.*

33. The Bank of England *Quarterly Bulletin* (1989) provides a survey of technical analysis used in the foreign exchange markets.

34. See *The Economist*, 23rd. September 1989, p.135.

35. Allen and Taylor (1990) show that if chartists in the foreign exchange market have bandwagon expectations, excess volatility in exchange rates will result from the destabilising influence of these expectations on the market.

36. "Fidelity Changes Tack," in *The Economist*, 8th.August, 1992, pp.67-68.

37. See Bernard (1984) for a useful guide to the Value Line Investment Survey.

38. *Wall Street Journal*, 3rd.October, 1989. USA: Pros Outperform Investment Dartboard in Stock Picking.

39. *Investors Chronicle*, 25th.November, 1991, p.16. Doing the Random Walk - Selecting a Share Portfolio.

BETTING MARKETS

CHAPTER THREE

WEAK FORM INFORMATION EFFICIENCY IN RACETRACK BETTING MARKETS: CONCEPTS, DEFINITIONS AND TESTS

3.1 Introduction

This chapter examines the idea of weak form information efficiency as it has been applied to racetrack betting markets, and reviews in particular the definitions, distinctions and tests which are associated with this concept.

It has been shown in Chapter Two that weak information efficiency is the notion that current prices incorporate all the information available from a study of past prices and price movements. In consequence, in a financial market which is weakly efficient it should not be possible to earn abnormal returns through a strategy of predicting future prices from past information on prices. Indeed, any such strategy should on average yield the same return.

Many studies of weak form information in betting markets have adapted this idea to examine the possibility for earning differential (or even abnormal) returns in the future, from betting on the basis of past information about the yield to bets at identified prices. In a betting market, these prices take the form of "odds". Odds of 3 to 1 laid against an outcome, for example, imply a return to a successful bet of three times the initial stake, plus the initial stake. An unsuccessful bet loses the entire stake. The theoretical point is that in a betting market which is weak form efficient the expected return to betting at any identified odds or odds grouping should be identical, unless there are differential costs or risks associated with betting at the various prices. Indeed, Snyder (1978a) argues that if horse race betting markets are weakly efficient,

"...then the expected rate of return for all types of bets would be identical" (p.1110)

Section 3.2 of this chapter contains a review of studies which have investigated the expected returns to bets at differing odds, identifies systematic biases in these returns, and considers various explanations for these biases derived from the preferences of bettors. An assessment is made of the implications of these findings for the existence of weak form efficiency in racetrack betting markets. Section 3.3 investigates an explanation of the observed biases framed in terms of the rational behaviour of profit-maximizing odds-setters who face bettors who possess potentially superior information. Section 3.4 reviews the literature on 'technical systems' of betting. These systems employ and utilize the information contained in current odds and the pattern in such odds, with the purpose of identifying and exploiting market inefficiencies so as to make above-average or abnormal returns. Section 3.5 summarizes the chapter and draws some conclusions.

3.2 Measuring the expected returns to bets at differing odds

This section reviews studies which have investigated systematic patterns in the expected return to bets placed at various odds levels and ranges of odds.

Tests for the potential existence of a differential return at different odds can be traced to laboratory experiments by Preston and Baratta (1948), Yaari (1965) and Rosett (1971). They each found evidence of a systematic tendency by subjects (under controlled conditions) to underbet or undervalue events characterized by high probability, and to overbet or overvalue those with low probability. Preston and Baratta calculated an indifference point below which subjective probabilities are objectively too large, and above which they are too small. They found this indifference point to lie close to the geometric mean of their series. These findings, if reproduced among real bettors, would imply that at lower odds the subjective probabilities attached by such bettors to a successful outcome would tend to understate the objective probabilities, while the reverse would be true at higher odds. Such an effect has come to be known in the literature as the 'favourite-longshot bias'[1] or simply the 'longshot bias.'

Tests for the existence of this bias in non-laboratory conditions can be traced to Griffith (1949),[2] who investigated the pool ("parimutuel") betting markets characteristic of U.S. racetracks. In these markets, winning bets share the pool of all bets. The objective probability, in the sense of the percentage of winners, was calculated for each odds grouping and compared with the subjective probability implicit in the established odds. He found that the subjective probabilities were close to the objective probabilities of winning. This point was developed and clarified by Hoerl and Fallin (1974), who ranked horses within a race by their track odds (for races categorized by the number of runners), and compared the average subjective probability implied in the odds with the actual finishing positions. They found a close correspondence between the subjective and objective probabilities, and that the average finishing position fell monotonically in the direction predicted by the odds. They concluded that bettors were able on average to "discriminate small differences" (p.230) in the probability of events occurring. Griffith (1949) also confirmed a tendency for bettors to undervalue events characterized by high probability, and to overvalue those with low probability, an effect which is consistent with higher expected returns at lower odds than at higher odds. Like Preston and Baratta, he calculated an indifference point below which subjective probabilities were objectively too large, and vice-versa, reaching broadly similar conclusions. Moreover, it was almost invariant as between samples taken from years with widely differing economic conditions[3], suggesting that the point of indifference is stable and independent of both the geometric mean and the amount of money available to bettors. Griffith (1961) extended the analysis to cover 'show' betting, i.e. betting on horses to finish third or better[4], for horses offering odds to win of less than 2 to 1, in the months of May 1949 and August 1960. For these data he was able not only to confirm the existence of a longshot bias, but also to demonstrate that a strategy of betting on all horses to show which started at odds to win of less than 1.4 to 1 against would have yielded a profit net of all deductions.

"As was to be expected, the tendency, which had been demonstrated with win betting, for horse race bettors to place too little money on the horses most likely to win is magnified in their even more conservative bets on the same horses to show." (p.81). Since then, others have identified evidence of mispricing in the place and show pools; for example, Hausch, Ziemba and Rubinstein (1981), Tuckwell (1981), Swidler and Shaw (1995).

McGlothlin (1956) used betting patterns and outcomes associated with a series of horse races[5] in order to determine the expected value of constant-size bets over a range of probabilities of success. Odds below 3 to 1 (against) yielded a positive expected value[6] after correcting for track deductions, odds above 8 to 1 a negative expected value[7], and odds of 3 to 1 to 6 to 1 an expected value approximately equal to zero.[8] McGlothlin located the indifference point at a value between 0.15 and 0.22 (i.e. between 3.5 and 5.5 to 1 against), findings which are consistent with those of Griffith (1949).

Performing the same analysis on sub-samples of the data corresponding to races classified by their position in the eight-race order yielded no surprises for the first six races on the cards. The patterns found were not significantly different from those established for the whole sample. Significant differences were, however, identified for the final two races of the day. Uniquely, the relatively high expected values displayed by the shortest odds groupings was not reproduced for the seventh (i.e. penultimate, usually feature) races on the card, although a significantly higher than expected return appeared in the 6 to 7.95 odds classification,[9] and the horses in the odds range from 16 to 25.95 displayed an exceptionally low expected return. For the eighth (i.e. last) races of the day, the expected return to betting on horses in the very shortest odds category chosen by McGlothlin, i.e. 0.5 to 1.95, was particularly high[10], dropping sharply for odds in the 3 to 3.95 category.[11]

Snyder (1978a, 1978b) provided surveys of the published evidence. On the basis of the existing literature[12] and his own data on U.S. horse race betting markets, he concluded that lower odds tend to be associated with higher returns, and vice-versa. Indeed, all of the studies indicated that bets placed at odds below 5 to 1 would have yielded above-average returns. However, the returns were not large enough to yield a profit after allowing for standard deductions from the pool.

Snyder (1978a) argued that "... the evidence collected for the weak form test shows that the public has a clear and strong bias which substantially affects the expected rate of return for various odds-groups, but that bias is not large enough to overcome track takes of nearly 20 per cent." (p.1114).[13]

Of the authors surveyed by Snyder (1978a), Fabricand (1965)[14] and McGlothlin (1956)[15] also found evidence of monotonically decreasing rates of return from the lowest odds to the highest, while Ali (1977)[16] offered evidence of a greater bias at smaller tracks. McGlothlin (1956) reported a systematically lower bias in the feature race of the day.

Snyder (1978a, 1978b) found the same bias in the predicted odds of various track experts. Using data gathered about 7657 horses running in 1975 at Arlington Park racetrack, Chicago, together with the predicted odds of the official track handicapper, of the *Daily Racing Form*, and of three major Chicago newspapers, he found that each of the 'experts' exhibited this same bias. However, he ascribed this more to perceived constraints on the range at which the 'experts' quoted the horses than on any inherent preference for longer odds. Another explanation is that the 'experts' are simply trying to predict the odds rather than what they perceive to be the actual winning chances, and so will reproduce any such bias. Snyder noted, however, that this in itself does not explain his finding that the degree of bias exhibited by the experts is greater in every instance than that demonstrated by the public, as indicated in the final track odds.

Asch, Malkiel and Quandt (1982) examined the relationship between the subjective and objective probabilities of a horse winning a race, as evidenced by the parimutuel odds[17] and actual outcomes respectively.[18] Although they found a close relationship between a horse's place in the order of favouritism and the likelihood of it winning, they also found that bettors tend to overbet horses offered at particular odds and to underbet others. In particular, whereas the objective probability of a horse winning was significantly greater[19] than the subjective probability for the favourites examined, the bias was gradually reversed until the subjective probability was significantly greater than the objective probability for the ninth horses in the order of favouritism. The implication is that betting on shorter-priced horses would tend to produce a higher rate of return than those on offer at higher odds, and this was borne out by an analysis they undertook. Specifically, they calculated the rates of return for bets at odds groupings varying from 0 to 2 to 1 against at one extreme, and of odds ranging from 25 to 1 upwards at the other. They also derived results employing the same odds groupings, but limiting the sample to the last two races of the day. Their findings for the total data set are consistent with the existence of the longshot bias already noted by Griffith (1949), McGlothlin (1956), Ali (1977), Snyder (1978a, 1978b) and others. Moreover, their analysis of the later races confirmed an earlier finding of McGlothlin (1956), Ali (1977), and Kahneman and Tversky (1979), that the bias toward underbetting short odds and overbetting long odds is particularly strong in such races. In contrast, Metzger's (1985) analysis[20] of the betting public's first and second favourites revealed virtually identical patterns of betting in the first and last races of the day, these patterns being different to all other races. Omitting the first races of the day, however, Metzger identified a significant underestimation of the true probabilities of favourites winning the last two races as compared with earlier races.

A survey of the overall picture, by Thaler and Ziemba (1988),[21] assessed the evidence from a wide range of previously published studies to calculate the expected market return at various odds. Net of deductions from the pool, expected returns confirmed the conventional bias, turning positive at a cut-off point of about 4.5 to 1. At odds of below 0.3 to 1, they even report a positive expected return gross of deductions, i.e. an expected profit.[22] This direction of bias is also documented in Hausch and Ziemba (1990), and has been confirmed for Australian data (Bird, McCrae and Beggs, 1987), and for New Zealand data (van Zijl, 1984).

Even so, there is not universal consistency in the published studies. A notable exception to these findings is reported by Busche and Hall (1988), for Hong Kong racetrack betting markets, and by Busche (1994) for Hong Kong and Japanese racetracks. In these markets they found no evidence of a positive bias, if anything the bias operating in the opposite direction. This sort of effect was also reported by Swidler and Shaw (1995) for a small U.S. racetrack.

Busche and Hall's (1988) study used data gathered from 2653 races at Hong Kong racecourses between 1981-1982 and 1986-1987. In line with earlier studies, such as Hausch, Ziemba and Rubinstein (1981), Busche and Hall's methodology involved asking how far the returns to random bets across differing odds categories were equal. Their logic is that if those placing bets are risk-neutral and also make accurate and unbiased predictions, then the returns should be equated across horses characterized by differing win odds (to reflect the winning proportions).[23] In other words, if a regression line is drawn through the scatter of points generated on a graph described by observed win odds on the horizontal axis and the actual available betting odds on the vertical axis, then risk-neutrality is consistent with a regression line demonstrating a slope of one. Similarly, a slope of greater than one is consistent with risk-aversion, and a slope of less than one with risk preference. Their actual results are (standard errors in brackets):

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

The slope estimate was significant beyond the 0.001 level, indicating evidence in this sense of risk aversion.[24]

Allowing for the existence and structure of measurement errors they were unable, however, to reject a hypothesis of equal average returns across groups of horses. They concluded that no evidence existed that Hong Kong bettors underbet favourite and overbet longshots.

Busche (1994) reported analogous results from a later sample of 2690 new Hong

Kong races (1987-1992), by pooling the new and original data into a total sample of 5343 Hong Kong races (i.e. 1981-1992), and separately for 1738 Japanese races from 1990.

Swidler and Shaw (1995), as noted above, also found no evidence of a favourite-longshot bias in their study of a small U.S. racetrack. The track, Trinity Meadows Raceway, was selected for study precisely because it is small ("a second tier Texas track", p.306). In this context, the small pool and the cost of obtaining accurate information might be expected to produce a population of relatively "uninformed" bettors. Their data set covered 2946 horses, running in 288 races, between June and December, 1991. Although the subjective and objective win probabilities were highly correlated, the application of a Spearman rank correlation coefficient to the returns in different odds groups revealed no significant bias (at the five per cent level). At less strict levels of significance, there was a reverse bias, i.e. bettors tending to overbet the favourite, and vice versa.

Common to the studies of Busche and Hall, Busche and Swidler and Shaw is that they examine behaviour in parimutuel markets. An investigation of the U.S. baseball betting market by Woodland and Woodland (1994) is distinguished by the fact that this is a fixed-odds betting market, in the sense that the odds can be agreed with odds-setters (bookmakers) at the time a bet is placed. As such, a bettor is able to ascertain in advance the eventual payoff to a successful bet. It is also equivalent to a set of two-horse races, inasmuch as bets are either on one team or its unique opponent. In this sense, bettors are in effect buying or selling an asset. The data set consisted of 24,603 major league baseball games for the 1979 to 1989 seasons. Woodland and Woodland tested for efficiency by applying z-tests and regressions (after Asch and Quandt, 1987, 1988) to ascertain whether there were systematic differences in the subjective and objective probabilities of the longshot ("underdog") winning. Their methodology is based on the premise that there should not be any significant differences if the market is weak form efficient. Their results suggest some evidence of market inefficiency (at the ten per cent level of significance), with baseball bettors tending to overbet the favourites relative to the longshots. A strategy of betting only on the longshot produced a higher average return than would be consistent with this definition of efficiency, at the five per cent level of significance. Any inefficiencies were not great enough, however, to yield a positive return net of deductions.

Most studies of betting with bookmakers have, however, been conducted using British racetrack betting data. In the markets from which this information is derived, bettors can take posted odds or else, in horse and greyhound racing, sometimes "take the Starting Price." The Starting Price is the independently determined assessment of the general price at which a "sizeable" bet could have been placed about any particular outcome with bookmakers on the course at the start of the race. This option is often, even usually, taken by off-course bettors. Bettors taking posted odds (or "Board Prices") are unaffected, however, by any subsequent odds movements.

These studies are traceable to studies of betting patterns undertaken by Figgis (1951, 1974a, 1974b, 1976), and quoted in part in the report of the Royal Commission on Gambling (1978), and point in the same direction as most of the U.S. parimutuel data. Figgis' evidence for 1950, 1965 and 1973, using starting prices, demonstrates that for the shortest odds examined, i.e. 0.4 to 1 (5 to 2 on) or less, the average pre-tax return varied from 97.2 per cent in 1950 to 108.1 per cent in 1965 to 108.5 per cent in 1973.[25] Calculations of the returns in 1975 and 1976, performed for the Royal Commission on Gambling, found rates of return of 112.1 per cent and 107 per cent respectively. Figgis' calculations for the longest odds range, i.e. 20 to 1 and over, on the other hand, demonstrate much lower average returns, varying from 23.8 per cent in 1950 to 37.3 per cent in 1965 to 23.2 per cent in 1973. These returns to extreme longshots are qualitatively in line with the U.S. findings reported by Snyder, although much more pronounced in extent. Although not a monotonic relationship, Figgis produced evidence of a persistence in this tendency over the intervening odds ranges. Over all odds ranges, the average return was about 80 per cent.

Dowie (1976) calculated the expected return at each of a wide variety of starting prices for the 1973 flat season, and derived the expected rate of return to a pattern of betting a unit stake on each and every horse at the starting price. He also derived the expected rate of return to a policy of betting on every horse so that the return at its starting price would yield a constant return. Whereas he calculated that the first approach would have yielded a pre-tax loss of 39.4 per cent, the second approach would have yielded a loss of 20 per cent, although most of the disparity occurred when examining odds in excess of 20 to 1 against. His sample of 2777 races also revealed evidence of a significant longshot bias, which he examined by sub-dividing the results into actual returns and cumulative returns to a policy of level staking, and also to a policy of staking to yield a constant return. He noted a profit even after tax at odds up to 4 to 6 (often termed 6 to 4 on). An examination of his figures reveals a cumulative profit before tax for all wagers struck at less than evens (odds of 1 to 1), given either of the two staking methods he explores. Again, the return to longshots (especially extreme longshots) was far worse in extent than that reported by Snyder for U.S. parimutuel markets.

Henery (1985) examined later evidence from the U.K. flat racing season, selecting 883 races in 1979 and 1980. The average return to a unit stake was calculated over various odds ranges, showing a similar bias to that offered by Figgis (1976), i.e. ranging from 97.9 per cent in the odds range 0 to 0.396 to 1, to 10 per cent in the odds range 38.12 to 1 and above. This inverse tendency, though not systematic through all odds ranges, was preserved as a general trend over all intervening odds classifications.

The Ladbrokes Pocket Companion, Flat Edition (1990) provides findings for the flat racing seasons from 1985 to 1989, showing evidence again of a systematic bias against the expected return at long odds, a result even more clearly illustrated by grouping the odds. The results suggest that a positive rate of return was available at strategies involving the consistent placing of bets at odds of 1 to 2 (2 to 1 on) or shorter. In particular, betting at odds of 1 to 5 (5 to 1 on) to 1 to 15 (15 to 1 on) would have yielded a 6.5 per cent profit,and at 1 to 8 or shorter a positive return every time.

Cain, Law and Peel (1992) examined the evidence for the existence of a favourite-longshot bias in U.K. greyhound racing betting markets. They compared the probability of winning implied by the starting prices, standardized to deduct the (ex-post) bookmakers' margin, with the realized win probabilities. The average returns were calculated using the returns to a unit stake on every greyhound at the starting price, and also by the average return from placing a stake to win a unit return at each starting price, i.e. the reciprocal of the starting price. While they offer no conclusions supporting a positive linkage between expected returns and shorter odds, they did find that the realized win probabilities exceeded the win probabilities implied by their standardized starting prices at all odds up to 1.5 to 1 (6 to 4 against). This is evidence of a favourite-longshot bias. They were unable, however, to translate any such inefficiencies into a strategy capable of yielding abnormal returns.

Thus, the weight of the evidence, at least for horse racing betting markets in the

U.S. and U.K., is in favour of the existence of some positive relationship between the expected rate of return to betting and the placing of bets on those horses most expected in the market to win. The implication is that higher average returns can be earned by betting on horses offered at particular identified odds (generally lower) than others (generally higher), and particularly so by betting at extremely short odds. As such this not only violates one definition of weak form efficiency, but also requires explanation in terms of rational economic behaviour.

3.2.1 Risk, Return and the Favourite-Longshot Bias

In this section various ideas are advanced to explain the existence of a favourite-longshot bias in racetrack betting data. In the context of these proposed explanations, the significance of the existence of this bias for weak form market efficiency is evaluated.

Much of conventional capital market theory assumes that a higher return is required to compensate investors for the incurrence of higher risk. Higher risk in the context of horse race betting may be associated with betting at higher odds, inasmuch as such odds are usually associated with lower probabilities of winning and a higher variance of return. The implication of such theory, therefore, is that a higher expected return would be required to compensate for greater risk, implied by the greater odds. Evidence from the behaviour of bettors in horse racing markets in the U.S. and U.K. suggests the opposite.

Attempts to provide an explanation founded in economic theory (for this apparent incongruity) can be traced to work by Rosett (1965) and Weitzman (1965). Rosett asked whether such observed behaviour was reconcilable with the existence of a sophisticated, rational betting public. A sophisticated bettor in this sense is one who satisfies three conditions. First, if the probabilities of winning are equal, he will choose that with the greater return to winning; second, if the returns to winning are equal, he will choose the one with the greater probability of winning; and third, in a choice of bets, he will always prefer a bet which exhibits both superior returns and a higher probability of winning than the alternatives. These conditions are referred to by Rosett as the "rationality hypothesis" (p.596). To test this rationality hypothesis, Rosett examined and compared the distinct risk/return profiles associated with different types of bets. He concluded from an empirical examination of the evidence from actual betting behaviour that bettors are sophisticated and rational, but that in their choice of betting strategies they displayed a strong preference for low-probability, high-return bets.

Weitzman (1965), using the same data and assuming that a proposed representative bettor[26] obeys the expected utility hypothesis, constructed a representative utility of wealth function from the relationship between the subjective and objective probabilities implied by the odds and the results, employing a weighted least-squares method that corrects for heteroscedasticity. For the sums of money examined, he found a range of values which implied increasing marginal utility (convexity in the utility function), signifying that bettors exhibit risk-loving behaviour. The implication of this convexity is that expected utility maximization would generate the observed longshot bias. Weitzman suggests that these findings coincide with the range of increasing marginal utility proposed on theoretical grounds by Markowitz (1952) as an amendment to the utility of money curve offered by Friedman and Savage(1948).[27] Asch and Quandt (1986) adhere to a similar conception of betting behaviour, i.e. that the utility function of horse race bettors may well be convex above the current level of wealth and concave below it. An obvious problem common to these approaches, as Thaler and Ziemba (1988) note, is whether they would explain behaviour displayed by bettors in other contexts;

"We venture a guess that when it comes to retirement saving, Professors Asch and Quandt would not be willing to accept a lower mean return in order to obtain a higher level of risk." (p.170).

An approach favoured in a number of studies of risk-taking behaviour is to propose such behaviour to be very context-specific (see Slovic, 1972). In particular, Thaler and Ziemba employ a concept of "mental accounting" (see Kahneman and Tversky, 1984; Thaler, 1985), whereby "... people adopt mental accounts and act as if the money in these accounts is not fungible" (p.171) in order to demonstrate how one may be risk-seeking at the racecourse, but risk-averse with respect, for example, to one's pension provisions.

Other studies, however, frame the issue of variations in attitude to risk in terms of the availability of ready capital to the bettor, or changes in such over the course of the betting period. A seminal study along these lines was undertaken by Ali (1977), who estimated subjective and objective winning probabilities from a database of 20,247 harness horse races at three tracks between 1970 and 1974. He confirmed the tendency for the odds to understate the likelihood of outcomes with a high probability (the subjective probability understates the objective probability), and to over-state those with a low probability. Employing Weitzman's concept of the "representative bettor", he found that such bettors did exhibit behaviour consistent with adopting a riskier approach at the smallest of the three tracks, i.e. at Saratoga. Ali offered this in support of the view that "... the more capital the representative has, the less he tends to be a risk lover" (p.815). The implication is that bettors possess an increasing marginal utility of money function[28], gambling being explained in terms of reallocating consumption possibilities in response to this. Ali also reported increased risk acceptability in the last race of the day compared to the first two races of the day (an effect first suggested by McGlothlin, 1956), a result he interprets as due to the influence of a change (decrease) in capital, as the day progresses, on attitude to risk (increasing risk-loving behaviour).

Two other studies, one by Kahneman and Tversky (1979) and the other by Asch, Malkiel and Quandt (1982), also identified greater apparent risk acceptability in races occurring later on the racecard, in the form of a greater longshot bias. Asch, Malkiel and Quandt explained this, like Ali, in terms of a proposed change in the risk attitude of bettors with respect to variations over the course of the day in their available betting capital. Their explanation is couched in terms of the fact that bettors are seeking to recoup their overall losses on the day. This conclusion should be examined in the context of the findings of a study by Hamid, Prakash and Smyser (1996) of Florida greyhound races. That study supported the view that bettors' aversion to risk declined as their losses increased during a racing session, causing an increase in the favourite-longshot bias in later races.

Although Metzger (1985) was unable to confirm any significant differences in risk acceptability between the last and first races on the card (possibly due to special features in these races), there was evidence that bettors' first choices were underestimated in races 8 and 9 (overall mean of 94) compared to races 2 through 7 (overall mean of 105). No significant pattern was detected, however, for the sample of second choices. Metzger offered these results as evidence in support of Tversky and Kahneman's (1981) proposition that variations in reference points for the framing of outcomes produce variations in the acceptability of risk.

"In particular, given that the reference point is the status quo at the beginning of the racing day and that the public expectation is negative, outcomes are framed increasingly in terms of getting-even versus loss rather than gain versus loss, producing fewer bets on favourites over the day. The public should increasingly underestimate the chances of favourites in later races." (Metzger, p.883).

If true, this tendency by bettors produces a clear implication, and one lent empirical support in parimutuel markets by Kopelman and Minkin (1991), that "The Best Time to Bet the Favourite is in the Last Race" (p.701), known as Gluck's Second Law. A study of behaviour by bettors in the fixed-odds arena of U.K. off-course bookmakers by Johnson and Bruce (1993) offers quite different conclusions. Employing a random sample of 1212 real bets, placed at betting offices throughout the U.K. between March 12th. and April 18th. 1987, they found a tendency for bets on races later in the day to be placed on horses at shorter odds, even allowing for disparities in field size (and therefore mean odds size). Moreover, there was a tendency for the mean stake size to increase in later (the last three) races compared with bets on the first two and first three races. The expected return to bets placed on later races also tended to exceed that to bets on earlier races, although only one of the early/late race comparisons was significant at the five per cent level. These results are consistent with the suggestions of empirical work by Thaler and Johnson (1990) and Garling, Romanus and Selart (1994) - see above - that prior losses tend to produce less risk-seeking/more loss-aversive behaviour. Even so, any interpretation of these findings should, as Johnson and Bruce point out, take account of potential differences in off-course behaviour (studied here) and on-course behaviour. It may well be that two separate influences are at work; a tendency for prior losses to increase risk-aversion, but also an over-importance placed by on-course bettors on the need to at least break even on the day.

Hamid, Prakash and Smyser (1996), using their Florida greyhound data, sought more basically to distinguish in bettors' utility functions between preference for variance and preference for skewness (i.e. preference for non-symmetry in the distribution of payoffs implied by a sample of parimutuel odds). Employing a standard von Neumann-Morgenstern expected utility of wealth function, they concluded, on the basis of their observation of the relevant payoffs, that the representative bettor exhibited behaviour which demonstrated a preference for variance and an aversion to positive skewness. This is consistent with the conclusions of a study by Quandt (1986) which showed how a favourite-longshot bias could arise as a natural and necessary consequence of equilibrium in a market characterized by risk-loving bettors, with homogeneous beliefs, in the context of a mean-variance framework.

Chadha and Quandt (1996) demonstrate an alternative scenario in which a favourite-longshot bias can arise in the context of risk-neutral bettors, each of whom optimizes given the bets of all other bettors. Simplifying (though not necessarily realistic) assumptions are that the aggregate of bettors arrive simultaneously at a Nash equilibrium, and that there are no arbitrage opportunities between parimutuel betting and betting with bookmakers. In this model the bias is a consequence of random, rather than systematic, errors by bettors in their perception of the true underlying probabilities.

The existence of a favourite-longshot bias has been explained in quite different terms by Golec and Tamarkin (1995), who ask simply whether bettors prefer longshots because they are risk-lovers or because they are over-confident. In order to compare the validity of these hypotheses, they identify a data set which is able to distinguish, it is claimed, the influence of over-confidence from that of risk-loving behaviour. The data set is composed of so-called 'teaser' bets. These bets are a variation on a normal point-spread bet, in which the bookmaker sets the margin of victory (the spread) of one team over another. In normal point-spread bets, the bettor chooses whether the spread will be greater or less than this. If the actual margin of victory equals the spread, the bet is void. Otherwise, winning bets earn 10/11 of their stake (plus stake returned), losing bets lose the whole stake. In 'teaser' bets, on the other hand, the bettor can be wrong by a given number of points ('teaser points') and yet still win. For instance, if the spread of team A over team B is 8 points, and team A actually win by 12 points, a bet on B would lose (four points out) without any "teaser" adjustment. With five teaser points, the bet is clearly a winning one. Bookmakers may adjust the agreed payout to the bettor downwards to compensate for the higher probability of winning. In order to win an n-team teaser bet, all of the bets (adjusted by the teaser points) must win in order to earn the agreed payoff. A losing bet loses the entire stake. In order to win a standard multiple (or "exotic") bet, all of the bets (unadjusted) must win. The bettor choosing the teaser bet has a higher objective probability of winning (and lower risk), but will usually receive a lower payoff to a winning bet. If bettors under-estimate the likelihood of making large errors (as suggested by Tversky and Kahneman, 1974; Tversky, Slovic and Kahneman, 1990; De Long, Shleifer, Summers and Waldman, 1991), they will overvalue the teaser points. Since teaser bets reduce risk (return variance), however, risk-loving bettors should require an additional return in order to bet on teasers. Golec and Tamarkin tested this empirically by comparing a given teaser's expected return with the expected return to other bets that have similar or greater objective win probabilities. They found that the teaser bets had larger win probabilities but much smaller returns. These findings are consistent, they argue, with the hypothesis of over-confident bettors, but not with risk-loving bettors.

In contrast, Hurley and McDonough (1995) and Terrell and Farmer (1996) propose explanations of the favourite-longshot bias which require neither a hypothesis of overconfidence nor of risk-loving behaviour. Instead, the bias can arise in a risk-neutral, confidence-neutral environment, as a consequence of positive transactions/information costs.

Hurley and McDonough consider the case of two types of risk-neutral bettor occupying a parimutuel betting market - 'informed handicappers', who know the 'true' probabilities, and 'uninformed handicappers', who do not. Since the uninformed bettors are unable to distinguish good bets from bad they will, in the simplest case of a two-horse race, bet a roughly equal amount on a favourite as on a longshot. If there are no transactions/information costs, the informed bettors should take advantage of this mispricing in the pool to bet on the horse with the highest objective probability of winning (defined as the favourite). In the model, it is assumed that there are a large number of informed bettors, and that the objective probability, net of track deductions, is greater than the probability implicit in the bets of the uninformed. The expected profit from this strategy is positive so long as the advantage of being informed is not outweighed by the costs of betting. The presence of transactions and information costs, however, cause the subjective probability that the "favourite" (the horse with the highest objective probability) wins to diverge systematically below the objective probability. This systematic divergence produces a favourite-longshot bias, and the bias increases as the costs increase. Nevertheless, laboratory evidence using groups of students, some of whom were exposed to betting costs and some of whom were not, was unable to confirm the theory at an experimental level. On this basis Hurley and McDonough conclude that "...the bias on the favourite is not explained by costly information and transactions costs." (p.953)

Terrell and Farmer employ a similar formulation, composed of informed bettors (who in their case purchase the true probabilities of events), and uninformed bettors (who do not). They model the decision as to whether to become informed explicitly, in terms of the costs of becoming informed, and the wagers of other informed bettors. In this model, if all bettors are uninformed, then the expected loss to any random betting pattern is equal to the track take-out. There is no favourite-longshot bias. The addition of informed bettors complicates the issue, however, as these bettors will bet on horses whose true probabilities of winning exceed the probabilities implicit in the wagers of the uninformed bettors, so long as the net expected return to a bet is greater than one. The size of the bet (and therefore the net expected return) will depend on the size of bets in the uninformed pool, the extent of the divergence of initial market odds from the true probabilities, and the number of other informed bettors in the pool. Informed bettors will therefore act so as to lower the odds on events with high expected returns and increase the odds on events with low expected returns. In consequence, low odds events will tend to be associated with higher expected returns than high odds events. This is the favourite-longshot bias. As transactions and information costs fall, however, the number of informed bettors rises, the expected profit on each bet tends to zero and subjective probabilities (implicit in market odds) converge to the objective probabilities. At this point the observed favourite-longshot bias disappears. The bias is, therefore, a consequence of costs involved in the betting process, such as the track take-out. Terrell and Farmer tested their hypothesis using a sample of 4,121 races at a Kansas City greyhound racetrack[29] in the 1989-1990 season, and also data from the 1993-1994 season. Calculating the return to a random betting strategy revealed an expected payout of 78.3 per cent, compared with a track deduction of 18 per cent. They explain the difference (3.7 per cent) as income to informed bettors. They also found evidence of the traditional longshot bias. The empirical evidence is offered in support of their model of betting behaviour.

3.2.2 Explaining the favourite-longshot bias in terms of utility-maximization

All the above explanations are couched within a financial framework. There is, however, no general agreement that bettors' motivations are best addressed from this perspective. Competing explanations of the favourite-longshot bias seek rather to distinguish bettors as utility-maximizers rather than profit-maximizers (although profit may be an element in the utility function). For instance, bettors may derive utility specifically from selecting longshots. Snyder (1978a), for example, notes that "... the main reward of horse betting comes from the thrill of successfully detecting a moderately long-odds winner and thus confirming one's ability to outperform everyone else" (p. 1113), a motivation reported by Livingston (1974) for the case of a member of Gambler's Anonymous. This gambler always bet heavily on a longshot in the last race, ostensibly so that if he did win he could brag about a winner. Unfortunately for him, according to Livingston he was not able in practice to brag very often. Thaler and Ziemba (1988) suggest that bettors may even derive utility just from holding a ticket on a longshot, while the tendency among a section of the betting public to bet for reasons totally unconnected with any serious assessment of the objective probabilities, for example because they like a horse's name, may contribute to a cut in the odds offered against longshots.

Letarte, Ladouceur and Mayrand (1986) examined the behaviour of 45 subjects who had never played roulette, selected from the general public via advertisements, in the context of a simulated roulette playing exercise.[30] They found that the amount of money bet increased as a function of the number of trials, that the type of bets became more risky as the game went on, and that subjects having frequent wins took significantly more risk than individuals having infrequent wins. They explained their findings in terms of the acquisition by gamblers of a sense of personal, albeit illusory, control which increases in line with increased familiarity with the gambling process and with increased frequency of success (see Langer, 1975; Langer and Roth, 1975; Langer, 1983).[31] One possible implication of their findings is a tendency for bettors who win disproportionately, i.e. those who bet at short odds, to follow this trend toward risk by switching gradually to longer odds. So long as this tendency is not compensated fully by a movement in the opposite direction by other bettors, a 'longshot bias' will result.

An approach favoured by Bruce and Johnson (1992) is to examine the motivations which cause people to bet at all. One such motivation which they specify is excitement. In this context, they comment that

"The excitement experienced by bettors with an interest in the race is naturally heightened by the risk to which they have exposed their stakes and the anticipation of possible success." (p.204).

Bruce and Johnson also note that "... the successful bettor who makes known this success may expect to receive... peer-group esteem associated with perceived 'skill'." (p.205). This effect may be more pronounced in the case of successful prediction of a longshot, there no doubt being some asymmetry in the reporting of failures to successes.

It is, on the other hand, possible to argue that more excitement and peer group esteem is furnished by a succession of successful, if more predictable, short-priced winners.

The desire for excitement, heightened pleasure, and social esteem may thus on certain, though not all, interpretations offer explanations of a longshot bias in terms of the maximization of expected utility as opposed to expected profit. A further possibility is that bettors are confused in their assessment of the expected returns. Bruce and Johnson tested both these types of explanation. By dividing bets into categories based on the timing of the bet, they sought to distinguish inputs into a non-monetary utility function, broadly classified as excitement and social interaction and intellectual challenge, from inputs into a predominantly monetary utility function, i.e. maximization of financial gain; and offered the possibility of assessing the influence of confusion or "non-cognitive constraints" on the ability to make effective decisions.[32] Specifically, they argue that bets placed early in the day contain a disproportionate number of bets placed to meet a need for intellectual challenge. Those placed later contain a disproportionate number of bets designed to meet a need for social interaction; those placed later still tend to satisfy a need for excitement; and those placed latest contain a disproportionate number of bets placed for the specific purpose of maximizing financial gain. Of these, those in the third subset are likely to contain the largest ratio of those subject to some form of cognitive overload or 'decision paralysis,' because these bettors are, observe Bruce and Johnson, "... subject to rapid and continual changes in the information set (e.g. prices of horses, horses' pre-race behaviour)", which "... may tend to distort the meaning of information, suspend vigilant search and be characterised by selective inattention" (p.211).[33] On the basis of a random selection of 1200 bets placed throughout the U.K. between March and April, 1987, they concluded that according to a variety of measures of actual financial return, the third subset performed worst and the fourth subset, i.e. the very late bettors, best. The group of bettors best identified as profit-maximizers (i.e. the very late bettors) displayed the highest propensity to bet on favourites, and the lowest propensity to bet on longshots (i.e. the lowest longshot bias).

The implication of this sort of approach is that one cannot fully explain betting behaviour within the framework of a totally rational and unconfused cognitive process, which strictly adheres to the goal of maximising financial return. Any understanding of the longshot bias must allow for this.

An alternative possibility is to explain the phenomenon in terms of differences in the staking patterns. Findings offered by Filby and Harvey (1988)[34], for example, on the link between amounts staked and other variables provide some interesting support for the idea of a longshot bias linked to staking levels. In particular, they identified a clear relationship between the size of stake and the type of bet, larger bets being associated with lower risk bets such as singles. An examination of the relationship between the probability of a positive return and the size of stakes revealed that their largest bets, i.e. over £20, were more than three times as likely to yield a positive return as bets under 50 pence (36.8 per cent to 12.1 per cent). Higher stakes were also associated with the pre-payment of betting tax[35], notable inasmuch as it can be shown that for any given total stake the expected return to a bet is greater if tax is pre-paid rather than paid on the winnings.

A compromise position between explanations which are proposed in the context of assumed profit-maximizing behaviour and those drawing upon a broader utility-maximizing approach is offered by Busche (1994). This explanation, which seeks to reconcile the absence of a longshot bias in his Hong Kong data (see above) with its prevalence elsewhere, is formulated in terms of two distinct types of bettors, those who bet to maximize their money, and those who bet as a consumption activity. Since expected returns are limited by the size of the betting pool, Busche proposes that money maximizers may in consequence dominate tracks offering large pools, whereas betting as a consumption activity may dominate the market at smaller tracks. In this context, it should be noted that average stakes in a day at Hong Kong racetracks were estimated at 120 million U.S. dollars in 1993 (Heckerman ,1994), several times greater than that at the leading U.S. racetracks. Significantly perhaps, Ali (1977, p.813) also observed a much greater bias at the smallest of the racetracks he examined.

3.2.3 The implications of a systematic under-estimation by bettors of their losses for our understanding of a favourite-longshot bias

One of the simplest, and perhaps most under-rated explanations of the favourite-longshot bias, and one which has evaded citation in most of the subsequent literature, can be found in Henery (1985). Henery argues that a favourite-longshot bias can arise as a consequence of bettors discounting a fixed fraction of their losses, i.e. they underweight their losses compared to their gains. This is an explanation which "rings true," whether it is true or not (see Gilovich, 1983, for a psychological perspective on this issue). It is also capable of explaining in a clear manner an observed link between the sum of the implied probabilities in bookmakers' odds (termed the 'over-round') and the number of runners in a race.

The rationale for Henery's hypothesis is that punters will tend to explain away and therefore discount losses as atypical, or unrelated to the judgement of the bettor. If, for instance, the true probability of a horse losing a race is q, so that the true odds against winning are $q/(1-q)$, then the bettor, he argues, will assess the chance of losing not as q, but as Q which is equal to fq.[36] If f, for example, is 3/4, and the true chance of a horse losing is 1/2 (i.e. $q=1/2$), then the bettor will rate subjectively the chance of the horse losing as $Q = fq$, i.e. $Q = (1/2).\ 3/4 = 3/8$.

Another way of looking at this approach is in terms of a £100 wager placed on a horse to win a race. If the true chance of this occurring is 1 in 2, the bettor can expect to lose £50, i.e. £100 one time in two (we are here ignoring any compensatory returns to a winning outcome). The objective odds against him are therefore evens, i.e. a 0.5 chance of losing divided by a 1 - 0.5 chance of winning.

The subjective assessment of the expected loss, following Henery's logic, with f = 3/4, is 3/4 x £50 = £37.50, i.e 0.375 of the stake. The subjective odds against the bettor are therefore 0.375 (the subjective probability of losing) divided by 1 - 0.375 (the subjective probability of winning), i.e. 3 to 5 against (or 5 to 3 on). This means that he will be just indifferent (if he is risk-neutral, and is motivated solely by profit-maximization) at odds of 0.375 to 0.625, i.e. 0.6. In general terms, the bettor facing true odds of $q/1-q$ will evaluate the true odds as $fq/1-fq$.

Listing, for purposes of exposition, some objective odds, together with their subjective counterparts, on the basis of $f = 3/4$ reveals the following:

Objective odds	**Subjective odds**
evens	3 to 5 against
4 to 1 against	6 to 4 against
infinity to 1 against	3 to 1 against
0 to 1 against	0 to 1 against.

The implication of the above is that, for instance, for a given f of 3/4, 3 to 5 is perceived as fair odds for a horse with a 1 in 2 chance of winning. In fact, however, £100 wagered at 3 to 5 yields £160 (3/5 x £100, plus the stake back) half of the time, i.e. an expected return of £80 (or 0.8 times the stake). £100 wagered at 6 to 4 yields £250 (6/4 x £100, plus the stake back) one fifth of the time, i.e. an expected return of £50 (or 0.5 times the stake). In fact, the higher the odds the lower is the expected rate of return on the stake, although the relationship between the subjective and objective probabilities remains at a fixed fraction throughout.

Assuming that f is some fraction between 0 and 1 it follows that the maximum odds offered will be f/(1-f), corresponding to q (the 'true' probability of losing the race) being equal to 1.

Henery does not demonstrate, under these assumptions, the general relationship between the expected return to a unit bet and the odds. It is proposed here to do so.

For simplicity, take the case of a two-horse race, in which the favourite has a true probability of winning of s and the non-favourite (the longshot) of 1-s. The subjective probability of losing from a bet on the favourite is now f.(1-s) and on the longshot f.s.

A subjective expected return of zero on each horse now implies odds of:

$$\frac{f(1-s)}{1-f(1-s)} \; ; \; \frac{fs}{(1-fs)} \qquad (3.1)$$

The objective expected returns from a unit stake on the favourite are as follows:

$$E(Ret) = \frac{f(1-s)}{1-f(1-s)} . s - (1-s) \quad (3.2)$$

$$E(Ret) = \frac{sf(1-s) - (1-s)[1-f(1-s)]}{1-f(1-s)} \quad (3.3)$$

$$E(Ret) = \frac{-[1-f(1-s) - s]}{1-f(1-s)} \quad (3.4)$$

$$E(Ret) = \frac{-(1-s) + f(1-s)}{1-f(1-s)} \quad (3.5)$$

$$E(Ret) = \frac{(f-1)(1-s)}{1 - f(1-s)} \quad (3.6)$$

This function is negative for all values of f and s less than unity. As the function is continuous in the relevant range, the expected return is lower when the winning probability is lower.

$$\frac{dE(Ret)}{ds} = \frac{(1-f+fs)(1-f) - (s-sf+f-1)f}{[1 - f(1-s)]^2} \quad (3.7)$$

$$\frac{dE(Ret)}{ds} = \frac{1-f}{[1-f(1-s)]^2} \quad (3.8)$$

$$\frac{dE(Ret)}{ds} > 0 \text{ if } f < 1. \quad (3.9)$$

Henery (1985) compared two estimates for the probability that horses with given starting prices (SPs) will lose the race. The empirical lose probability q is taken as the relative frequency of losing, i.e. the fraction of horses which lose for the given odds. The subjective lose probability is implied in the bookmakers' odds. Specifically, $Q = SP/(1+SP)$. Plotting the empirical against the subjective lose probabilities for Figgis' (1976) data for the 1973 flat racing season and his own data for the 1979-80 flat racing seasons, the model $Q = fq$ provided a reasonable fit. Henery also found the fit to a straight line to be greater for his own data, a result he suggested may be explained by the data being chosen from later in the season, and from races with greater prize money, both factors which might contribute to a greater reliability in odds setting.

The same assumption about bettors' behaviour can be used to demonstrate a relationship between the over-round (sum of bookmakers' prices - 1) and the number of runners in a race, in particular that the book 'over-round' should vary linearly with the number of runners, and that the slope of the line should be (1-f). The over-round can be calculated by translating the odds offered about every horse into percentages, and summing these percentages. A book is over-round to the extent that the percentages sum to more than 100. Similarly, a book is over-broke or under-round if the percentages sum to less than a hundred. The over-round is expressed either as the sum, e.g. 109.2 per cent or the difference from 100, in this case 9.2 per cent. Another way of looking at the over-round is in terms of "the sum of money which would have to be wagered in order to ensure a return of £1 no matter which horse wins" (Henery, 1985, p.345). The over-round is the excess of this sum over the guaranteed return of £1. For example, if the minimum stake, at the prevailing odds, required to guarantee a return of £1 is £1.20, then the over-round is 20 per cent.[37]

Applying Henery's hypothesis that bettors only count a fixed fraction f of their losses yields the following predictions for the size of the over-round. Again take the case of a two-horse race. In such an event, to be sure of obtaining a £1 return, the bettor has to stake £1/(x+1) on both horses, where x represents the odds offered about each horse. If x=1 for both horses the bettor will break even whichever horse wins simply by staking £1 on each. The over-round is zero. If x<1, however, the sum staked must exceed £1 to guarantee a return of £1 whatever the outcome, and the book is, therefore, over-round. The converse applies if x>1.

Assuming now that the objective probability of losing is 1/2, then the true odds are 1/2/(1-1/2), i.e. 1 (or evens). Yet the subjective probability (on Henery's hypothesis) is 1/2f/(1-1/2f). If, for instance, f= 3/4, then the subjective probability is:

(1/2 x 3/4) divided by (1 - [1/2 x 3/4]) = 3/8 divided by 5/8 = 3/5.

The over-round implied by odds of 3 to 5 offered against two horses in a two-horse race is:

5/(3+5) + 5/(3+5) = 10/8 = 1.25,

i.e. an over-round of 0.25 or 25 per cent (often referred to as 125 per cent).

In effect, the bettor has to stake £10 to ensure a return of £8.

Generally, in a field of n horses, when the objective probability of losing is n-1/n the subjective probability is:

$$f. (n-1)/n \qquad (3.10)$$

The subjective odds implied by this are given by:

$$\frac{f(n-1)/n}{1-[f(n-1)/n]} \qquad (3.11)$$

The over-round may be derived as follows:

$$\text{Over-round} = \frac{1-f(n-1)/n}{f(n-1)/n + [1-f(n-1)/n]}.n - 1 \qquad (3.12)$$

$$= \quad n(1-f.[n-1]/n) - 1 \qquad (3.13)$$

$$= \quad n - f(n-1) - 1 \qquad (3.14)$$

$$= \quad (1-f).(n-1) \qquad (3.15)$$

The implication of this model is that the average over-round in a race should be related linearly and positively to the number of runners, and that the slope of the line representing this relationship should be (1-f).

An even more general result is derived in Vaughan Williams and Paton (1997a), and demonstrated as follows; where OR = over-round, q_i = objective probability of horse i losing, and p_i = objective probability of horse i winning:

$$OR = \sum_{i=1}^{n} (1-fq_i) - 1 \qquad (3.16)$$

$$OR = n - f.\sum^{n} q_i - 1 \qquad (3.17)$$

$$i=1$$

$$OR = n - 1 - f. \sum_{i=1}^{n} (1-p_i) \quad (3.18)$$

$$OR = (n-1) - f.(n-1) \quad (3.19)$$

$$OR = (n-1). (1-f) \quad (3.20)$$

Thus, the over-round is linearly related to the number of runners.

Henery did in fact use data from the 1979-80 flat racing seasons to plot the average over-round against the number of runners in these races. He found a good fit for the aggregate of races, although consistent aberrations were produced by some, notably prestige, races.

There are, therefore, a number of explanations of the favourite-longshot bias, linked by the common thread that they explain the data solely in terms of the demand side of the market. The bias is explained, therefore, as the natural outcome of bettors' pre-existing perceptions and preferences. This is quite consistent with a market efficiently processing the information available to it. Moreover, there is little evidence that the market offers opportunities for market players to earn abnormal returns or positive profits. Thus although possibilities clearly exist for earning above-average returns on the basis of weak-form information, there is no convincing evidence that this contradicts a wider conceptualization of this type of information efficiency.

A challenge to this demand-side orthodoxy was first proposed by Shin (1991, 1992), which is able to explain the existence of a favourite-longshot bias in a market characterized by odds setters, without recourse to any demand-side influences (which is not to say that these influences do not exist as well). The idea is that odds-setters (bookmakers) face an adverse selection problem when they are faced by bettors who know more than they do ("insiders"), the extent and identity of whom are unknown. The bookmakers' optimal pricing strategy in this environment is to contract odds, particularly where potential losses are greatest, i.e. at higher odds. The consequent differential contraction at different odds levels leads to a favourite-longshot bias. If true, this supply-side explanation indicates some form of information inefficiency, as conventionally defined.

3.3. The favourite-longshot bias as a supply-side phenomenon

Common to all the above explanations of the favourite-longshot bias is that they are couched in terms of bettor behaviour, and are as such demand-induced. Recent contributions to the literature have sought to explain the favourite-longshot bias instead as an optimal supply-side response to market uncertainties.

Shin (1991) examined the fixed-odds bookmaking system as a case of adverse selection in which the bookmaker faces a number of bettors who possess superior information, the proportion and identity of whom are unknown to the bookmaker. In the simplest modelling of this situation, insiders are assumed to know with certainty the result of a race while the rest of the betting population, i.e. outsiders, are simply noise traders. Shin employed a comparative static method using a model with linear pricing to solve the bookmakers' profit maximization problem in this context. Shin (1992) extended the model for a horse race with n horses and three players, i.e an incumbent bookmaker, a potential bookmaker and the insider. The outsiders exist as before and are not modelled formally as players. Assuming that the incidence of insider trading is not larger when a favourite is tipped to win than when a longshot is tipped to win (or more precisely, so long as the ratio of insider trading to the winning probability falls as the probability of winning rises), it is shown that equilibrium prices will exhibit a favourite-longshot bias. The intuition is clear. If insiders (who make up a proportion z of all bettors) know with certainty the outcome of a race, than bookmakers face a greater loss from insiders at higher odds (lower implied winning probabilities) than at lower odds. So long as it is not the case that the proportion of insiders (the value of z) is relatively higher at lower odds, bookmakers will face greater expected losses to insiders at higher odds. A simplifying assumption is that where insiders play no part market prices correspond to the true probabilities, and so any deviation from this can be attributed to insider trading. The consequence of price-setting behaviour by bookmakers facing this uncertain environment is for the normalised betting odds to understate the winning chances of favourites and to overstate the winning chances of longshots. This is the traditional favourite-longshot bias. Shin (1993) provided an estimate of the extent of insider trading based on a proposed link between the size of the bid-ask spread in the market and the prevalence of insider trading in the market. The key to the analysis lies in the idea that the direct effect of insider trading on the sum of bookmakers' prices will tend to increase as the number of runners (and therefore the size of the odds) increases, and that this regularity can be used to isolate the proportion of insiders in the total betting population.

3.4 Technical systems of betting

'Technical systems' of betting employ information contained in the odds and odds movements in an attempt to earn above-average or abnormal returns. If successful, they constitute evidence of weak-form inefficiency in these markets.

Attempts to exploit any favourite-longshot bias so as to secure systematic profits is an example of what is termed a 'technical system' of betting. Most of the evidence from studies of win betting suggests that while the longshot bias does usually exist (at least outside of Hong Kong and Japan, it is insufficiently strong to permit systematic abnormal profits after allowing for track or bookmaker deductions (see, for example, Hausch and Ziemba, 1990).

Research traceable to Griffith (1961) has, however, tended to show evidence of inefficiency in the U.S. place and show bet markets, with some evidence that this can be translated into systematic abnormal profits. Subsequent work along these lines was surveyed by Hausch, Ziemba and Rubinstein (1981), in an article which, supported by an analysis of their own data, concluded that there existed evidence of inefficiency in a weak form sense in the place and show markets, and thus the possibility of using technical analysis to make substantial positive profits from these pools. The analysis assumed, however, that the bettor is able to bet after the final odds are set, i.e. after all other bets have been made. In practice the bettor has to balance the advantages of betting as late as possible so as to minimize inaccuracies with being able to perform all the necessary calculations on a given data set. In fact, though, an examination they undertook of odds changes in the last two minutes before the off found that although expected returns did change, profitable bets based on the odds displayed two minutes before the close of betting tended to stay profitable based on the final odds.

Nevertheless, their model in its precise form is a complex non-linear optimization problem that may be difficult if not impossible to solve quickly. As such, they propose approximate solutions using regression procedures which, by limiting the data input, make the system operational in real time. Their approximations, developed for a range of constrained initial wealth levels and a given track take, were able to yield profits of about eleven per cent in the place and show markets.

They also developed a model to show that the abnormal profits were due to the proper identification of market inefficiencies rather than pure chance.

Hausch, Ziemba and Rubinstein's (1981) analysis was developed by Hausch and Ziemba (1985) and Ziemba and Hausch (1984, 1987) into a system known as the Dr.Z system for exploiting inefficiencies in place and show betting. Ziemba and Hausch (1994) present a modified version of the Dr.Z system for place betting at British racetracks. In the British context, place betting is normally on a horse to finish in the first three. They reported the possibility of earning a positive profit from the application of their system, although in light of the higher track take in Britain compared to the U.S., most notably in the Tote place pool (26.5 per cent in Britain compared to less than 20 per cent in the U.S.), they were unable to state how often profitable bets would exist or to assess the likely long run scenario. Incidentally, this deduction from the U.K. Tote place pool is somewhat higher than the deduction from the U.K. Tote win pool, the regressive impacts of which are considered by Dowie (1992a, 1992b).

Swidler and Shaw (1995) also employed an analysis following Hausch et al. (1981) to identify mispricings in the place and show betting pools. They found evidence of opportunities for a positive expected return on 61 occasions (out of 288 races) in which there were disparities between the place and win pool payouts, allowing a positive net expected return. The value of this strategy was, however, limited by the high operational costs (in terms of time), the risk of last-minute odds adjustments, and the deflating effect of a large bet on the pool payout. They examined also, therefore, the value of a simple betting system based on an identified underpricing of win favourites in the place and show markets.

"...bet the win favorite (second favorite) if it is not the favorite (first or second favorite) in the place pool. The strategy in the show market is to bet the win favorite (second favorite) if it not among the first two horses in the show pool." (p.310).

Tuckwell (1981) used data from the win and place betting markets at Melbourne and Sydney racetracks in order to examine the relationship between win odds and place odds. He observed that the relationship was inconsistent, i.e. horses with given win odds did not consistently possess the same place odds. In order to assess whether this reflected genuine differences in the actual probabilities of the various possible outcomes, he used the bookmakers' starting price odds to estimate the true win probabilities, and on the basis of these win probabilities to estimate the probabilities of being placed. His results indicated that a strategy of betting on horses to place in the totalizator (parimutuel) pool when the actual place odds exceed the implied place odds was capable in theory of generating positive profits. However, two practical difficulties were noted. First, the effect of the bet may be to depress the relevant odds, and so reduce or eliminate the profitability. Second, the calculations assume that the bets can be and are taken at the starting price, i.e. the bettor can lay the final bet. The finite time taken to perform the relevant calculations render this unlikely.

The results of these studies taken in aggregate are consistent, therefore, with the existence of some form of weak efficiency in the market for win bets, and of mispricing in the place and show betting pools. There has been less success in utilizing this information so as to make significant abnormal profits from a 'technical' approach to betting. Although some success has been reported in generating such rules in the market for place and show bets some of this has disappeared under close re-examination. Indeed, most systems based on the identification of win/place/show mispricings depend on an ability to perform difficult calculations (see Thaler, 1992), and to operationalize complex decision-making procedures in limited real time. The size of any positive returns are also limited by the size of the pool, and may only offer a small return to time invested.

3.5 Weak form information efficiency in betting markets: summary and conclusions

Many studies of weak form information efficiency in betting markets have adapted the idea of price predictability to examine the possibility for earning differential (or even abnormal) returns in the future from betting on the basis of past information about the yield to bets at identified prices (odds) or odds groupings. In a betting market which is weak form efficient, the expected return to betting at any identified odds or odds grouping should be identical, unless there are differential costs or risks associated with betting at the various odds. The existence of a differential return at different odds has been identified in a number of laboratory studies, the evidence pointing to a systematic tendency by bettors to underbet events with a high objective probability of occurring relative to those with a low probability of occurring. Many studies of actual betting behaviour at racetracks have confirmed the existence of this bias, the expected return (measured after the event) to bets placed at lower odds tending to exceed that to bets placed at higher odds. The implication of this 'favourite-longshot bias' is that bettors can make above-average (though not usually profitable) returns by betting at lower odds. This dependence of future returns on existing prices (odds), derivable from a study of past patterns of returns and prices, is adduced as evidence of weak-form information inefficiency in betting markets displaying it. The existence of a bias in the other direction (a 'longshot-favourite bias') has similar implications, but requires a converse betting strategy. In fact, the usual bias (against higher odds events) has been found in most (but not all) studies of U.S. parimutuel betting markets (in which winning bettors share some fixed proportion of the pool of all bets), and also in markets characterized by bookmakers, such as can be found in the U.K. (for horses and greyhounds) and in Australia. Some studies, however, find the absence of a bias, notably in Hong Kong parimutuel betting markets, and even some evidence of a reverse bias in the U.S. baseball betting market. A large new data set taken from the U.K. racetrack (discussed in chapter 5) is employed to test for the existence of a bias in recent British racetrack betting markets (chapter 6).

Another finding common to a number of studies is that the behaviour of bettors in later races differs significantly from that in earlier races. Explanations have been offered in terms of variations in the amount of capital available to bettors, or a shift in the degree of risk-aversion caused by prior losses. Whether these indications of weak inefficiency constitute a genuine information inefficiency depends upon the reasons for the observed behaviour, and in particular for the favourite-longshot bias. Some studies explain it as a rational outcome in a market characterized by risk-loving bettors, or else that the market is responding efficiently to bettors who just like longshots. Variations of the latter argument are contained in theories of utility-maximizing (rather than profit-maximizing) behaviour by bettors. Such explanations can, however, be so broadly drawn that they are as difficult (or impossible) to refute as to prove. A number of other theoretical explanations fall into the same trap, either being untestable (and unrejectable) or else proposing restrictive (even heroic) assumptions. A good explanation of the favourite-longshot bias should, perhaps, not only explain the bias where it exists, but also the studies where it does not exist, and even other regularities in the data.

An explanation offered by Henery (1985) can in principle explain two of these - the existence of a favourite-longshot bias in U.K. fixed-odds racetrack betting markets, and also an observed correlation in these markets between the sum of bookmakers' prices (the over-round) and the number of runners in a race. Shin (1991, 1992, 1993) offers a different explanation, which also justifies these two regularities. Henery's explanation is based on the premise that bettors discount a fixed proportion of their losses, whereas Shin models the bookmaker's optimal pricing decisions in a market characterized by an unknown proportion of bettors who know the outcome in advance. Since both these explanations explain the data, more discerning tests are required, which are able to distinguish between these hypotheses. These tests are proposed and employed in chapter 8. In the same chapter a new model of optimal behaviour by bookmakers faced by an adverse selection problem is proposed and tested. This model makes less restrictive assumptions about the knowledge of insiders. None of these models is sufficient to explain the variations in behaviour observed across betting markets. The bookmaker-based models do not even attempt to model parimutuel behaviour. Recent studies by Hurley and McDonough (1995) and Terrell and Farmer (1996) explain the longshot bias in parimutuel markets as a consequence of transactions/information costs in the context of two types of bettor (informed and uninformed). A model is proposed in chapter 9, based on these contributions, which seeks to explain generally the existence (and otherwise) of a bias (positive or negative) in racetrack betting markets (parimutuel and fixed-odds).

A number of other studies have attempted to produce technical trading rules, based on other patterns in the odds (or odds movements), or apparent opportunities for arbitrage between the various pools available in a parimutuel market. While some evidence of mispricing (and hence inefficiency) has been found, there is less consensus on how far this can be used to earn abnormal returns in the real market.

In summary, while there is an indication of weak form efficiency in the form of systematic biases in the expected returns at different odds, it is important to understand the reason for the bias if we are to draw the correct conclusions about the existence of genuine information inefficiency. Other indications of weak inefficiency, in particular apparent mispricings across parimutuel pools, seem convincing. The evidence that these can be exploited in real time so as to earn abnormal returns is less convincing.

ENDNOTES

1. See, for example, Thaler and Ziemba (1988), p.163.

2. In a study of 1386 U.S. horse races during August 1947, and subsequently of all U.S. horse races during August 1934.

3. 0.18 for the 1934 data as compared with 0.16 for the 1947 data.

4. Betting on horses to finish second or better is known in the U.S. as 'place' betting.

5. The sample used is one of 9248 races over a period of 1156 days taken from *The Daily Racing Form Chart Book*, vols. 53-59, 1947-1953, Triangle Publications Inc., Los Angeles.

6. The expected value of 0.08 exceeded zero by four standard errors.

7. Significant at the 5% level of confidence.

8. Significant at the 5% level of confidence.

9. Significant at the 5% level of confidence.

10. Significantly above zero at the 0.1% level of confidence, and above the first seven races beyond the 2% level of significance.

11. The difference between this expected value and that for the first seven races was not quite significant at the 5% level.

12. See Fabricand (1965), Griffith (1949), McGlothlin (1956), Seligman (1975), and Weitzman (1965).

13. For the record, Ali (1977) for New York data, 1970-1974, reported an average track take of 15%; Asch, Malkiel and Quandt (1982), for Illinois data, a take of 17%; Hausch, Ziemba and Rubinstein (1981), for California data, a take of 16.8%. Busche and Hall (1988), for Hong Kong data, report a track take of 17%.

14. In an analysis of ten thousand races between 1955 and 1962.

15. In a study of 9248 races, mostly from California tracks, between 1947 and 1953.

16. In a survey of 20247 races at three U.S. tracks, i.e. Saratoga, Roosevelt and Yonkers, between 1970 and 1974.

17. The subjective probability may be defined as proportional to the reciprocal of the market return, c.f. Griffith (1949), McGlothlin (1956), Ali (1977).

18. Their data set is provided by the 729 races making up the 1978 thoroughbred racing season at the Atlantic City, New Jersey racecourse.

19. Significant at the 5% level.

20. For 11,313 races run on nine-race cards at thoroughbred tracks in the U.S. in the months of May, June and December of 1978.

21. Using data in Ziemba and Hausch (1986).

22. This finding was first presented in Ziemba and Hausch (1986).

23. The odds must be adjusted so that the fractions of money wagered and the suggested probabilities add up to one.

24. Hausch, Ziemba and Rubinstein (1981), in contrast, for U.S. data, found the following:

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

25. The calculation of the return is undertaken on the assumption that horses are backed to return a given pre-tax stake.

26. The idea of proposing a single racetrack bettor, representative of all the bettors at the racetrack (Mr.Avmart- i.e. the average man at the racetrack) with a single utility function, can be traced to this article.

27. Friedman and Savage (1948) posited a utility function which is concave at low levels of wealth, convex at intermediate levels, and concave again at higher levels, with the first point of inflection coinciding with the current level of wealth. Markowitz (1952) amended this utility function to include a convex portion at low levels of wealth.

28. See also Weitzman (1965, p.26).

29. Woodlands Greyhound Park.

30. The subjects were allowed to keep 5% of their final winnings in the interests of realism.

31. Bruce and Johnson (1992) link the "illusion of control" argument for cases to the idea of an "intellectual challenge" motivation for gambling.

32. Quoting Eiser and van der Pligt (1988, p.100).

33. See Janis and Mann (1977) for more on this. Also Eiser and Pligt's (1988, p.101) contention that "Firstly, individuals tend to use simpler and less optimal choice rules as the information load increases. Usually accuracy declines considerably when the number of features or the number of alternatives increases. Secondly, the reliability with which choice rules are used tends to decrease as the decision-maker's information load increases."

34. In a survey of over nine thousand betting slips collected from three Birmingham betting offices over one week in June 1984.

35. Such pre-payment of tax (termed 'tax on' bets) accounted for 57.8% of all bets, 38.8% of bets under 50 pence, and 71.9% of bets over £20.

36. Since in reality the true win probabilities are unknown, the problem is in fact one of estimating the win probabilities from the starting price odds. In effect, then, q should more precisely be identified with the empirical average lose probability for given SP odds than with the empirical lose probability itself. See Henery, 1985, p. 347, for a more complete discussion of the issues.

37. At odds of X, the over-round, $R = \Sigma\ 1/(1+X) - 1$

CHAPTER FOUR

SEMI-STRONG AND STRONG FORM INFORMATION EFFICIENCY IN RACETRACK BETTING MARKETS: CONCEPTS, DEFINITIONS AND TESTS

4.1 Introduction

So far the analysis has concentrated on weak form information efficiency in racetrack betting markets. In this chapter the focus turns to semi-strong and strong efficiency. Semi-strong form information efficiency is the notion that current prices incorporate all publicly available information. In consequence, in a financial market which is semi-strong efficient it should not be possible to earn above-average or abnormal returns on the basis of information which is currently publicly available. In a market which is strong form efficient it should not be possible to do so on the basis of all information, including private information. Indeed, any such strategy should on average yield the same return, unless there are differential costs or risks associated with these strategies.

The existence of semi-strong efficiency in betting markets would imply, therefore, that the expected returns to any bet, or type of bet, placed about identical outcomes on the basis of publicly available information, should be identical (subject to identical costs and risks). The same applies with respect to strong efficiency when assessed in respect of all information. Otherwise bettors could use this information to increase their expected returns. In a semi-strong efficient market, for example, the expected return to a bet placed on a horse on the parimutuel (or 'tote') should be identical to that available with bookmakers, should both options be available. Similarly, it should not be possible to identify patterns in the returns which can be used to yield above-average or abnormal returns. For example, the expected return to bets on favourites after a preceding losing favourite should be identical to the expected return after a preceding winning favourite. In a strongly efficient market, it should not be possible for those with access to all information, including private, monopolistic information, to secure a higher expected return (at least net of costs and risk) than those with access to all publicly available information. Prices set later in the market, after those trading on the basis of private information might have been active, should not in this type of market incorporate any more information than those set earlier in the market. It is ambiguous whether information contained in forecasting (tipping) services is publicly available or is private information. Whichever, the distinction must be borne in mind in assessing the significance of this information for the existence of semi-strong/strong form betting market efficiency.

In Section 4.2, the evidence is assessed with respect to the returns about different types of bet. In Section 4.3, evidence is considered as to whether there exists an identifiable market anomaly in the form of a 'gambler's fallacy', i.e. an over-reaction to recent information. Section 4.4 reviews the evidence as to the usefulness of betting systems based on a range of published information (fundamental betting strategies). Section 4.5 assesses the value of racetrack forecasts and forecasting services, and what this tells us about the existence in racetrack betting markets of semi-strong and strong form information efficiency. In Section 4.6, a review is undertaken of strong form tests of efficiency. Such tests are based on an evaluation of the extent of insider activity in racetrack betting markets. Section 4.7 presents a summary and conclusions.

4.2 Employing expected returns to different types of bets placed about identical outcomes to test for semi-strong efficiency

In this section the evidence on the expected returns to different types of bet are considered and assesses. Conclusions are drawn from these findings for the existence of weak form information efficiency in betting markets.

4.2.1 Efficiency and exotic betting markets

Studies of U.K., U.S., Canadian and Hong Kong racetracks have attempted to assess the expected returns to different types of bet, each of which are placed about identical outcomes. In an informationally efficient market, these returns should converge, at least net of differential costs of implementation and risk. Evidence from the U.S., Canadian and Hong Kong racetracks suggest that they do.

Much of this evidence is based on an analysis of so-called "exotic bets", i.e. bets involving two or more horses. An 'exotic bet' is a bet involving two or more horses. Since these bets can be constructed in different ways they offer a convenient test of the hypothesis that the actual returns to bets with identical probabilities of success will themselves be identical. The idea behind these tests is that differential actual returns would indicate evidence of market inefficiency.

Ali (1979) tested the hypothesis of differential returns to two forms of 'exotic bet', known as the 'daily double' and the 'parlay'. In a win bet 'daily double' (also known as a double) the bettor selects the winner of two consecutive races before the first race is run, securing a return only if both horses win. In the 'parlay', the bettor selects a series of horses, betting the total proceeds of each win on the next, until a pre-determined number of wins, or until one loses. The usefulness of comparing these two types of bet is that one can be constructed from the other. In particular, a parlay can be constructed to duplicate the double by selecting the same two horses as in the double before the first race. The added stipulation is that any return from the first race is bet in full on the pre-specified horse in the second race to win. In this form, the win probabilities of this daily double and this parlay are identical. However, the market returns will not necessarily be identical. In a parimutuel system, the return to a daily double depends on the amounts bet on all possible daily doubles involving the two races in question. Similarly, the return to a parlay is determined by the returns to win bets in these races, the win bet return in each race being dependent on the relative amounts bet on all possible outcomes in each particular race. Ali's test of market efficiency is based on the idea that in an efficient market bets will be valued according to their probability distributions alone, and so the return to a daily double will be the same as that of the corresponding parlay.

Using data from 34 racetracks in the U.S. and Canada between September and December 1975, Ali (1979) compared the return to daily double bets with the corresponding return to equivalent parlays. His results are consistent with the hypothesis that both bets are identically priced and therefore, that the efficient market hypothesis cannot be rejected.

Asch and Quandt (1987), employing data from 705 races at the U.S. Meadowlands racetrack between May and August 1984, performed a similar exercise to Ali (1979), comparing the returns to winning daily doubles with the returns to the corresponding parlays. They found the daily double bets to be significantly more profitable than the parlays, although this difference loses its statistical significance if allowance is made for the fact that the track take is applied twice to the parlay (since this consists of two separate bets) but only once for the daily double.

Using data from Meadowlands (1984) and Hong Kong (1981-1989), Lo and Busche (1994) compared the mean returns to various types of double bets with those for corresponding parlays. Although they found that the various types of double revealed a higher expected return than the equivalent parlays at conventional levels of statistical significance, the difference disappeared if allowance was made for the differential track takes associated with the different types of bet. Taken as a whole, their findings on the difference in the expected payoffs to doubles and to the corresponding self-constructed parlays are consistent with those of Asch and Quandt (1987). Evidence from the U.K. tells a different story, and is based upon a comparison of the two types of betting market which co-exist in the U.K., i.e. the Tote (totalisator) and the bookmakers. This evidence is discussed below.

4.2.2 Employing expected returns to bets at 'tote' odds and at starting prices to test for semi-strong form efficiency

In the Tote system, which is the British version of the parimutuel, winning bets share the pool of a fixed proportion of all bets. The final dividend is not known with certainty until after the race. Bookmakers on the other hand set fixed odds, although the Starting Price is often taken off-course, which is the price at which a sizeable bet could have been placed with bookmakers on-course at the start of a race. This division of the market provides a useful test of information efficiency, the expectation being that in an efficient market any differences in Tote and bookmakers' prices should be eliminated by the end of trading (when starting prices are returned). Although there is some evidence that this does occur in Australia, which possesses an off-course tote monopoly (see Bird and

McCrae, 1994), the U.K. evidence, which can be traced to Gabriel and Marsden (1990, 1991) points to a semi-strong form inefficiency. Specifically, Gabriel and Marsden compared the returns to two types of betting; first, bets struck with bookmakers at the starting price; and second, bets placed with the Totalizator Board at 'tote' odds.[1] The starting price was chosen in preference to any other agreed fixed odds price since it possesses the dual strengths of forming a significant part of the betting market, and also being fully and properly recorded. In particular, the data can be regarded as accurate, widely available and easily accessible. Comparing the odds implied ex-post in the starting price and in the tote return, Gabriel and Marsden's main point is that "Since the differing bets are two options for purchasing exactly the same item (a bet to win on a specific horse), we would expect the odds to converge." (p.877)

They tested this hypothesis using data drawn from the first 1427 flat races of the 1978 racing season in England. The year 1978 was chosen because the general absence of mechanical or electronic tote boards in that year limited the information available to bettors on betting patterns or likely final odds. The idea here is that in order to equalize the risk of betting with the tote or with bookmakers, bettors should be equally uncertain under either betting system of the exact odds until after the race starts.

"Thus a rough test of such [semistrong] market efficiency is simply to compare the average tote and starting price payoffs after races." (p.878)

In fact, an examination of the difference in the mean tote and starting price, employing standard t-tests and a Wilcoxon matched-pairs signed-rank test (because it requires no assumption about the shape of the underlying distributions), revealed a significantly greater expected return to the aggregate of bets placed on the tote.

To test for the possibility that the differences could be explained in terms of a few very large tote payouts, they deleted tote payouts above a certain level. Although the average difference fell as the size of the payout decreased, they still found that the difference persisted at a statistically significant level.

To test for a learning effect as the season developed (which would tend to reduce the importance of insider information) they performed the calculations for three successive periods. While they found lower average tote and starting price payouts as the season progressed, as well as generally lower differences between tote and starting price payouts in the later periods, these were not sufficient for semi-strong efficiency. Moreover, the differences after excluding higher tote payoffs (specifically tote odds of more than 20 to 1) remained not only statistically significant, but also showed no indication of convergence.

How far this evidence of semi-strong efficiency bears testimony to strong form efficiency, i.e. efficiency with respect to all information including insider or private information, is a matter of interpretation. Gabriel and Marsden argue that insiders would tend to avoid tote betting in favour of placing bets at fixed odds, the effect of which would provide a market signal to bookmakers to cut the odds, so depressing the starting price relative to earlier prices and to the tote odds.

Even with insider information, in an efficient market one might in any case expect the market to absorb evidence of such information, and market participants should assimilate this into their choices of how and where to bet. Thus, if starting prices are being artificially depressed relative to tote odds, rational bettors should switch their bets from bookmakers to the tote, and vice-versa when the opposite occurs. This should lead to convergence. The fact that Gabriel and Marsden's results suggest it does not lead to such convergence is prima facie evidence of some form of market inefficiency.

Ultimately, though, as Gabriel and Marsden themselves accept, "...is it ever possible to...separate an inefficient market from one in which the participants are pursuing the satisfaction of nonmonetary preferences?" (p.885)

A further test of inefficiency is included by Gabriel and Marsden, following de Leeuw and McKelvey (1984) and Zuber et al. (1985). This is an estimation of the parameters of the following equation:

$$TOTE_i = \alpha_0 + \alpha_1 SP_i + \mu_i ,$$

where $TOTE_i$ is the tote payout in the ith. race, SP_i is the starting price payout in the ith. race, and μ is the error term.

Following Zuber at al. (1985, pp.800-801), they use a standard F-test to test the null hypothesis that $\alpha_0 = 0$ and $\alpha_1 = 1$ jointly.

Applying this test for the whole season, for races through April, through May and through June separately, they found that the null hypothesis is rejected for all these data sets at the one per cent level of significance, providing additional evidence of market inefficiency.

Gabriel and Marsden (1991) corrected their 1990 conclusions to rectify an error in calculating the tote returns in their sample of Irish races, which led to an over-statement of these returns. Allowing for this, and re-calculating their original figures, they reported continued broad support for the proposition that tote returns are on average significantly better than starting price returns, although they are unable to reproduce the finding at all the original levels of sub-aggregation. They concluded that "Simply put, the corrected results are not as strong as those reported earlier." (p.564)

Blackburn and Peirson (1995) and Vaughan Williams and Paton (1997c) provide further evidence of significant differences, and show that starting price returns are actually superior to tote returns at lower odds.

Unless a convincing explanation can be offered for these persisting differences, they constitute evidence in a limited sense of information inefficiency at a semi-strong form level. It is not clear, however, that this information is capable of being traded upon so as to earn abnormal returns or positive profits. To this extent, any rejection of the hypothesis of semi-strong information efficiency in these markets is less satisfactory.

The Australian evidence provides somewhat contrasting findings. Bird and McCrae (1994) found for their Australian data that any difference between bookmaker and parimutuel (totalizator) odds tended to evaporate as the start of a race approached. Whereas the totalizator take was about seventeen per cent, the bookmaker take, as implied by the market odds (the 'over-round') varied from about 26 per cent at the opening of the market to a level roughly equivalent to the tote 'take' at the off, i.e. about 17 per cent. Even so, the starting prices laid by bookmakers tended to be lower than those available 'on the tote' in those cases where the odds lengthened, and tended to be higher in those cases where the odds contracted, although this pattern could not be used to create a profitable betting strategy. Bird and McCrae (1994) also found that most of the odds movement occurred in the first half of the betting fraction, implying perhaps that insiders are not strict monopolists of superior information, and so bet early at advantageous odds before these odds disappear.

4.3 Testing for the existence of a 'gambler's fallacy'

The gambler's fallacy is the proposition that bettors, instead of accepting an actual independence of successive outcomes, are influenced in their perceptions of the next possible outcome by the results of the preceding sequence of outcomes, e.g. throws of a die, or spins of a wheel. Terrell (1994) states it thus:

"The 'gambler's fallacy' is the belief that the probability of an event is decreased when the event has occurred recently, even though the probability of the event is objectively known to be independent across trials." (p.309)

This idea was proposed in a generalized form by Kahneman and Tversky (1974, 1982). Their notion of a 'winner's blessing and loser's curse', as it is commonly known, is a reported tendency for people, in revising their beliefs, to overweight new information and underweight older information. Such a hypothesis has been extensively tested in financial markets, e.g. De Bondt and Thaler (1985), Brown and Harlow (1988) and Stock (1990). Each of these studies found evidence of the existence of such an 'anomaly', and have yielded, therefore, the idea of trading upon a contrarian strategy. Insofar as such a strategy is based on the historical pattern of past prices, it provides at least prima facie support for a hypothesis of weak form efficiency in these markets. Since it also implies a failure by traders to allow for all public information, it is also indicative of semi-strong form inefficiency. However, some authors (e.g. Chan, 1986) explain any above-average or abnormal returns which can be elicited by acting on the basis of the above as simply fair compensation for additional risk or other factors.

The existence of a 'gambler's fallacy' has been documented in laboratory studies (Ordohook and Morrissey, 1984); lottery-type games (Clotfelter and Cook, 1991, 1993); and lotteries (Terrell, 1994). In particular, Clotfelter and Cook (1991, 1993) found (in a study of a Maryland numbers game) a significant fall in the amount of money wagered on winning numbers in the days following the win, an effect which did not disappear entirely until after about sixty days. This particular game was, however, characterized by a fixed-odds payout to a unit bet, and so the gambler's fallacy had no effect on expected returns. In parimutuel games, on the other hand, the return to a winning number is linked to the amount of money bet on that number, and so the operation of a systematic bias against certain numbers will tend to increase the expected return about those numbers. Terrell (1994) investigated one such parimutuel system, the New Jersey state lottery. In a sample of 1785 daily drawings from 1988 to 1993, he constructed a sub-sample of 97 winners which repeated as a winner within the sixty day cut-off point suggested by the Clotfelter and Cook (1991) findings. He found that these numbers had a higher payout than when they previously won on 80 of the 97 occasions. In order to determine the relationship more precisely, Terrell also regressed the payout to winning numbers on the number of days since the last win by that number. The expected payout on a number increased by 28 per cent one day after winning, and decreased from this level by about 0.5 per cent each day after the number won, returning to the original level after sixty days or so. The size of the gambler's fallacy observed in New Jersey, while significant, was nevertheless not as great as that found by Clotfelter and Cook (1993) for the fixed-odds Maryland numbers game. It is as if irrational (certainly non-profit-maximising/loss-minimizing) behaviour exists, but reduces as the cost of the anomalous behaviour increases.

Two studies of a 'gambler's fallacy' in racetrack betting found the same effect; Metzger (1985) in U.S. horserace betting, and Terrell and Farmer (1996) in U.S. greyhound racing.

Metzger (1985) set out to test one prediction consistent with the concept of a gambler's fallacy, specifically that bettors will tend to overestimate the chances of a favourite winning after a series of wins by longshots compared to the situation after a series of wins by favourites. On the basis of an examination of a sample of U.S. horse races, Metzger concluded that there was indeed such a tendency shown by bettors in the aggregate. In particular, a series of wins by favourites (longshots) made less (more) favourable a bet on a favourite, which in turn produced underbetting (overbetting) of favourites. Terrell and Farmer (1996) calculated the return to a strategy of betting the greyhound in the starting trap occupied by the previous winner. This yielded a 9 per cent profit, and as such constituted "... the only strategy earning positive profits" (p.864). The finding is consistent with the hypothesis that bettors were under-estimating the probability of a repeated outcome and that as such, they were victims of a gambler's fallacy.

If confirmed, such trends or patterns could be exploited, at least in principle and perhaps in practice, to yield above-average or abnormal returns. If such configurations can be shown to exist in racetrack betting markets at an appropriate level of confidence and to constitute more than fair compensation for other factors, such as changes in the incidence of risk, then this constitutes potential evidence in contradiction of the existence of informationally efficient betting markets.

4.4 Fundamental strategies and tests for semi-strong form efficiency in racetrack betting markets

In this section evidence is derived from racetrack betting markets which addresses the issue of whether bettors can apply decision rules, based on fundamental information, which can be employed to earn above-average or abnormal returns. This evidence is used to draw conclusions as to the existence of semi-strong information efficiency in these markets.

Hausch, Ziemba and Rubinstein (1981) offered a convenient interpretation of 'fundamental strategies' as decision rules which "utilize past data available from racing forms, special sources, etc. to 'handicap' races. The investor then wagers on one or more horses whose probability of winning exceeds that determined by the odds by an amount sufficient to overcome the track take." (p.1435).[2] Similarly, Benter (1994) defines the idea of 'fundamentally' handicapping a race as "... to empirically assess each horse's chance of winning, and utilize that assessment to find profitable wagering opportunities." (p.183). Such strategies are different from the systems based on utilizing information contained in the betting odds, and usually more complex to make operational.

Vergin (1977) examined six such strategies, in the form of published betting systems[3], on the basis of a sample of 102 races run in January and February 1972 at Santa Anita Park, California. Of these, only one, i.e. the McQuaid (1971) elimination rule, produced a profit to level stakes, and that only for win bets. McQuaid's "elimination rule", reproduced in Vergin (1977) is stated below:

"Eliminate any horse which:

a. has not had one race at today's track;

b. has not run today's distance at today's track (+ 1 furlong);

c. has not raced within one month of today's race;

d. has not won a race;

e. in its last race did not finish in the money;

f. did not finish within eight lengths of the winner in its last race;

g. in its last race lost more than 3/4 lengths in the stretch;

h. has a speed rating at today's distance which is not within five points of the highest speed rating for any of the competing horses for the past four races;

i. has a consistency rating which is not within five points of the highest consistency rating for any of the competing horses unless the horse's speed rating is as high or within one point of the highest speed rating." (Vergin, p.43).

McQuaid proposed three ways of improving this rule:- eliminating low odds bets, deleting ineffective individual elimination rules, and adding rules drawn from other sources where this would reduce the number of non-winning horses predicted as winners.

Although this increased the return on investment from 17 per cent to 78 per cent, this is hardly surprising since the returns are calculated using data from which the modifications were derived. Applying the same system to new data still yielded a return above what could be expected by chance, although less than that yielded by the original data. Nevertheless, even after modifying the amended system in order to generate more bets he admits that "this sample is still too small for anything approaching the level of statistical tests of significance..." (p.44).

Canfield, Fauman and Ziemba (1987) questioned whether trading rules based on a knowledge of one example of a persistent bias in racetrack outcomes could be used to earn abnormal returns. In particular they tested for the existence of a post-position bias, employing a sample of 3345 races at Exhibition Park, Vancouver, Canada, between 1982 and 1984. As such they built upon the ideas of Quirin (1979) and Beyer (1983). Beyer noted that "At tracks less than a mile in circumference, the sharp turns and short stretch almost always work to the advantage of the front runners and the horses on the inside." (p.42). An examination of mile racetracks by Quirin (1979) indicated that the inside six positions produced winners more often than would be expected by chance, with the inside under most conditions the most advantageous of all.

Canfield, Fauman and Ziemba examined win bets for the whole sample, and for sub-samples comprising fast tracks (which might favour the inside post) and off tracks (in which the wet conditions might disadvantage the horse holding the inside position) separately. They also examined a sample of longer races in order to test the hypothesis that the greater number of turns in longer races might make the bias against outside positions particularly pronounced, and assessed the influence of the circumference of the track in eliciting the same bias. Although a strategy of betting on particular post positions, under particular track conditions, could be devised which appeared to offer positive profits, it was not possible to reject a hypothesis of non-positive profits in the long run from any such strategy at conventional levels of significance. Their explanation of the inability to translate post position bias into significant net profits rests on the idea that, after allowing for track deductions, bettors overbet the favourable positions to the extent of negating any potential advantage from a betting strategy based on the bias. Applying their data to an examination of the incidence and consequence for expected returns of post-position bias in the market for 'exotic' bets, they reached similar conclusions.

Betton (1994), in an analysis based on 1,062 races from the same racetrack as Canfield, Fauman and Ziemba (1987), re-examined the incidence of post-position bias. A t-test was employed to compare the average post position of the first three horses finishing with the average post position expected if no post position bias existed. The bias was found to be significant for the first two places but not for the third.

Betton concluded that while "knowledge of the post position significantly improves the information available from the odds rankings, the relatively low overall explanatory power of these models suggests that more is unknown than known in the determination of racing results." (p.520).

Terrell and Farmer (1996) employed a large data set of 4,121 races (at the Woodlands Greyhound Park, Kansas City) for the 1989-1990 season, and follow-up data from 1994, to investigate the post-position 'anomaly.' They found a consistent pattern over time of significant differences in the win rates of starting boxes, and the expected returns to bets on these boxes. The best return was to bets on dog one (the 'inside trap'), the worst to dog seven (the 'outside trap'). The implication is that bettors tended to over-bet the outside traps, although this disparity was not sufficiently strong so as to yield a positive return (net of deductions) to bets on the inside traps.

Bolton and Chapman (1986) developed a stochastic utility model, parametrized in the form of a multinomial logit model, employing horse, jockey and various race characteristics (ten basic horse and jockey independent variables in total), to evaluate the worth (or 'utility') of racehorses. Their database consisted of two hundred races reported in the publication, the Daily Racing Form. Finding that the signs of the coefficients on their explanatory variables were consistent with their 'a priori' theoretical expectations, they tested the usefulness of their model across various alternative betting strategies. By constraining wealth to a level at which bets have negligible influence on the track odds, they were able to obtain expected returns significantly greater than would be generated by a random betting strategy, though not sufficient to enable a positive expected return to be earned across the sample of races. Applying various side constraints they eventually decided to eliminate those bets for which the logit model provided poor estimates of the winning probabilities, i.e. longshots. In other words, bettors should confine their bets to horses whose estimated probability of winning exceeds some minimum value, which they estimate. They were able to express, subject to further modifications, some optimism that positive returns could be made from a specified betting strategy, although the strength of their conclusions are limited by their small sample size and the existence of positive estimation errors.

Chapman (1994) extended this type of analysis, using a much larger database with a larger number of covariates. In particular, he applied a twenty variable pure fundamental multinomial logit handicapping model to two thousand races in Hong Kong[4], which included handicapping variables linked to horse, jockey, situational context (e.g. track and distance conditions) and observable current performance signals (e.g. recent track workouts). There was no evidence of the usual longshot bias, in line with analyses of the Hong Kong market by Busche and Hall (1988) and Busche (1994), and there was a close correspondence between the expected win frequencies and the actual winning frequencies. Chapman tested for the possibility that positive profits could be made on the basis of the sophisticated handicapping model he proposed, concluding that there was a potential to make positive expected returns from such a model. These findings are consistent in principle with the results published by Bolton and Chapman (1986), but much greater in size. Moreover, the returns are found to be higher still if the logarithm[5] of the public's win probabilities, as revealed by their actual win bets, is included as an additional (i.e. 21st.) variable in the multinomial horse race handicapping model. To make the latter operational as an optimal betting system in a practical sense would, nevertheless, require replacing in real time the estimates of the true win probabilities with the final public win bets, and acting upon the final information set.

Employing the same data, Benter (1994) constructed a computer model designed to estimate current performance potential. This involved the investigation of variables and factors with potential predictive significance and the refining of these individually so as to maximize their predictive accuracy. In order to counter any tendency to overfit on the basis of past data, Benter was careful to test his refinements on out-of-sample data. The sole criterion for inclusion of a variable as a predictor of performance was, therefore, improvements in the model's goodness-of-fit at an acceptable degree of statistical significance. This was the paramount consideration even when any theoretical explanation of the variable's effect was either missing or else unconvincing. Implementing a betting strategy based on this model, he reported an overall net profit for five years of large scale betting, although a loss was reported in one of the five seasons. He concluded that "... at least at some times, at some tracks, a statistically derived fundamental handicapping model can achieve a significant positive expectation." (p.196). As such modelling becomes more widely available, however, he was led to suspect that the market will become efficient to such predictions, i.e. "The profits have gone, and will go, to those who are 'in action' first with sophisticated models." (p.196).

In summary, although there is an array of evidence which suggests that forecasting strategies are capable to some degree of forecasting racetrack outcomes, less evidence is available that such in improvement can be used to make abnormal returns. Where such evidence does exist, the strength of the findings is either linked to the ability to operationalize a model incorporating changing variables in limited real time, or else subject to variations in the return which may permit extended short-term losses. Even if these models are accepted as ex-post evidence of the possibility of earning abnormal profits from a fundamental handicapping system, bettors seeking to operationalize the model in the future are subject to the risk that the findings have already been incorporated into future odds.

4.5 Racetrack forecasts and betting market efficiency: An appraisal of theory and evidence

The relationship between racetrack forecasting services and the concept of information efficiency in betting markets in which such forecasting services operate, is an example of the need for a clear distinction between the semi-strong and strong forms of information efficiency.

This issue is highlighted by Snyder's (1978a, 1978b) tests of five forecasting services, which he termed "five strong tests of market efficiency" (1978a, p.1117). This designation was subsequently disputed by Losey and Talbott (1980), who argued that Snyder's "experts" are not "insiders" in the sense of possessing monopolistic access to information. To the extent that the odds which they quote are based on publicly available information, and are published and disseminated widely before each race, Losey and Talbott compare their role to that of advisory service financial analysts. In this sense, they argue that Snyder's tests may be more accurately viewed as tests of semi-strong efficiency.

It seems that there is a need, therefore, to distinguish between forecasts which are widely available, such as the tips provided in a national daily newspaper, and those which are provided to a small group of clients who normally pay for this service. There is no clear dividing line, however, and the strength of the test for efficiency is inevitably, therefore, a matter of interpretation. In this chapter the tests are of widely available forecasts, and the conclusions must be interpreted in light of this.

4.5.1 Efficiency and the information content of publicly available racetrack forecasts

Work on the information content of racetrack forecasts can be traced to Snyder's tests of the hypothesis that individuals or groups with special or expert information are able to outperform the general betting public. To do this Snyder compared the performance of "expert"[6] advice with that of actual bets struck. To do so he calculated the rate of return to bets placed at parimutuel odds and also at the odds forecast by the "experts", using these as if they were the real payout odds. He found that the experts' rate of return was both larger at low odds and smaller at high odds than that of the public. Indeed, all of the 'experts' displayed a greater bias than the public.

Losey and Talbott re-examined Snyder's work using a similar data set. Their objective was to determine whether and to what degree a knowledge of the divergences noted by Snyder could be used by bettors to improve their expected return. The point is that if the market imperfectly incorporates all the information supplied by experts, a bettor acting upon such imperfections may be able to make above-average returns. As such, a test for the existence of such an imperfection can be used to test for the existence of market efficiency. Losey and Talbott identified those cases in which the parimutuel odds exceeded the odds listed by the "Racing Form" handicapper, the extent of this divergence being calculated as the simple ratio of the final parimutuel odds to the handicappers' odds. This is termed the 'overlay ratio'. The idea behind their approach is that if the handicappers do possess superior information, then higher than average expected returns should be derivable from betting on horse with higher than average overlay ratios. In fact, they observed that bettors employing such strategies tended to earn lower than average returns. Losey and Talbott conclude that parimutuel markets are inefficient in the sense of reflecting all available information, but that they have insufficient evidence to claim the existence of inefficiency in the sense of a systematic profit potential.

A different type of analysis of the relationship between expert opinion and betting odds was offered by Figlewski (1979), who employed a multinomial logit probability model to assess the information content of professional handicappers' published forecasts, and related the observed frequency of winning to the handicappers' odds and the final odds. The study is based on an examination of data on win bet odds from 189 thoroughbred races run at Belmont racetrack, New York, in June and July 1977. Figlewski observed that while substantial information was contained within the handicappers' advice, most of this was already incorporated into the betting odds. It could not, therefore, be used to improve significantly the forecast accuracy of the betting odds as determined in the market. While he found that on-track bettors discounted the information in full, this was not true of off-track bettors. The implication is that on-course bettors at least were able to attach the appropriate weight to the handicappers' information in placing their bets.

Studies of the relationship between expert opinion and betting odds, using Australian data, offer mixed conclusions. The case for semi-strong form inefficiency in Australian racetrack markets can be traced to Tuckwell (1983), who showed in his data set how a strategy of betting on horses whose market odds deviated significantly from forecast odds could be used to generate positive returns. This finding has not, however, been reproduced in later studies. In particular, Anderson, Clarke and Ziegler (1985), employing a two-stage regression, found that such information was indeed impounded into the market odds at the off, and knowledge of the forecast prices could not be used to generate profitable trading rules. Bird and McCrae (1987) examined "expert" opinion in the form of the forecasts supplied by ten so-called experts which were published on the morning of each race in a Melbourne newspaper. Since all the newspaper tipsters provide a first, second and third selection in each race, these selections were pooled and aggregated on the basis of three points for a first selection, down to one point for a third selection. The rate of return was calculated for a strategy of placing level stakes on horses ranked on the basis of the poll of these experts. The equivalent level stakes strategy was applied to horses ranked at a particular level of favouritism. No strategy based on the above produced a positive rate of return. Neither were the rates of return from bets placed on horses ranked according to the experts' poll significantly different from those based on the level of favouritism or the odds. Bird and McCrae (1987) concluded that the betting market is efficient with respect to information supplied by the newspaper experts, and that this information is incorporated into the bookmakers' odds.

Thus, although there is some evidence that genuine information is contained in forecast prices, there is less evidence that such information is unincorporated into the payout odds available in the market, and even less that any such unincorporation can be systematically exploited so as to yield abnormal returns.

4.6 Inside information, insider trading and tests of strong form efficiency

In this section betting markets are investigated for the existence of inside information which can be traded upon to earn above-average or abnormal returns. Conclusions are drawn from the evidence for the existence of strong information efficiency in these markets.

In order to test for the existence of strong form efficiency in the Fama (1970) sense[7], i.e. efficient with respect to all information including private information, Dowie (1976) noted the importance of drawing the distinction between odds available or predicted at differing times. His methodology centres on the fact that the starting price, being determined extremely late in the market, should incorporate more information than any of the various odds generated early, and should include any inside or monopolistically held information. Indeed, since those possessing such information will exploit this continuously up to the point at which the starting price is determined, then the correlation between the starting price returns and the realized probabilities should, in the presence of insider activity, be higher than that between any odds formed earlier in the market and these probabilities. Certainly it should be higher than for any odds assigned before the market is formed. To test this Dowie used a very large data set of races, covering the 1973 flat racing season (29,307 horses in 2,777 races). He calculated two relevant correlations. The first was between the realized probabilities and the starting prices, while the second was between the realized probabilities and odds forecast before the formation of the market (the *Sporting Life* forecast was chosen). Because the over-round is not standard as between starting prices and forecast prices, he standardized these prices for each race, adjusting the prices by a percentage implied in the respective over-rounds so as to generate hypothetical perfectly round books (i.e. no over-round or under-round). He found no significant difference between the two sets of correlations, and therefore, whilst acknowledging data deficiencies at the short end of the odds market, concluded that there exist "serious doubts as to the significance of inside information." (p.150)

Crafts (1985) extended Dowie's study to allow for the possibility of profitable arbitrage. Crafts' point is that insiders may take advantage of market dynamics to bet at odds greater than the starting price. Although this is not an option in a parimutuel system, it is in a bookmaking system, where bettors can 'take a price' at any time during the course of the market. To allow for this, Crafts proposed a test based on separating out cases where there is a "marked" difference between the forecast price and the starting price. He reasoned that a shortening (lengthening) of the odds available about a horse during the course of the market may indicate evidence of insiders who knew the true probabilities of that horse winning were greater (less) than those implied in the odds offered early in the market. He applied an equivalent test also to marked differences in the prices actually offered on the course during the period of trading. The data set consisted of 16,769 runners in total, over a period from 11 September 1982 to 8 January 1983, and employed various arbitrary definitions of a marked difference, e.g. forecast price greater than or equal to twice the starting price. In all, he identified 2,280 horses (13.6 per cent of the total sample) which fell into one of his categories. He also cleaned the data to allow for idiosyncrasies in the compilation of the forecast prices. For example, he eliminated all observations which did not have either a starting price or a forecast price of 10/1 or more, because of the tendency to collate higher forecast prices into a residual category (the 'bar'), and also cases where the sum of implied probabilities in each of these set of prices did not diverge from each other too greatly.[8] If insider trading is the cause of these 'marked' price movements, then where these marked differences occur the expected return to bets placed earlier in the market should be significantly different to those placed later in the market. Moreover, this should be even more the case in races which offer particular scope for insider trading.

In order to isolate races which might be the target of above-average levels of insider activity, Crafts divided the sample into handicap and non-handicaps, the idea being that these are distinguished by the amount of established public form available about them. In handicap races horses are weighted by past form (which must be established over a series of races) in an attempt to equalize their chances. Such races are less likely, therefore, to offer as much scope for insiders to trade as non-handicap races, where the form need not be so exposed to public scrutiny. Crafts' results suggest that horses displaying a marked shortening of the odds between the forecast and starting price stages are indeed characterized by an exceptionally high expected return at forecast prices. Moreover, this is particularly strong for non-handicap races. An examination of the scope for profitable arbitrage during the formation of starting price odds in the actual on-course market produced similar findings, although splitting the sample as before (into handicaps and non-handicaps) failed to reproduce the earlier result.

Crafts offered supporting evidence in descriptions of betting patterns, published in the *Sporting Life*, about the previous day's on-course trading. In particular, he identified examples of large sums of money placed on horses with poor previous form, which went on to shorten considerably in the market and to win.

On the basis of this evidence, Crafts concluded that British horse race betting markets do offer insiders the opportunity for profitable exploitation of information not publicly available. As such, it is not strongly efficient.

Crafts (1994) developed this work by identifying a category of horses about which inside information is likely to be particularly useful, i.e. horses which have not run for a long time (the season before last). As in his 1985 paper, he used the existence of a 'marked' shortening of the odds ((forecast price/starting price>1.5) to indicate positive insider trading. Presumably to eliminate the potential negative bias to returns implied by the favourite-longshot bias, all horses starting at odds of 7 to 1 against or greater were eliminated from consideration.

For 88 observations taken over a period between September 1982 and November 1987, the rate of return was 261.9 per cent at forecast prices, and 55.8 per cent at starting prices. The return at forecast prices would appear to contradict strong form efficiency, and the return at starting price to contradict the existence of less strong forms of efficiency. The implication is that a good strategy for outsiders in the British horserace betting market is to identify and selectively follow information provided by insiders. If Crafts' results are significant and reproducible, then those seeking to continue their advocacy of information efficiency in these markets may do worse than accept Crafts' lifeline, i.e. that there are substantial costs incurred in collecting the data required to make the decision rule operational.

Crafts also examined the existence of market efficiency in British betting markets by analyzing the performance of three marketed betting systems. Each of the systems provided special decision-making rules based on public information, and were tested on data generated prior to the publication of the systems, and again for data generated after publication. No evidence of significant post-tax profits were revealed either before or after publication, although the return to one of the systems, which showed some evidence of pre-tax profitability, declined markedly after publication. This is indicative, perhaps, of a market response to knowledge contained therein. Further tests of nine racing systems and five ratings services (i.e. services providing tips rather than decision rules), based on results reported in Roberts and Newton (1987), also failed to show profits. These results, Crafts (1994) noted, are consistent with the operation of the Efficient Markets Hypothesis in British betting markets. He concludes:

"The continued sale of these systems suggests that the participants in British horserace betting include many gullible outsiders." (p.547).

Because in parimutuel markets it is not possible to arbitrage at different prices, it is more difficult to test for the existence of insider trading. However, Asch and Quandt (1986) do offer some evidence. Employing data from 729 races at the Atlantic City racetrack, they report that the final parimutuel odds tended to be lower than the predicted (or 'morning line') odds for winners. Moreover, in the case of winners, the later in the betting period the money was laid the stronger was this effect. Asch and Quandt proposed an explanation of this finding as the withholding of "smart money" (money bet by people with superior information) until late in the betting period, in order to avoid giving out market signals which could depress the final payout odds.

The influence of inside information on betting markets was also the subject of an investigation conducted by Schnytzer and Shilony (1995). This approach compared two mutually isolated groups, one of which they propose has access to inside information and one of which does not. The informed group are identified as on-course bettors in Australian (Victoria) racetrack betting markets who are able to detect a significant shortening of the odds offered by on-course bookmakers about particular horses, an effect consistent with a 'plunge' on these horses by insiders. Although they may miss the value with bookmakers by the time they identify the 'plunge' there may still, it is hypothesised, be opportunities to benefit from this second-hand information by betting on the 'tote'. Off-course bettors, on the other hand, are isolated from the information about the on-course bookmaker market, as the 'tote' in Australia has a monopoly of off-course betting. They found, using a standard and a Spearman rank correlation coefficient, that the betting behaviour (in terms of the proportions bet on given horses) of their two groups was significantly different, the on-course betting population accruing a 13 per cent advantage compared to the off-course bettors, and also tending to bet more on longshots. In order to distinguish whether this is due to differences in the relative availability of inside information or to differences in the ability of the two populations to understand and process public information, they calculated the predictive power of the two populations, and compared the variance of the predictions. Using a multinomial logit model, they estimated that the variance of the prediction by those with access to both public and private information was less than that by those with access only to public information. This is, they conclude, consistent with the hypothesis of valuable inside information. A separate test of the value of inside information was to estimate whether proxies for positive and for negative inside information are statistically significant predictors of race outcomes. Evidence that a significantly greater proportion of money is bet by on-course bettors on a particular horse is used to proxy for positive inside information, and vice-versa. If there is useful positive private information about a horse it might be expected, they argue, that the ensuing 'plunge' would cause a shortening of that horse's odds. This information should constitute a useful predictor of race outcomes. They found that it did. Negative inside information about a horse may also cause a lengthening of its odds, but odds lengthen also as an attempt by bookmakers to attract money away from the 'plunge'. Because of this ambiguity they argue that proxies for negative inside information cannot readily be used to predict race outcomes. They found that this was indeed the case.

This evidence is adduced in further support of the contention that differences in betting behaviour between on-course and off-course bettors is explained by differences in their access to inside information rather than in their capacity to process public information.

It is proposed here that this reasoning is faulty in principle, although it may be true in practice. Subsection 4.6.1 explores this issue.

4.6.1 Inside Information in a Betting Market: A Comment

Schnytzer and Shilony (1995) investigate the influence of inside information in an Australian racetrack betting market. To do this they compared two mutually isolated groups, one of which (on-course bettors) has access to inside information and one of which (off-course bettors) does not. Evidence that a significantly greater proportion of money is bet on a particular horse by on-course bettors than by off-course bettors was used to proxy for positive inside information about that horse, and vice-versa. If there is useful positive private information about a horse, they argue that the ensuing 'plunge' would cause a shortening of that horse's odds. This information should constitute a useful predictor of race outcomes. Negative inside information about a horse may also cause a lengthening of its odds, but odds lengthen also as an attempt by bookmakers to attract money away from the 'plunge'. Because of this ambiguity, they argue that proxies for negative inside information cannot readily be used to predict race outcomes.

It is argued here that this reasoning is faulty, since it makes an unwarranted assumption of asymmetry. Where negative inside information about a horse causes its odds to lengthen, bookmakers seeking to balance their books are just as likely to shorten the odds of others. This creates a signal which may be no less ambiguous than that contained in a lengthening of the odds. On Schnytzer and Shilony's implicit assumption of an asymmetry in the odds-setting process, the bookmakers' over-round (or sum of prices) should be significantly lower in races in which a horse's odds have lengthened than in those other races where a horse's odds have shortened.

This is tested using bookmaking data in a sample of 4689 horses. Table 4.1 compares the over-round for races in which a horse's odds have increased from Forecast Prices to Starting Prices (FP < SP) with those whose odds have shortened (FP > SP). FPs are being used here to proxy for earliest prices. They are normally published prior to the start of trading and, as such, are likely to be free of the influence of inside information. For the whole sample the over-round is, in fact, *higher* in the case of the former. Even controlling for the influence of the number of runners on the over-round, in only four out of 22 cases is the over-round significantly lower (at the 5 per cent level) for horses whose odds have lengthened. This is consistent with these signals of positive and negative inside information being equally ambiguous.

Table 4.1: Over-round where Odds Lengthen compared to Over-round where Odds Shorten

	FP < SP			FP > SP		
Runners	**Over-round**	**SD**	**N**	**Over-round**	**SD**	**N**
All *	127.91	11.98	2858	127.13	12.33	1318
3	108.23	2.044	13	107.68	2.029	11
4	110.28	3.382	25	110.41	3.228	18
5	110.59	2.156	75	111.07	2.054	36
6	112.43	2.852	109	112.81	2.636	55
7	114.73	5.066	108	115.27	4.727	63
8	116.70	3.557	258	116.84	3.656	152
9	117.76	5.936	142	118.14	5.956	82
10	121.01	6.197	206	121.22	6.485	114
11 ***	122.52	7.757	148	125.79	9.077	76
12	126.49	6.132	188	127.72	5.926	86
13 **	129.62	7.173	177	131.85	7.763	90
14	130.51	7.625	257	131.06	8.519	113
15	132.05	6.699	239	131.80	7.329	81
16	133.86	8.457	210	134.98	8.953	77
17	138.30	8.127	130	140.10	7.580	37
18 **	138.59	7.383	127	141.03	7.788	63
19	137.01	9.455	104	136.24	10.24	42
20 ***	139.85	9.760	164	144.61	12.61	53
21 *	144.39	6.432	79	142.02	7.382	38
22	145.89	8.981	37	144.40	7.090	15
24	136.69	4.924	32	136.04	5.643	7
25 **	159.22	1.361	21	157.78	1.314	5

Note:
* Indicates that the t-statistic, testing for difference in over-round on a two-tailed test, is significant at 10% level; ** at 5% level; *** at 1% level.

As a further test, the over-round is regressed on a dummy variable (Up) for those horses whose odds increase from Forecast Prices to Starting Prices, including both the number of runners (rn) and its interaction with "Up" as control variables. This gives the following result:

$$\text{Over-round} = \underset{(1.029)}{101.02} + \underset{(0.703)}{0.057}\,Up + \underset{(0.100)}{2.080}\,rn - \underset{(0.067)}{0.051}\,rn \times Up \qquad (4.1)$$

Adj R^2 =0.643; N=4689

Figures in brackets are standard errors, adjusted due to clustering (within races) in the sample and are robust to heteroscedasticity.

Both the constant and the coefficient on rn are different from zero at the 5 per cent level of significance.

The positive (although insignificant) coefficient on the "Up" variable again suggests that there is little evidence that the over-round is lower in races where a horse's odds have lengthened.

Whatever the merits of this argument, Schnytzer and Shilony believe that a judicious use of information contained in on-course/off-course betting patterns can be used to secure positive profits net of deductions. One such method they advocate is to "compare the last off-course tote odds with the bookmakers odds a minute or two before the end of betting and back those horses for which the ratio of bookmakers' odds to tote off-course odds is greatest, in proportion to that ratio." (p. 970). In an exercise they undertook on this basis of hypothetical betting results, they found that betting on horses where there existed a marked difference between the on-course and off-course win bet ratios at the time of the race gave a net profit up to 32.5 per cent. Whatever the difficulties may be of operationalizing this in real time, this is certainly evidence of inside information in the Australian betting market under consideration.

4.6.2 Measuring the incidence of insider trading in a betting market

An approach which seeks to estimate the actual extent of insider trading can be found in Shin (1993). This analysis models the betting process in a fixed-odds system along the lines of Shin (1991, 1992), and brings the debate full circle, to the earlier consideration of the reasons for a favourite-longshot bias in these markets. Shin (1991, 1992) first proposed that the bias could be explained not only by the demand-side influences of bettors, but in fixed-odds betting markets also as the result of an optimal pricing strategy by bookmakers who face a number of bettors who possess superior information to that publicly available ("insiders"). He assumes that there is perfect competition among risk-neutral bookmakers, no transactions costs, and that insiders can identify the outcome with certainty, while the rest of the betting population, i.e. outsiders, are simply noise traders. Assuming that the incidence of insider trading is not larger when a favourite is tipped to win than when a longshot is tipped to win (or more precisely, so long as the ratio of insider trading to the winning probability falls as the probability of winning rises), it is shown that equilibrium prices will exhibit a favourite-longshot bias. The intuition is clear. If insiders (who make up a proportion z of all bettors) know with certainty the outcome of a race, than bookmakers face a greater loss from insiders at higher odds (lower implied winning probabilities) than at lower odds. So long as it is not the case that the proportion of insiders (the value of z) is relatively higher at lower odds, bookmakers will face greater expected losses to insiders at higher odds. A simplifying assumption is that where insiders play no part market prices correspond to the true probabilities, and so any deviation from this can be attributed to insider trading. The consequence of price-setting behaviour by bookmakers facing this uncertain environment is for the normalised betting odds to understate the winning chances of favourites and to overstate the winning chances of longshots. This is the traditional favourite-longshot bias. Shin (1993) provided an estimate of the extent of insider trading based on a proposed link between the size of the bid-ask spread in the market and the prevalence of insider trading in the market. The key to the analysis lies in the idea that the direct effect of insider trading on the sum of bookmakers' prices will tend to increase as the number of runners (and therefore the size of the odds) increases, and that this regularity can be used to isolate the proportion of insiders in the total betting population. Shin estimated this proportion by applying an iterative estimation procedure on linearized versions of his equations, although Jullien and Salanie', show how this can be achieved using standard nonlinear estimation procedures. While Shin's modelling can explain a favourite-longshot bias in betting markets characterized by odds-setters, and also a link between the number of runners and the bookmakers' over-round, it can be shown (see Vaughan Williams and Paton, 1997a, p.152) that identical results may result from demand-side explanations, such as the behaviour of bettors who discount a fixed fraction of their losses, a suggestion proposed by Henery (1985). Because of this, Vaughan Williams and Paton (1997a) employ a much larger data set than that in Shin (1993), and distinguish between two types of race, on the basis of their relative potential for insider trading. It is shown that in those races in which there might be expected to be very limited opportunity for the non-disclosure of inside information, the link between the number of runners and the over-round disappears. This lends empirical support to Shin's supply-side explanation of the phenomenon.

Thus an estimate (of the extent of inside information) can be derived, from the form of relationship between the sum of bookmakers' prices and number of runners. Shin (1993), Jullien and Salanie' (1994)[9], and Vaughan Williams and Paton (1997d)[9] , all calculate a figure of about 2 per cent of all sums staked. Fingleton and Waldron (1996), however, who model the issue on the assumption that bookmakers are risk-averse, engage in anti-competitive behaviour and/or face significant transactions costs, estimate a figure for their Irish data of 3.1 to 3.7 per cent. Future research might usefully build upon the sort of models adopted by Shin (1993) and Fingleton and Waldron (1996), in particular to amend some of the more heroic assumptions of those models, and also to develop alternative empirical tests which can arbitrate between demand-side and supply-side explanations of the data. It is unlikely that such an approach would affect the qualitative findings, although they may well affect the estimate of the extent of insider trading.

A more direct test of the extent of insider trading is provided by Terrell and Farmer (1996). They propose a model of a betting market characterized by informed bettors (who purchase the true probabilities) and uninformed bettors (who do not). The presence of these uninformed bettors is sufficient to cause the market odds to diverge from the objective probabilities. Any difference between the return to a random betting strategy based on market odds and the actual return to bettors (the pool minus deductions) is thus interpreted as income to informed bettors. Since the expected return to random bets was calculated to be 78.3 per cent of total stakes, compared to a pool payout of 82 per cent, the difference (3.7 per cent) is classified as the reward to being informed.

Extrapolating across all races at the Woodlands track in the 1989 season yielded an estimate of $2.23 million as this reward.

4.7 Semi-strong and strong form information efficiency in betting markets: Summary and conclusions

In testing for the existence of semi-strong form efficiency in financial markets, two main approaches have been adopted. One is to assess the impact of new public information on prices. The other is an exploration of opportunities for earning systematic abnormal returns on the basis of identifiable circumstances (so-called 'market anomalies').

Semi-strong form tests of betting market efficiency employ the same methodology, applied to the special circumstances pertaining to betting markets. In a financial market which is semi-strong form efficient it should not be possible to earn abnormal returns through a strategy of predicting future prices on the basis of information which is currently publicly available. Indeed, any such strategy should on average yield the same return, unless there are differential costs or risks associated with these strategies. Likewise, the existence of semi-strong form efficiency in betting markets would imply that the expected returns to any bet, or type of bet, placed about identical outcomes on the basis of publicly available information, should be identical (subject to identical costs and risks). In a semi-strong efficient market, for example, the expected return to a bet placed on a horse on the parimutuel (or tote) should be identical to that available with bookmakers, should both options be available, as should the actual return to different bets which have the same expected return. Similarly, it should not be possible to identify patterns on the basis of information publicly available, in the returns which can be used to yield above-average or abnormal returns. For example, the expected returns to bets on a Monday should be the same as on a Thursday.

A number of studies have sought to identify significant differences in the actual return to different types of bet available in parimutuel markets. To do this 'exotic bets' (bets involving two or more horses) are used, since they allow a comparison of the actual return to bets characterized by equivalent probabilities of success. These studies offer mixed support for a hypothesis of market efficiency.

In a semi-strong form market, the expected return to identical bets placed in different arenas (where both are equally available) should also converge, or else there are opportunities for arbitrage. A study by Gabriel and Marsden (1990) purports to show just such a divergence (between tote odds and bookmaker odds), but the significance of their findings is reduced somewhat when an error in their method was corrected (Gabriel and Marsden, 1991). Further evidence of significant differences in starting price and tote returns is provided by Blackburn and Peirson (1995) and Vaughan Williams and Paton (1997c).

As part of the more general issue of market 'anomalies', some studies of betting markets have tested for the existence of a 'gambler's fallacy', i.e. a tendency for bettors to believe that the probability of an event is decreased when the event has occurred recently, e.g. a number drawn in a lottery or a sequence of previous winning favourites. This sort of bias has been observed in lottery play and in studies of the U.S. racetrack.

A quite different approach to testing for semi-strong form efficiency has been the application of 'fundamental' strategies to racetrack data, i.e. assessing the value of decision rules based on publicly available information. These range from simple policies such as always betting the inside track, to complex handicapping models. While there is evidence that such fundamental strategies are capable to some degree of forecasting racetrack outcomes, there is less evidence is available that such in improvement can be used to make abnormal returns. Where such evidence does exist the strength of the findings is either linked to an ability to operationalize the model incorporating in limited real time, or else it is subject to variations in the return which may permit extended short-term losses.

It is ambiguous whether information contained in forecasting (tipping) services should be classed as private information. In assessing financial market efficiency, forecasting performance was classified as evidence of strong form inefficiency, albeit loosely defined. Information implicit in the prices forecast about horses in daily publications is, however, clearly publicly available. Studies which have sought to generate rules based on these prices have usually found that any information is impounded into the odds by the off. Even where there is some indication that the use of this early price information could generate above-average returns, there is very little evidence of any opportunity to earn to earn abnormal profits. There has been very little academic work on the performance of private professional forecasting services (tipsters), although a study by Crafts (1994) of racing systems and ratings services revealed little evidence of inefficiency.

Studies which investigate the existence of strong form efficiency in racetrack betting markets address the issue of the value of inside information in these markets. Three distinct methodologies have been adopted in the literature. One such approach is to compare prices (odds) at different stages in the development of the market (Dowie, 1976; Crafts, 1985). Forecast prices are sometimes used as a convenient representation of pre-market or earliest prices. The idea behind such studies is that in a fixed-odds market devoid of inside information, there should be no opportunities to earn a higher expected return at earlier prices (before insiders can take the price) than at later prices. If such opportunities can be identified, this is prima facie evidence of strong form inefficiency in the market.

There has been some evidence produced (Crafts, 1985) that horses whose odds shorten in the market demonstrate a higher expected return (especially at earlier or forecast prices) than those whose odds do not, and vice-versa. Moreover, this effect seems to be greater in races where insider information is likely to be more prevalent. Another approach, identified by Schnytzer and Shilony (1995), compares two mutually isolated groups, one of which it is proposed has access to inside information and one of which does not. Empirical tests of differences in the betting behaviour of these two groups is consistent, they argue, with the presence of inside information which possesses a positive value. In particular, the group with access to the inside information were better (and more consistently) able to predict the outcomes of races than the other group. The third approach (Shin, 1993; Fingleton and Waldron, 1996; Vaughan Williams and Paton, 1996, 1997a) is to derive or imply the incidence of insider trading indirectly by an analysis of the divergence of the pricing behaviour of bookmakers relative to what would be expected were there no possibility of insider activity. Such studies imply that bookmakers perceive the existence of some positive level of insider trading, estimated to be between about two and four per cent of all sums staked.

ENDNOTES

1. Tote odds, as in all parimutuel systems, are not fixed or guaranteed, but depend on the size of the pool and the number of winning ticket holders. A large pool, with few winning tickets, for example, pays higher odds than a small pool with a large number of winning tickets. The odds are displayed as a dividend to a £1 stake, and rounded downwards to the nearest 10 pence. A tote dividend of £4.50 thus implies odds of 3.5 to 1 against. Such odds are the actual return and are not subject to any further deductions, the deductions being accounted for in the dividend.

2. The same idea can be applied to the British bookmaking system by substituting 'to earn a positive profit' for 'to overcome the track take'.

3. 'System 73' by Ainslie, in Ainslie (1973); 'Singularly Best Race and Speed', by 'Cohen and Stephens, and 'Singularly Easy Win' by Cohen and Stephens, both in Cohen and Stephen (1963); 'An Elimination Rule' by McQuaid; 'The Consistent Horse System' by McQuaid, in McQuaid (1971); and 'A Breaks and Trial System' by Reynolds, in Reynolds (1971).

4. See Benter (1994) for details of the database.

5. The log transformation of the public's win probabilities is used instead of the actual win probabilities because of evidence that it improves the statistical fit. See Asch and Quandt (1986, pp.123-125) and Benter (1994).

6. The "experts" chosen were from the *Daily Racing Form* publication, three Chicago newspapers, and the track handicapper.

7. Where a section of the population have monopolistic access to information not available to others, Dowie prefers to describe this as inequitable rather than inefficient; "... we will talk of a market as efficient to the extent that it passes the weak and semi-strong tests and equitable to the extent that it passes the strong test." (Dowie, 1976, p. 140).

8. In cases where starting prices exceeds forecast prices, Crafts includes only those where the sum of the forecast price probabilities were greater than the sum of the starting price probabilities by less than 0.01 per horse after disregarding the residual (or 'bar') category.

9. Using Shin's data set.

10. Using their own data set but Shin's estimation procedure.

CHAPTER FIVE

CONSTRUCTING A NEW DATA SET TO TEST FOR INFORMATION EFFICIENCY IN BRITISH RACETRACK BETTING MARKETS

5.1 Introduction

Most data sets derived from racetrack betting markets are limited to a very small number of variables. The most basic include only the finishing position of the horse (or, at least, whether the horse finished first or not), and the starting price. This is sufficient to calculate the expected return at general prices prevailing at the off. As a consequence, simple tests for the existence of a favourite-longshot bias compare the expected returns at different starting prices, e.g. Henery (1985), or at different parimutuel odds, e.g. Snyder (1978a, 1978b). If the expected returns differ significantly between odds levels (or odds categories) this is a bias. If the expected return tends to be greater at lower odds, this is the traditional favourite-longshot bias. Some studies extend the analysis by comparing the expected return to races classified by their position in the race order, e.g. McGlothlin (1956), Ali (1977), Asch, Malkiel and Quandt (1982), Metzger (1985). There is a suggestion in these studies that the size of a longshot bias may be different in later races than it is in earlier races. Johnson and Bruce (1993) adapt this approach to consider differences in the behaviour of off-course bettors at different times of the day. Asch, Malkiel and Quandt (1982) also compare the expected return to horses categorized by their order of favouritism. Henery (1985) suggests that the time of the season from which the data is drawn may affect the findings. Other studies, such as Swidler and Shaw

(1995), have analyzed data which may show evidence of mispricing across the win, place and show pools, with a potential for profitable arbitrage. Tests of semi-strong efficiency have highlighed tote/starting price differentials, e.g. Gabriel and Marsden (1990, 1991), Bird and McCrae (1994), Blackburn and Peirson (1995), Vaughan Williams and Paton (1997c), and the possibility of a gambler's fallacy resulting from clear sequences of winning favourites (or longshots - see Metzger (1985). Tests along the lines of Gabriel and Marsden (1990, 1991) require evidence on tote returns and the corresponding starting price, sub-divided perhaps into various parts of the season to examine the influence of learning effects. Tests, like that of Metzger, which focus upon the gambler's fallacy, require information on patterns of winning favourites and longshots. Other information has been gathered and analyzed about the actual return to different bets possessing identical probabilities (e.g. Ali, 1979; Asch and Quandt, 1987; Lo and Busche, 1994). Forecasting models have ranged from the use of specific predictors, such as post position, as in Beyer (1983), to sophisticated multinomial logit models, as employed by Bolton and Chapman (1986). The number of potential variables which can be employed here is almost limitless. A final category of tests (which in some circumstances could arguably be seen as tests for strong efficiency) seek to establish whether prices forecast by professional judges of form can be employed, alone or in conjunction with other information, to make an above-average return. Studies by Losey and Talbott (1980) and Bird and McCrae (1987) typify the field. This category of tests require information on forecast prices and final expected returns at obtainable odds. There are no large academic studies of the performance of professional tipping services, although Crafts (1994) has investigated this issue.

Tests for strong form efficiency are attempts to identify the presence of valuable inside information in a betting market. The idea behind some of these is that starting prices should incorporate more information than earlier prices, and especially forecast prices, which are 'set' even before the market opens. Differences between the expected return at forecast prices and at starting prices, in the whole sample or in identifiable parts of the sample (such as when prices shorten markedly during a market), may provide evidence of insider activity. To perform these tests information has been collected about forecast and starting prices (after Dowie, 1976; Crafts, 1985), and also about the longest price available at any time in the on-course market (Crafts, 1985). Shin (1993), however, has sought to identify the incidence of insider trading through an examination of the link between the sum of bookmakers' prices in a race and the number of runners. Tests of this model require information about the sum of prices implied in the odds about each race, together with an identification of the number of runners competing in that race. Schnytzer and Shilony (1995) compare the predictive ability of two groups of mutually isolated bettors, one of which consists of uninformed bettors only and one of both uninformed and informed (insider) bettors. The hypothesis is that if the latter group can predict better (and more consistently) than the former this constitutes evidence of the value of private information. Their data is unique to the Australian bookmaking/tote on/off course set-up, and involves a comparison of the sums laid on the tote about particular horses on and off course.

This data set covers all the variables considered above insofar as they can be derived from British racetrack betting markets, and many more besides. Some of these, such as the prize money offered about a race can cross-check particular comments or claims made by authors of previous articles - for example, Henery's (1985) suggestion that odds are a more reliable guide to objective probabilities in races characterized by higher prize money. In other cases, it is necessary for the implementation of new tests proposed here. In some cases, the scope of the information extends beyond the immediate needs of this research programme, but may be used profitably as the basis for further work in this field.

The information contained in this data set is divided into four sections. The first three sections correspond approximately to the first, the second, and the final third of the 1992 Flat racing season. The fourth section consists of data collected for those races in the 1993 Flat season upon which the major bookmakers offered odds prior to the opening of the course market.[1] The information was gathered from Raceform Flat Annual (1992), the Sporting Life Flat Results (1992), the Sporting Life Flat Results (1993), editions of the Sporting Life newpaper covering the 1992 and 1993 Flat Racing seasons, and televised racing coverage on BBC1, BBC2 and Channel 4. All races in the specified periods on standard racedays are included, i.e. days on which six or seven races took place; and in which no horse was withdrawn too late for a fresh book to be formed. The data set excludes all-weather racing in order to maintain consistency with earlier studies of the U.K. racetrack.

The variables included in the new data set are set out below, with additional, more detailed commentary given in Appendix 1.[2]

5.2 List of the variables included in the new data set

1. The finishing position of the horse in each race.
2. A race number for each race.
3. A number allocated to each race to indicate the order of that race in the daily running order of all races at the racecourse in question.
4. The starting price of each horse in every race.

5. The opening price of each horse in every race. This indicates the price about each horse at the first show in the market.
6. The best price available about each horse in every race.
7. The worst price available for each horse in every race.
8. The forecast price of each horse in every race.
9. The prize money associated with each race.
10. The distance of each race.
11. The ages of the horses competing in each particular race.
12. The number of runners in the race.
13. The type of race.
14. The month of the year in which the race took place.
15. The day of the week on which the race took place, e.g. Tuesday.
16. The date of the month in which the race took place, e.g. 21.
17. The official "going", or state of the ground, for the race.
18. The forecast going, as listed in the Sporting Life, on the morning of the race.
19. The racecourse.
20. The over-round.
21. The grade of handicap, where applicable.
22. The "going correction", as listed in the Sporting Life Flat Results (1992, 1993).
23. The tote odds returned to a bet placed on the winning horse.
24. The order of the race on the racecard.
25. The finishing position of the favourite.
26. The finishing position of the favourite at the opening show, as indicated by the shortest opening price in the market.
27. The early odds offered about each horse running on a particular day (where applicable) by the William Hill bookmaking organisation, as reported in the Sporting Life on the day of the race.
28. The early odds offered about each horse running on a particular day (where applicable) by the Corals bookmaking organisation, as reported in the Sporting Life on the day of the race.
29. The early odds offered about each horse running on a particular day (where

applicable) by the Ladbrokes bookmaking organisation, as reported in the Sporting Life on the day of the race.

30. The starting price of the favourite in each race.

31. The starting price of the second favourite in each race.

32. The finishing position of the forecast favourite, as indicated by the shortest forecast price offered in the Sporting Life on the day of the race.

33. The finishing position of the bookmakers' favourite, as indicated by the early odds offered about each horse in a race by the three major bookmaking organisations, i.e. William Hills, Corals and Ladbrokes.

34. The Bar price.

35. The position of the race winner in order of favouritism.

36. The proportion of the returns to a winning bet deducted by bookmakers when a horse or horses do not run, but where bets have been placed in advance of the start on the basis that they are runners.

37. An alternative measure of the returns to a winning bet which should be deducted by bookmakers when a horse or horses does not run, but where bets have been placed in advance of the start on the basis that they are runners.

38. The Dual Forecast odds.

39. The Computer Straight Forecast odds.

40. The place odds returned by the tote about the winning horse.

41. The place odds returned by the tote about the horse
finishing second, where applicable.

42. The place odds returned by the tote about the horse finishing third, where applicable.

43. The place odds returned by the tote about the horse finishing fourth, where applicable.

44. The Tricast odds returned by the tote.

45. The number of horses contained in the Sporting Life 'Bar' about each particular race, i.e. the number of horses about which no individual odds are forecast in that race.

46. The draw of the horse.

47. The jockey.

48. The trainer.

49. The age of the horse.

50. The weight carried by the jockey.

51. The number of pounds the jockey is allowed to deduct from the weight allocated to the horse, i.e. the jockey's claim.

52. The number of pounds a horse is carrying in excess of the weight allocated to it, where applicable.

53. A measure of the movement in the betting odds, as reported in The Sporting Life Flat Results (1992).

54. A measure of the movement in the televised betting odds.

55. The opening price quoted about the opening favourite, i.e. the horse with the shortest opening price.

56. The opening price of the second favourite in each race.

57. The forecast price quoted about the forecast favourite,
i.e. the horse with the shortest forecast price (as quoted in the Sporting Life).

58. The forecast price of the forecast second favourite in each race.

ENDNOTES

1. Period 1 covers races from Thursday, March 19, 1992 to Saturday May 16, 1992 inclusive; period 2, from Monday, May 18, 1992 to Saturday, August 3, 1992; period 3, from Monday, August 3, 1992 to Monday, November 3, 1992 (inclusive). Period 4 consists of data collected for selected races from the 1993 flat season, i.e. those in which at least one of the three major U.K. bookmakers (Ladbrokes, William Hill and Corals) offered odds prior to the opening of the course market, and covers the period from Thursday, March 25, 1993 to Monday, November 8, 1993 inclusive.

2. The following variables are presented for the whole of the data set :

1, 3, 4, 9, 13, 24.

The following variables are presented for the whole of periods 1 and 4 (except where indicated in the data set):

1 to 58 (inclusive) excluding 46, 47, 48, 50, 51, 52.

CHAPTER SIX

EMPLOYING A NEW DATA SET TO TEST FOR WEAK FORM EFFICIENCY IN BRITISH RACETRACK BETTING MARKETS

6.1 Introduction

In this chapter, the new data set is used to test the hypothesis that there exists evidence of weak information inefficiency in the form of a favourite-longshot bias in recent British racetrack betting markets. Tests are conducted at two levels: first the whole 1992 flat racing season and second, for the 1992 season sub-divided into approximate chronological thirds. Previous studies have differed in the odds classifications used and also whether to examine differences in the bias based on the position of a race in the race order (and if so, on what basis should the races be divided up). Three previous studies are chosen as representing the broad theme of the literature, and most clearly and directly addressing the most relevant issues, and the results of the new empirical tests are presented here.

The first is adapted from Dowie's (1976) presentation (for the 1973 flat racing season) of the ratio of winnings to stakes for ranges of starting prices considered cumulatively, in the summary form employed in Crafts (1985) and Shin (1991).

The second is Asch, Malkiel and Quandt's (1982) tests of weak form efficiency in U.S. parimutuel betting markets. In their study, they calculated the rates of return from bets on horses in various odds categories for all races, and compared the findings with the rates of return from bets on horses in these same odds categories for the last two races of the day (in their cases, races 8 and 9).

The third is Henery's (1985) explanation of a favourite-longshot bias in British racetrack betting markets. Henery's investigation of British racetrack betting patterns, using data from the 1973, 1978 and 1979 flat racing seasons, shows the ratio of winnings to stakes for various betting odds, and also proposes an explanation based on a systematic divergence between subjective and objective probabilities.

"The suggested model says that, in assessing the chances of a horse losing the race, the punter only counts a fixed fraction, f, of his losses" (Henery, 1985, p.344).

Henery regressed the subjective lose probability (implied by the odds) against the objective lose probability (implied by the proportion of losers), and found that the former was significantly less than the latter. This is re-examined here with respect to the new data set. The data is also re-analyzed at different levels of sub-aggregation, based on prize money and time of the season, in order to test suggestions in Henery that there is a closer fit between subjective and objective probabilities in races later in the season, and for races with higher prize money.

"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 June 6 onwards) and the merits of the horses would be better known, and also because only those races were chosen where the prize money was more than £1000 (thereby excluding some selling plate races). Odds quoted in these races may also be more reliable than those in the early part of the season when the form of the runners is not known." (Henery, p.345).

In undertaking these tests, particular allowance is made for potential problems arising from heteroscedasticity in the error terms.

6.2 An examination, employing the new data set, of cumulative returns to level stakes - based on an approach in Dowie (1976).

In this section, Dowie's (1976) idea of calculating the cumulative return to level stakes bets up to particular starting prices is adopted and applied to the new data set constructed for this study. This is presented in the summary form suggested in Crafts (1985) and Shin (1991).

Dowie's investigation (1976) of the 1973 British flat reason shows the ratio of winnings to stakes for various ranges of betting odds. These are displayed in Crafts (1985) and in Shin (1991) as follows:-

up to evens;

up to 5/1;

up to 10/1;

up to 16/1;

All.

The findings here, however, are presented not only in aggregate form, but also sub-aggregated by the position of the race on the racecard. There are four levels of sub-aggregation. The 'first', 'penultimate' and 'final' races refer to the order of the race on the racecard at respective racetracks on a particular day. The set of 'middle' races is defined for these purposes as those races which are neither 'first, nor 'penultimate' nor 'final', in the sense described above.

The results are presented in tables 6.1 to 6.5.

Table 6.1 **Cumulative returns to level stakes for all races in periods 1, 2, 3 and 4**

Starting Price	≤1	≤5	≤10	≤16	**All**
Field size	288	3652	8427	12620	18653
Winners	169	943	1401	1583	1675
Actual return	283.36	3422.29	7404.29	10048.29	12546.29
Expected return (%)	98.39	93.71	87.86	79.62	67.26

Table 6.2 **Cumulative returns to level stakes - for all "first races of the day " in periods 1, 2, 3 and 4**

Starting Price	≤1	≤5	≤10	≤16	**All**
Field size	55	506	1086	1649	2609
Winners	36	136	188	214	235
Actual return	60.57	449.91	918.41	1290.41	1937.41
Expected return (%)	110.13	88.92	84.57	78.25	74.26

Table 6.3 **Cumulative returns to level stakes - for all "middle races of the day" in periods 1, 2, 3 and 4**

Starting Price	≤1	≤5	≤10	≤16	**All**
Field size	138	2045	4891	7364	10788
Winners	63	514	800	906	955
Actual return	139.64	1946.09	4400.59	5959.59	7260.59
Expected return (%)	101.19	95.16	89.97	80.93	67.30

Table 6.4 **Cumulative returns to level stakes - for all "penultimate races of the day" in periods 1, 2, 3 and 4**

Starting Price	**≤1**	**≤5**	**≤10**	**≤16**	**All**
Field size	63	582	1222	1745	2579
Winners	32	163	222	242	253
Actual return	54.11	549.42	1063.42	1349.42	1610.42
Expected return (%)	85.89	94.40	87.02	77.33	62.44

Table 6.5 **Cumulative returns to level stakes - for all "final races of the day" in periods 1, 2, 3 and 4**

Starting Price	**≤1**	**≤5**	**≤10**	**≤16**	**All**
Field size	32	519	1228	1862	2677
Winners	17	130	191	221	232
Actual return	29.06	476.87	1021.87	1448.87	1737.87
Expected return (%)	90.81	91.88	83.21	77.81	64.92

6.2.1 Summary of results

For the whole sample, there is a monotonic inverse relationship between the length of the odds in each odds grouping and the size of the expected return (a favourite-longshot bias). This pattern was repeated for all first, 'middle', penultimate and final races. For penultimate races, the expected return was slightly lower in the shortest odds category than in the two categories above (i.e. of longer odds), and for final races it was slightly lower for the shortest odds category than for the immediate category above. Since the sample size is smallest for the 'aberrant' odds group, i.e. up to evens, and allowing for the cumulative nature of the classification system, the significance of the findings have to be taken with care. However, the direction of the findings are not consistent with earlier studies of a late race phenomenon, such as Ali (1977) and Asch, Malkiel and Quandt (1982), both of which suggest a greater bias for later races, and McGlothlin (1956), for final races. Although McGlothlin did find a contrary effect in penultimate races, he explained this in terms of the special nature of such races, as feature races of the day. The issue is examined in more detail in the next sub-section.

6.3 An examination, employing the new data set, of rates of return in various odds categories - based on an approach in Asch, Malkiel and Quandt (1982).

Asch, Malkiel and Quandt (1982) calculated the rates of return from bets on horses in various odds categories for all races, and compared these findings with the rates of return from bets on horses in these same odds categories for late races only (specifically the last two races of the day, in their cases races 8 and 9). They classified the odds groupings into the following intervals:

a) less than or equal to 2;

b) greater than 2 but less than or equal to 3.5;

c) greater than 3.5 but less than or equal to 5;

d) greater than 5 but less than or equal to 8;

e) greater than 8 but less than or equal to 14;

f) greater than 14 but less than or equal to 25;

g) greater than 25.

The returns are based on a parimutuel system, in which the deductions from the pool are estimated by Asch et al. as roughly 18.5 per cent. Thus a rate of return of greater than -0.185 reflects a profitable return before deductions.

These data are divided into three sections, corresponding to approximately the first, the second, and the final third of the 1992 flat racing season. Calculations are made of the average return for all the races at each of the odds groupings chosen by Asch et al., as well as of the average return for the subset consisting of the penultimate and the last races of each day taken together.

Three distinct changes have been made to the Asch et al. approach, without changing its sense or purpose. First, since the data set employed in the present study applies to a bookmaker system, there is no direct comparison with their results. In their case, the actual returns should be viewed in the context of an 18.5 per cent deduction from the payout. Second, the findings are presented for successive parts of the 1992 flat racing season (approximately thirds), taken in chronological order. This helps detect, for example, any trends or time-specificity in the data, or learning effects. This follows suggestions in Henery (1985) and Gabriel and Marsden (1990). The findings are also presented for a separate later sample composed of races from the 1993 flat racing season, and selected on the basis that early odds are offered by the major bookmakers on these races (see above). These races tend to contain a disproportionate number of large handicaps, and prestige events. Third, all racecards in the present sample consist of six or seven races, so the equivalent to their eighth and ninth races are in this case the fifth and sixth races (where there are six races on the card), or the sixth and seventh races (where there are seven races on the card).

In each case the odds are grouped into the following categories-

SP(starting price)≤ 2; $2<SP\leq 3.5$; $3.5<SP\leq 5$; $5<SP\leq 8$; $8<SP\leq 14$; $14<SP\leq 25$; $SP>25$.

Results are presented for the aggregates and for various sub-categories with respect to the following:

total number of horses in each odds grouping ("field size");

total number of horses winning in each odds grouping ("winners");

total return to a unit stake on all the horses in each odds grouping ("actual return");

expected return to a unit stake on all the horses in each odds grouping, i.e. total return divided by total stake multiplied by a hundred ("expected return").

The results are presented in tables 6.6 and 6.7.

Table 6.6 presents the results for all the races in the first period analyzed, followed below by the results for all the races in the second period, followed below by the results for all the races in the third period, followed by the results for all the races in the fourth period.

Table 6.7 presents the results for the subset consisting of the penultimate and the last races of each day taken together. This is presented as in table 6.6, first for period 1, followed below by period 2, followed below by period 3, followed below by period 4.

Table 6.6 Expected returns to level stakes, in various odds categories - for all races, separately for period 1, period 2, period 3 and price 4

Starting price	**≤2**	**> 2 ≤3.5**	**> 3.5 ≤5**	**>5 ≤8**	**>8 ≤14**	**> 14 ≤25**	**>25**	
Field size	281	420	431	891	1449	1332	1044	Period 1
Winners	122	98	64	87	87	41	7	
Total Return	276.65	380.29	357	675.5	1048	855	303	
Expected return (%)	98.45	90.55	82.83	75.81	72.33	64.19	29.02	
Field size	382	500	537	917	1066	842	852	Period 2
Winners	174	109	96	98	57	27	4	
Total Return	387.55	426.91	532	772.5	668	526	283	
Expected Return (%)	101.45	85.38	99.07	84.24	62.66	62.47	23.83	
Field size	216	291	328	749	1288	1164	961	Period 3
Winners	96	66	58	80	74	33	7	
Total Return	208.89	255.41	322	624	889	672	271	
Expected Return (%)	96.71	87.77	98.17	83.31	69.02	57.73	28.2	
Field size	23	89	154	461	819	757	409	Period 4
Winners	7	24	29	50	57	20	3	
Total Return	14.57	97.5	163.5	395	704	421	97	
Expected Return (%)	63.35	109.55	106.17	85.68	85.96	55.61	23.72	

Table 6.7 **Expected returns to level stakes, in various odds categories - for all "penultimate and final races of the day", separately for period 1, period 2, period 3 and period 4**

Starting price	**≤2**	**> 2 ≤3.5**	**>3.5 ≤5**	**>5 ≤8**	**>8 ≤14**	**> 14 ≤25**	**>25**	
Field size	99	130	123	244	431	410	343	Period 1
Winners	46	33	11	21	30	13	0	
Total Return	106	126.87	63	166	361	295	0	
Expected return (%)	107.07	97.59	51.22	68.03	83.76	71.95	0	
Field size	118	136	189	304	362	240	254	Period 2
Winners	59	30	32	25	19	7	2	
Total Return	139.69	120.33	175	201	220	132	85	
Expected Return (%)	118.38	88.48	92.59	66.12	60.77	55	33.46	
Field size	78	87	98	199	368	320	269	Period 3
Winners	32	24	15	25	19	9	0	
Total Return	69.89	93.67	82.5	192.5	228	183	0	
Expected Return (%)	89.6	107.67	84.18	96.73	61.96	57.19	0	
Field size	2	11	30	81	154	123	409	Period 4
Winners	1	4	6	7	13	2	3	
Total Return	1.615	11.25	32.5	56.5	147	42	97	
Expected Return (%)	80.75	102.27	108.33	69.75	95.45	34.15	23.72	

In every single section of the analysis, the expected return in the shortest odds grouping (≤2) was greater than that in the longest odds grouping (>25). Moreover, with the exception of the part of Table 6.6 dealing with period 4 (a small sample), the expected return for the set of all races in each of the periods studied was greater in all odds classifications up to and including 8 than in all odds classifications at odds above 8. For the subset of races comprising the penultimate and last races, however, the expected return in the odds classification 8 <SP≤14 was higher in two of the periods examined (the first and the fourth) than in at least one lower odds classification.

In order to clarify the issue of significance, Tables 6.8 and 6.9 present calculations, for the whole data sample, of the expected return for later as compared with all races.

Table 6.8 A comparison of the expected returns to level stakes, at starting prices ≤ 8 with starting prices > 8, for all races in periods 1, 2, 3 and 4

Starting Price	≤8	> 8
Field size	6670	11983
Winners	1258	417
Actual return	5889.29	6657
Expected	88.30	55.55

Table 6.9 **A comparison of the expected returns to level stakes, at starting prices ≤ 8 with starting prices > 8, for all "penultimate and final races of the day" in periods 1, 2, 3 and 4**

Starting Price	≤**8**	> **8**
Field size	1929	3327
Winners	371	114
Actual return	1638.32	1693
Expected	84.93	50.89

There is very little evidence, with the odds grouped more generally, of a clear difference in the expected returns between the races aggregated into penultimate/final races compared with those of other races.

To clarify the issue, however, significance tests are performed on the set of all penultimate/final races (table 6.10) and the set of all other races (table 6.11), employing the odds classifications in tables 6.8 and 6.9, and restricting the sample to periods 1, 2 and 3 in order to eliminate any possibility of a distortion caused by the specific nature of the races covered in period 4. Although the distributions may not be normal, the use of standard t-tests to gauge significance may be justified by appeal to the Central Limit Theorem.

Table 6.10 Testing the significance of differences in the expected return to level stakes, at starting prices≤8, for all "penultimate and final races of the day", and all "other races of the day", in periods 1,2,3 (combined).

All penultimate and final races (periods 1,2,3).
For SP ≤8
Field size = 1805; winners = 353; actual return = 1536.42;
expected return = 0.8512 (85.12%)
Standard deviation (actual returns) = 1.976

All other races (periods 1,2,3).
For SP ≤8
Field size = 4138; winners = 795; actual return = 3682.29;
expected return = 0.8899 (88.99%)
Standard deviation (actual returns) = 2.073

For the two samples, t = 0.684.

Table 6.11 Testing the significance of differences in the expected return to level stakes, at starting prices>8, for all "penultimate and final races of the day" and all "other races of the day", in periods 1,2,3 (combined).

All penultimate and final races (periods 1,2,3).
For SP>8
Field size = 2997; winners = 99; actual return = 1504;
expected return = 0.5018 (50.18%)
Standard deviation (actual returns) = 2.949

All other races (periods 1,2,3).
For SP>8.
Field size = 7001; winners = 238; actual return = 3931;
expected return = 0.5615 (56.15%)
Standard deviation (actual returns) = 3.411

For the two samples, t = 0.8837.

The hypothesis that there is no difference in the expected return to bets placed earlier compared to those placed later in the day cannot be rejected at conventional levels of significance.

6.3.1 Summary of results

The existence of a favourite-longshot bias is confirmed for all races and for the subset of races combining penultimate and final races. There is some prima facie evidence of a lower, or even contrary bias for the last two races of the day, in identified periods, using Asch, Malkiel and Quandt's (1982) chosen odds categorizations. This is not confirmed, however, in an examination of the evidence undertaken at wider levels of aggregation.

6.4 An examination, employing the new data set, of the relationship between subjective and objective probabilities - using the approach in Henery (1985).

Henery (1985) regressed bettors' subjective estimated probabilities of losing (Q) against the objective probabilities of losing (q). He derived bettors' "subjective lose probabilities" from the odds, and the "objective lose probabilities" from the ratio of losers to total races. In particular, Q = SP/(1+SP), so where SP (the starting price) is 3 to 1, say, Q = 1/4 (or 0.25). If 100 horses actually lose at an SP of 3/1 out of a total of 250 starting at this price, q = 100/250, i.e. 0.4. His results, using two different data samples (from the 1973 and 1978-79 flat seasons), indicate a close correspondence to a linear relationship, of the form Q = fq, where f is significantly less than 1.

In this sub-section, the subjective lose probability, Q, is plotted against the objective lose probability, q, both as defined by Henery, and a weighted least squares fit applied. The results are repeated at two levels of sub-aggregation, the first to compare the results at different stages of the season, and the second to compare the results in races categorized by levels of prize money. This follows suggestions in Henery (p.345) that the closeness of fit of these variables to a straight line may be positively related to lateness in the season and for races offering higher prize money. Figgis' (1976) investigation of the 1973 flat racing season, as reported in Henery, calculated the average return to unit stakes for individual odds classifications, as opposed to classifying into odds ranges. The same approach is applied to the new data set derived for this study, but the number of odds classifications is increased from the 19[1] reported in Henery, to all 71 identified in the present data set.[2]

In this form 0.8, for example, represents odds of 0.8 to 1 against (or 4/5 or 5/4 on), 33 represents odds of 33 to 1 against.

There are five categories:

1. SP - the starting price at which each horse was returned, presented in decimalized form. For example, a starting price of 11 to 4 is recorded as 2.75.

2. Win ratio - the proportion of horses returned at each starting price which won. For example, 12 out of 17 signifies that 17 horses were returned at that starting price, and twelve won.

3. Rate of return (%) - The rate of return is the total return on all winning horses in a given starting price category (inclusive of a level unit stake) divided by the total of all level unit stakes on all horses running in that category. The formula is Winners X (SP + 1) divided by number of runners. For example, if 50 horses ran at an SP of 2 to 1, and 18 won, the rate of return = 18 X (2 + 1) divided by 50 = 54/50 = 108%. If only 9 won, the rate of return would be 27/50 = 54%.

4. Subjective lose probability (Q) - The probability that a horse will lose, as implied in the starting price. This is calculated by the formula $Q = SP/1+SP$. For example, a starting price (SP) of evens (1 to 1) implies a subjective lose probability of 1/2 (or 0.5).

5. Objective lose probability (q) - The fraction of horses which actually lost at the given odds. This is 1 minus the win ratio. For example, if 12 out of 20 horses won at evens, the objective lose probability at evens is 1-(12/20) = 8/20 (0.4).

Four tables of results are presented for the 1992 flat racing season in Appendix 2. In the first table (table 2A), results are presented for all races in periods 1,2 and 3 combined. In the second table (table 2B), results are presented for period 1 separately, and in this table (table 2C) for period 3 separately. In the final table (table 2D), results are presented for all races in periods 1,2 and 3 (combined) for which prize money exceeds the median prize money in each individual period.

Henery plotted Q (the subjective lose probability) against q (the objective lose probability), using the model Q = fq, where f, some fraction between 0 and 1, is defined as a fixed fraction of losses discounted by the bettor. Applying weighted least squares fits to his data (taken from the latter part of the 1978/1979 flat racing seasons, and excluding some lower quality races), and to data provided by Figgis for the 1973 flat season, he found "a reasonable fit on average" (p.345), although somewhat closer to a straight line in the later data. For the 1973 data, Q = 0.974q, for the 1978/79 data, Q = 0.978q.[3]

This exercise is repeated on the data set constructed for this study, for the whole of the 1992 flat racing season (table 6.12), and also at levels of sub-aggregation based on the time of season (table 6.13) and quality of race (indicated by prize money levels) - table 6.14.

Table 6.12 Estimating Henery's model, Q = fq, for all races in periods 1,2,3 (combined).

Returns from bets at given SP

(For all races in periods 1,2 and 3 inclusive).

Regressing Q against q, and quoting standard errors of the coefficients in parentheses,

$$Q = \underset{(0.0434)}{0.1365} + \underset{(0.0566)}{0.7539}q. \qquad (6.1)$$

$R^2 = 0.7230$

Employing Glejser (1969) tests[4] for the existence of heteroscedasticity suggests rejecting the null hypothesis of no heteroscedasticity at the 95 per cent level of confidence (t=2.126 for the difference between the coefficient of q and 0) in the following functional form:

$|e_i| = c_1 + cq_2 + v$

Removing just two outliers, however, restores the null hypothesis of no heteroscedasticity (t=-0.038) and yields the following regression:

$$Q = \underset{(0.0287)}{0.1158} + \underset{(0.0378)}{0.8116}q \qquad (6.2)$$

$R^2 = 0.8748$

The null hypothesis of homoscedasticity cannot be rejected before or after removing the outliers in the following functional forms.

$|e_i| = c_1 + c_2q^{0.5} + v$

$|e_i| = c_1 + c_2(1/q)+ v$

In the original form, t = 1.773 and t = -0.906 with respect to a and b. After removing the outliers, t = -0.012 and t = 0.225 respectively.

Table 6.13 Estimating Henery's model, Q = fq, for all races in period 1, and separately for all races in period 3.

Returns from bets at given SP

(For all races in period 1).

Regressing Q against q,

$$Q = \underset{(0.0384)}{0.1907} + \underset{(0.0494)}{0.7248}q \qquad (6.3)$$

$R^2 = 0.7909$

In this sample the null hypothesis of homoscedasticity cannot be rejected in any of Glejser's suggested functional forms:

$|e_i| = c_1 + c_2q + v$ (t = -0.196)

$|e_i| = c_1 + c_2q^{0.5} + v$ (t = -0.230)

$|e_i| = c_1 + c_2(1/q) + v$ (t = 0.220)

Returns from bets at given SP

(For all races in period 3).

Regressing Q against q,

$$Q = \underset{(0.0506)}{0.1938} + \underset{(0.0648)}{0.6978}q \qquad (6.4)$$

$R^2 = 0.6703$.

In this sample the null hypothesis of homoscedasticity cannot be rejected in any of Glejser's suggested functional forms:

$|e_i| = c_1 + c_2 q + v$ (t = -1.186)

$|e_i| = c_1 + c_2 q^{0.5} + v$ (t = 1.064)

$|e_i| = c_1 + c_2(1/q) + v$ (t = -0.450)

Table 6.14 Estimating Henery's model, Q = fq, for all races in periods 1,2 and 3, for which prize money exceeds the median prize money in each individual period.

Regressing Q against q,

$$Q = \underset{(0.0560)}{0.2100} + \underset{(0.0704)}{0.6709}q \qquad (6.5)$$

$R^2 = 0.6104$

In this sample the null hypothesis of homoscedasticity cannot be rejected in any of Glejser's suggested functional forms:

$|e_i| = c_1 + c_2 q + v$ (t = -1.562)

$|e_i| = c_1 + c_2 q^{0.5} + v$ (t = 1.242)

$|e_i| = c_1 + c_2(1/q) + v$ (t = -0.964)

There is no evidence of heteroscedasticity in all any of the sub-aggregated data samples, although it does pose a problem in the aggregated data set. This is removed, however, at the cost of simply removing two outliers in the data.

6.4.1 The relationship between subjective and objective probabilities: Summary of results

The data reveal a reasonable goodness of fits between Q and q in each of the samples, and the q coefficient in each of the regressions is less than one at all conventional levels of significance. This implies that a linear relationship between Q and q is a reasonable specification and that Q (the subjective lose probability) is significantly less than q (the empirical lose probability).

There is no evidence, however, for Henery's suggestion (p.345) that the closeness of fit to a straight line between these variables was greater for races later in the season or for races offering higher levels of prize money. If anything, the reverse was true.

The conclusion which may be drawn from this study is that bettors on average behave as if they attach a lower subjective probability to the likelihood of a horse losing (as implied in the starting price) than is the objective case (as revealed by the ratio of losers to runners). Henery explains this difference as a consequence of bettors discounting a fixed fraction of their losses. Another implication of this hypothesis is that it can also explain an observed relationship between the sum of prices implied in the odds in excess of 1 (the bookmakers' over-round) and the number of runners in a race. This is demonstrated below.

6.4.2 The bookmakers' over-round and the number of runners in a race: an implication of Henery's fixed loss discount model.

In a field of n horses, each of whom has an equal probability of winning, the objective probability of any particular horse losing any particular race is (n-1)/n.

In such a case, according to Henery's model, the subjective probability is:

f. (n-1)/n, where f is the fraction of losses discounted by the bettor.

The subjective odds implied by this are given by:

$$\frac{f.(n-1)n}{1 - [f.(n-1)/n]} \qquad (6.6)$$

The over-round may be derived as follows:

$$\text{Over-round} = \frac{1-f.(n-1)/n}{f.(n-1)/n + [1-f.(n-1)/n]} . n - 1 \qquad (6.7)$$

$$= n\,(1 - f.[n-1]/n) - 1 \qquad (6.8)$$

$$= n - f.(n-1) - 1 \qquad (6.9)$$

$$= (1-f).\,(n-1) \qquad (6.10)$$

The implication of this model is that the average over-round in a race should be related linearly and positively to the number of runners, and that the slope of the line representing this relationship should be (1-f).

The implication of this model is that the average over-round in a race should be related linearly and positively to the number of runners, and that the slope of the line representing this relationship should be (1-f).

The analysis employs data from period 1, i.e. races between Thursday, March 19th. and Saturday, May 18th. 1992 (table 6.15). Initially, the over-round is estimated against the number of runners for all races in the period. This is repeated for the sample of races in which the prize money exceeds the mean and again for those in which the prize money exceeds the median.

6.4.2.1 Estimation Results

Let or% = sum of prices in excess of one (the over-round), expressed as a percentage.

Let rn = number of runners.

Table 6.15 Estimating the relationship between the bookmakers' over-round and the number of runners in a race, for all races in period 1.

<u>For period 1.</u>

Regressing or against rn,

$$or\% = \underset{(0.2902)}{1.55} + \underset{(0.0202)}{2.03}\ rn \qquad (6.11)$$

$R^2 = 0.6258$

However, the Cook-Weisberg statistic (see Goldstein, 1992) suggests that there is a serious problem of heteroscedasticity. On the null hypothesis of homoscedasticity, the test statistic, distributed as σ(1) = 582.24, which is significant at the 0 per cent level. Employing standard errors which are robust to heteroscedasticity (table 6.16) do not alter the conclusions. The coefficient on the number of runners is still highly significant, and indicates an over-round of about 2 per cent for every runner.

Table 6.16 Estimating the relationship between the bookmakers' over-round and the number of runners in a race, for all races in period 1, using Huber standard errors.

<u>For period 1.</u>

Regressing or against rn,

$$\text{or\%} = \underset{(0.2466)}{1.55} + \underset{(0.0202)}{2.03}\ \text{rn} \qquad (6.12)$$

$R^2 = 0.6298$

The analysis is now repeated for all races in this period in which the prize money of the race exceeds the mean for all races. The results are presented in table 6.17.

Table 6.17 Estimating the relationship between the bookmakers' over-round and the number of runners in a race, for all races in period 1 in which prize money exceeds £4923.

<u>For period 1.</u>

Regressing or against rn,

$$or\% = \underset{(0.475)}{2.88} + \underset{(0.037)}{1.57}\ rn \qquad (6.13)$$

$R^2 = 0.6515$

Again the Cook-Weisberg statistic suggests a problem of heteroscedasticity. On the null hypothesis of homoscedasticity, the test statistic, distributed as $\sigma(1) = 375.47$, which is significant at the 0 per cent level. Employing standard errors which are robust to heteroscedasticity (table 6.18) do not alter the conclusions. The coefficient on the number of runners is still highly significant, although the over-round is now only about 1.5 per cent for every runner.

Table 6.18 Estimating the relationship between the bookmakers' over-round and the number of runners in a race, for all races in period 1 in which prize money exceeds £4923, using Huber standard errors.

For period 1.

Regressing or against rn,

$$or\% = \underset{(0.470)}{2.88} + \underset{(0.049)}{1.57}\ rn \qquad (6.14)$$

$R^2 = 0.6515$

The analysis is repeated for all races in this period in which the prize money of the race exceeds the median for all races. The results are presented in table 6.19.

Table 6.19 Estimating the relationship between the bookmakers' over-round and the number of runners in a race, for all races in period 1 in which prize money exceeds £2657.

<u>For period 1.</u>

Regressing or against rn,

$$or\% = \underset{(0.385)}{1.09} + \underset{(0.027)}{1.99}\, rn \qquad (6.15)$$

$R^2 = 0.6587$

Again the Cook-Weisberg statistic suggests a problem of heteroscedasticity. On the null hypothesis of homoscedasticity, the test statistic, distributed as $\sigma(1) = 562.54$, which is significant at the 0 per cent level. Employing standard errors which are robust to heteroscedasticity (table 6.20) do not alter the conclusions. The coefficient on the number of runners is still highly significant, although the over-round is about 2 per cent for every runner.

Table 6.20 Estimating the relationship between the bookmakers' over-round and the number of runners in a race, for all races in period 1 in which prize money exceeds £2657, using Huber standard errors.

<u>For period 1.</u>

Regressing or against rn,

$$or\% = \underset{(0.329)}{1.09} + \underset{(0.029)}{1.99}\, rn \qquad (6.16)$$

$R^2 = 0.6587$

6.4.3 Summary of results

In all three samples there is evidence of a positive and significant relationship between the over-round and the number of runners, which lends support to Henery's findings. The size of the coefficient is greater, however, for the sample of all races than for races whose prize money exceeded the median, and greater still than for those races whose prize money exceeded the mean (which was almost twice as great as the median). This is particularly interesting insofar as it suggests that the over-round per runner tends to fall as the quality of the race increases, and appears to decline quite sharply toward the upper end. This evidence is in line with Henery's assertion that "all 'prestige' races... have comparatively generous odds and must be regarded as atypical." (Henery, 1985, p.346).

In light of these findings, it would appear that Henery's hypothesis of bettor behaviour cannot be rejected unless it implies other results which are not consistent with the data, or unless an alternative explanation can be produced which is more consistent with all the observed evidence.

6.5 Employing the new data set to test for weak form efficiency in British racetrack betting markets: Summary and conclusions

In this chapter a totally new data set derived from the British flat horse race seasons (1992 and 1993) has been used to test for the existence of a favourite-longshot bias, and for a differential bias in races taking place at different times in the order of racing or in different parts of the season. The existence of a bias in expected returns against horses starting at higher odds is confirmed (the conventional favourite-longshot bias). However, there is no clear evidence of a differential bias in different races categorized in terms of race order, either on a particular day or across a season.

Tests are also performed of the hypothesis that bettors discount a fixed fraction of their losses, and of the implications of this for a link between bookmakers' over-round and the number of runners in a race. Weighted least squares regressions are applied, and reveal a reasonable goodness of fit for a specified linear relationship between the objective probabilities implied by actual race results and the subjective probabilities implied by bookmakers' odds. Moreover, that the subjective lose probability, as so defined, is significantly less than the empirical lose probability. This relationship was not significantly different for different parts of the season, or for different levels of prize money. However, a close relationship between the size of the over-round and the number of runners in a race was observed, although the size of the relationship was lower in races characterized by higher levels of prize money.

The implications of these findings for an understanding of the cause of the favourite-longshot bias in fixed-odds racetrack betting markets depends on an evaluation of the relative merits of an explanation couched in terms of bettors' behaviour (as in Henery, 1985) compared with one couched in terms of bookmakers' behaviour (as in

Shin, 1991, 1992, 1993). This issue is explored in greater detail in Vaughan Williams and Paton (1996, 1997a), and in chapter 8 of this thesis.

1. 1/2, 4/6 evens, 6/4, 2/1, 5/2, 3/1, 4/1, 5/1, 6/1, 8/1, 10/1, 12/1, 14/1, 16/1, 20/1, 25/1, 33/1, 50/1.

2. In decimalized from, these are:
0.111; 0.143; 0.167; 0.2; 0.222; 0.25; 0.286; 0.3; 0.333; 0.364; 0.4; 0.444; 0.5; 0.533; 0.571; 0.615; 0.667; 0.727; 0.8; 0.833; 0.909; 1; 1.05; 1.1; 1.2; 1.25; 1.375; 1.5; 1.625; 1.75; 1.875; 2; 2.125; 2.25; 2.5; 2.75; 3; 3.333; 3.5; 4; 4.5; 5; 5.5; 6; 6.5; 7; 7.5; 8; 8.5; 9; 10; 11; 12; 14; 16; 20; 22; 25; 28; 33; 40; 50; 66; 80; 100; 125; 150; 200; 250; 400; 500.

3. The regression figures are apparently presented the wrong way around in the original article, although the sense is clear.

4. See Gujarati (1992), pp.331-340.

CHAPTER SEVEN

EMPLOYING A NEW DATA SET TO TEST FOR SEMI-STRONG FORM EFFICIENCY IN BRITISH RACETRACK BETTING MARKETS

7.1 Introduction

In chapter 5, the new data set compiled for this project was described and explained. In chapter 6, this data set was used to test for the existence and nature of a favourite-longshot bias in recent British racetrack betting markets. Implications were drawn for the issue of weak form efficiency in these markets.

In the literature to date, tests of semi-strong efficiency in betting markets have highlighted differentials in the returns to separate types of bet which possess identical objective probabilities of success. They have also examined the possibility of constructing a successful forecasting model (or system), and tested for the existence of market 'anomalies', such as an over-reaction to recent information (or 'gambler's fallacy'). In this chapter, the new data set, as well as an extensive survey of the performance of professional forecasting services, are used to undertake new tests of the existence of semi-strong information efficiency in recent British racetrack betting markets.

In Section 7.2, an approach used by Crafts (1994) is adapted and developed. Crafts assessed the performance of a number of organisations which market betting systems (selling decision rules) and ratings services (selling selections). He writes:

"It seems reasonable to conclude that punters who buy racing systems in the hope of obtaining access to decision rules that will lead to profitable betting do so in vain- as a proponent of the efficient markets hypothesis would predict- despite the misleading claims made repeatedly in advertisements in the sporting press." (p.547).

This is a strong claim, and is based on an investigation of three systems marketed in the 1980s "at prices in the £10-£15 range" (p.545). The performance of six other racing systems and five ratings services were derived from a survey in Roberts and Newton (1987).

In order to test this claim, Section 7.2 investigates the recent performance of five leading tipping services are investigated here. Unlike in Crafts' survey, the cost of these services is upwards of several hundred pounds per annum. Each of the services is widely advertised in the media, and each continues to operate, presumably successfully (in the sense of business, if not tipping profitability). Because of the cost of these services, there is a need to distinguish between these types of service and other forecasts, such as newspaper tips, which are cheaply and publicly available.

In Sections 7.3, 7.4 and 7.5, three potential market 'anomalies' are investigated. First, the existence of a gambler's fallacy, already an area of investigation using U.S. racetrack betting data, by Metzger (1985) and Terrell and Farmer (1996). Other studies, such as Clotfelter and Cook (1993) and Terrell (1994) have also concentrated on its incidence in lottery and lottery-type play. The second test is for a calendar (specifically, a day-of-the-week) effect, suggested by work on the seasonality of returns in financial markets, traceable to studies by Bonin and Moses (1974) and Rozeff and Kinney (1976). The third is a rank order of favouritism effect suggested by, but not based upon, an analysis in Asch, Malkiel and Quandt (1982).

In Section 7.6, an investigation is conducted into the expected return to bets placed in different arenas but with equivalent objective probabilities. Two tests are proposed. First, tote and starting prices returns about identical outcomes are compared. Second, the expected return on the tote is compared with that of bookmakers about a similar 'exotic' bet, in this case forecasting the first two horses home. In particular, the tote Dual Forecast bet is compared with the bookmaker-run Computer Straight Forecast.

In Section 7.7, evidence that the differential between tote and starting price payouts can be linked to movements in bookmakers' odds is investigated. Implications are drawn for a view that the actual existence of information efficiency in a market may be related to a perception of information inefficiency.

7.2 The performance of professional racetrack forecasting services

In this section, an investigation is undertaken of the performance of professional racetrack tipping services over 1995. An extensive data set is collected for this purpose. Implications are drawn for the existence of information efficiency in these markets.

If forecasters can provide information which can be used to earn above-average returns, this constitutes prima facie evidence of inefficiency in the market which is forecast. The nature and extent of the inefficiency depends on the type of forecast and the reasons for the above-average returns.

Racetrack forecasts can roughly be divided into newspaper tips, published and widely available in a range of morning newspapers, subscription services and premium rate telephone services.

It is helpful to distinguish between the first set of forecasts, i.e. newspaper tips, which are cheaply and publicly available and the latter, i.e. premium rate and subscription services, which are normally provided to a small group of paying clients. Although there is no clear dividing line, and the strength of the test for efficiency is a matter of interpretation, the newspaper forecasts are nearer to a test of semi-strong efficiency, while the more privileged services can be construed as a slightly stronger test of efficiency.

The value of the distinction can be traced to the difference of opinion expressed in the work of Snyder (1978) and of Losey and Talbott (1980). Snyder performed tests of five forecasting services, which he termed "five strong tests of market efficiency" (p.277), an expression challenged by Losey and Talbott, who contended that Snyder's "experts" were not "insiders" in the sense of possessing monopolistic access to information. Since the forecasts chosen for examination by Snyder were widely and publicly available in advance of each race, Losey and Talbott proposed that Snyder's tests may be more accurately viewed as tests of semi-strong efficiency.

A similar interpretation may be placed on the analyses of newspaper forecasts and forecast odds undertaken by Figlewski (1979), Tuckwell (1983), Anderson, Clarke and Ziegler (1985), Asch and Quandt (1986), and Bird and McCrae (1987).

Crafts (1994) distinguished between betting systems, i.e. sets of decision rules which can be used by the purchaser to generate forecasts, and marketed forecasts, which are actual tips about the outcome of particular races. His tests of twelve racing systems and five forecasting services all failed to demonstrate evidence of the potential for earning ex-post profits.

In this section, tests are conducted exclusively of the performance of pay-as-you-learn professional forecasting services, and the conclusions must be interpreted in light of this. In selecting the appropriate services for analysis two factors have been taken into consideration. First, the success to date of these services as measured by the independent publication, the Racing Information Database. Second, the degree of public awareness of these services, as assessed by their advertising profile.

Four subscription services are chosen, three of which are widely publicized in the specialist racing press, and which cost upward of £99 per month (1995 subscriptions). Each of these has been assessed as one of the most profitable by the Racing Information Database at the time of selection. The fourth is advertised by direct mail to clients of the major bookmakers, and publishes detailed claims of previous success.

One premium rate telephone service is chosen. This is selected on two counts. First, it is identified as one of the most successful by the Racing Information Database at the time of selection; and second, its particularly high cost (compared to other premium rate lines) provides some indication of exclusivity.

Although every effort has been made to construct as complete a data set as possible, some data may be missing and errors in compilation/calculation may have arisen. The results of this academic study should not, therefore, be used to pass judgement upon the professional competence of any individual named service. In recognition of this, the forecasting services are referred to as Service A, Service B, Service C, Service D and Service E.

1. Telephone services.

Service A : £129 per month.

2. Subscription services.

a. Service B : £99 per month.

b. Service C : £945 per year.

c. Service D : £495 per year.

d. Service E : £120 per month.

7.2.1 The performance of 'tipping' service A

Service A is chosen because it is rated highly by the independent Racing Information Database (R.I.D.) which monitors the performance of racetrack forecasting services. In particular, for the category of premium rate telephone tipping services Service A was ranked second in the year from January 1, 1995 to February 28, 1995; and for February, 1995 taken alone (in the March, 1995 edition of the R.I.D.). No other service was ranked in the first two on both criteria. All profit/loss figures are calculated in the R.I.D. net of a ten per cent deduction to a win bet £100 level stake (£50 each way for each way advices).[1] All figures are calculated at starting price unless a specific early price is advised, and this advised price must be available from Ladbrokes, Corals, William Hill or the Tote at the time the advice is proofed to the R.I.D. The detailed findings are as follows, presented in table 7.1.

Table 7.1 Racing Information Database analysis of the performance of 'tipping' service A - for February, 1995; and for January/February, 1995. (Source: R.I.D., March, 1995).

	For February, 1995	For year to end of February, 1995
Number of advices given	67	114
Number of winners given	25	38
Proportion of winners to advices	37%	33%
Average telephone call length	10 mins 20 secs	10 mins 20 secs
Average call cost	£4.99	£4.99
Costs incurred	£129.00	£258.00
Net profit	£2510	£2468.00

This service provides almost daily forecasts of horses to be backed. No price proviso is included, and so the client is left to either attempt to obtain the best price or else take the starting price. For the purposes of the present assessment the return about winning bets is calculated at starting price. The results are presented before tax and also after tax, calculated at the ten per cent charged by the leading bookmakers at the time the advices were given[2], the eight per cent charged by some of the independents, and the six per cent charged by the on-course betting shops. Tax is calculated paid-on, i.e. paid in advance on the stake.

The profit (loss) to a level stake on each horse tipped is calculated, together with the standard deviation of the profit/loss. Service A in fact advises varying stakes, in principle from one to ten points (in fact, in this sample, from 2 to 10 points), and the same calculations are performed on the basis of this strategy also.

The sample is of bets advised between 31st.March, 1995 and 12th.June, 1995, and are presented in table 7.2.

Table 7.2 Analysis of the performance of Service A - 31st.March, 1995 to 12th.June, 1995.

Number of days service called and monitored	49
Average call length	10 minutes 20 seconds
Total cost of calls	£256.41
Number of bets advised	221
Total stake at level unit stakes	221
Total stake at advised stakes (1 to 10)	1487
Number of winning bets	68
Average number of winning bets	30.77%
Standard deviation (winning bets)	0.46
Longest winning sequence	8
Longest losing sequence	14
Average starting price of advices	3.24
Standard deviation of starting prices of advices	2.51
Profit to a level stake (pre-tax)	-5.66 (-2.56%)
Standard deviation of profit (loss)	1.67
Profit to a level stake (10% tax)	-27.45 (-12.42%)
Profit to a level stake (8% tax)	-23.34 (-10.56%)
Profit to a level stake (6% tax)	-18.92 (-8.56%)
Profit to advised stakes (pre-tax)	44.04 (2.96%)
Standard deviation of profit (loss)	11.72
Profit to advised stakes (10% tax)	-104.66 (-7.04%)
Profit to advised stakes (8%tax)	-74.92 (-7.76%)
Profit to advised stakes (6%tax)	-45.18 (3.04%)

Although a pre-tax profit could have been earned by backing all the Service A tips at the advised stakes before tax (i.e. on-course), this was very small and certainly not significant at conventional levels of significance (t= 0.24). Bets to a level stake yielded a loss as did bets to advised stakes at each of the levels of deduction operated and considered.

At the actual profit margin reported, to cover the cost of calls (£256.41) the required stake at pre-tax advised levels would have had to exceed £8657.63 in total, or £5.82 per point.

Although no significant profit was realizable at level or advised staking levels using the Service A advices, it is possible that a profit is realizable by betting only on the more heavily weighted tips. The rationale behind this expectation lies in the large number of regular advices associated with this service, many of which may simply be included to extend the length of the call. These more weakly advised horses (in the sense of a low advised points rating) would therefore simply act as noise about the genuine information. To test for this the profitability of the Service A advices sub-aggregated into the various points categories is presented below, in table 7.3. No one point bets were advised.

Tables 7.3 Analysis of the Performance of Service A, sub-aggregated by 'points strength' of advice; 31st March 1995 to 12th June 1995

Number of points advised	2	3	4	5	6	7	8	9	10
Number of bets advised	4	9	13	21	41	56	56	1	20
Total stake at level unit stakes	4	9	13	21	41	56	56	1	20
Number of winning bets	0	0	3	6	12	14	25	0	8
Average number of winning bets (%)	0	0	23.08	28.57	29.27	25.00	44.64	0	40
Longest winning sequence	0	0	2	1	3	3	5	0	3
Longest losing sequence	4	9	5	5	6	12	5	1	6
Average starting price of advices	9.50	9.56	5.15	3.84	3.18	2.73	2.09	1	2.12
Standard deviation of starting price of advices	0.87	4.70	2.23	2.23	1.64	1.46	1.20	0	1.15
Profit to a level stake (pre-tax)	-4	-9	-2.25	-3.125	3.32	-13.28	13.39	0	7.03
Standard deviation of profit (loss)	-	-	1.51	2.13	1.95	1.44	1.58	-	2.05
Expected profit to a level stake (pre-tax) (%)	-100	-100	-17.31	-14.88	8.10	-23.71	23.91	-100	25.15

The set of all advices involving recommended stakes of 8 and 10 points (the one advice at 9 points is hardly significant) showed a positive expected profit both pre-tax and post-tax. At recommended stakes below 8 points the case is less clear, with clear losses or very small gains in each category.

In order to test whether a sub-aggregation of Service A tips into 'genuine' advices (8 points and above) and 'noise' advices (7 points and below) is capable of generating significant positive expected returns, all advices are aggregated into these categories and tested for significance. The findings are presented below, in table 7.4.

Table 7.4 Analysis of the performance of Service A 'stronger advices' (8 points and above) and 'weaker advices' (below 8 points) - 31st.March, 1995 to 12th. June, 1995.

	7 points or below	8 points or above
Number of bets advised	144	77
Total stake at level unit stakes	144	77
Total stake at advised stakes (1 to 10)	830	657
Number of winning bets	35	33
Average number of winning bets	24.31%	42.86%
Standard deviation (winning bets)	0.43	0.49
Longest winning sequence	3	7
Longest losing sequence	14	5
Average starting price of advices	3.86	2.09
Standard deviation of starting prices of advices	2.80	1.19
Profit to a level stake (pre-tax)	-23.09(-16.03%)	17.42
Standard deviation of profit (loss)	1.66	1.65
Profit to a level stake (10% tax)	-37.49(-26.03%)	9.72 (12.62%)
Profit to a level stake (8% tax)	-34.61(-24.03%)	11.26 (14.62%)
Profit to a level stake (6% tax)	-31.73(-22.03%)	12.80 (16.62%)
Profit to advised stakes (pre-tax)	-106.42(-12.82%)	150.45 (22.90%)
Standard deviation of profit (loss)	9.86	14.41
Profit to advised stakes (10% tax)	-189.42(-22.82%)	84.75 (12.90%)
Profit to advised stakes (8% tax)	-172.82(-20.82%)	97.89 (14.90%)
Profit to advised stakes (6% tax)	-156.22(-18.82%)	111.03 (16.90%)

Although a strategy of betting on Service A advices with recommended stakes of 7 points or less would have yielded a pre-tax loss at both level and advised stakes, limiting bets to those with recommended stakes of 8 and above would have produced profits before and after tax at both level and advised stakes.

The next step is to test these profits for significance. The results are presented below, in table 7.5.

Table 7.5 Testing the significance of the profitability of Service A's 'stronger advices' (8 points and above).

	Level Stakes	Advised Stakes
Expected profit (pre-tax)	0.2262	1.954
Standard deviation (actual profits)	1.65	14.41
Significance of expected profit (pre-tax)	t = 1.20	t = 1.19

Neither of the expected profits is significant at any conventional level of confidence.

An examination of the actual profit margin which was available from betting on all advices with recommended stakes of 8 and above, regardless of significance, allows us to determine how much would have had to have been wagered in order to cover the cost of the calls (£256.41). The findings of such an examination are presented below, in table 7.6.

Table 7.6 Calculating the total stake required to "break even" on Service A's 'stronger advices' (8 points and above), net of information costs.

<u>For bets of 8 points and above.</u>

At level stakes (pre-tax) total stake required to cover cost of calls = £1133.39 (or £14.72 per bet).

At level stakes (10% tax) total stake required to cover cost of calls = £2031.23 (or £26.38 per bet).

At level stakes (8% tax) total stake required to cover cost of calls = £1720.93 (or £22.77 per bet).

At level stakes (6% tax) total stake required to cover cost of calls = £1542.47 (or £20.03 per bet).

At advised stakes (pre-tax) total stake required to cover cost of calls = £1119.72 (or £1.70 per point).

At advised stakes (10% tax) total stake required to cover cost of calls = £1987.75 (or £3.03 per point).

At advised stakes (8% tax) total stake required to cover cost of calls = £1753.43 (or £2.62 per point).

At advised stakes (6% tax) total stake required to cover cost of calls = £1517.26 (or £2.31 per point).

7.2.1.1 Service A premium rate telephone advisory service: Summary and conclusions

The Service A tipping service operates on a premium rate telephone number, and issues daily recommendations of horses to bet upon at a range of recommended stakes (from 1 to 10 points). An examination of the performance of these tips over a three-month period, covering 221 recommendations, reveals that a profit could have been made, even after tax, for a total stake exceeding £1987.75 (or £1119.72 before tax). This is based upon a strategy of betting only on those advices accompanied by the highest recommended stakes (i.e. 8 points or greater), and betting at the advised levels. A strategy of betting at level stakes in this upper range would also have yielded profits pre- and post-tax, albeit at the cost of a greater total stake (£2031.23 and £1133.39 respectively). For all bets, at the advised stakes, a total sum in excess of £8657.63 would have been required to provide a pre-tax profit. Given the standard deviations of the returns, however, none of the profits reported above could be confirmed as significant at conventional levels of significance.

At level stakes, bets on all advices would have yielded a pre-tax loss. At any of the standard rates of tax a strategy of betting on all advices (whether at level or advised stakes) would also have yielded a net loss.

While, therefore, it is possible to confirm the opportunity to have earned a positive profit from this Service over the period examined, given the size of the stakes required and the variance of the returns a strategy of following this service in the future is a risky one. Moreover, there is no convincing evidence that this additional risk will be compensated by additional positive net returns.

7.2.2 The performance of 'tipping' service B.

Service B was rated as the best performer among subscription services (in terms of profit in the year to the end of May, 1995) in the June 1995 edition of the Racing Information Database (R.I.D.). The service is also extensively promoted in the racing press and media.

As before, all profit/loss figures are calculated net of a ten per cent deduction to a win bet £100 level stake (£50 each way for each way advices). All figures are calculated at starting price unless a specific early price is advised, and this advised price must be available from Ladbrokes, Corals, William Hill or the Tote at the time the advice is proofed to the R.I.D.

The detailed findings are as follows, presented in table 7.7.

Table 7.7 Racing Information Database analysis of the performance of tipping service B - for January to May, 1995. (Source: R.I.D., June, 1995).

<u>For the year to the end of May, 1995.</u>

Number of advices given: 134

Number of winners given: 57

Proportion of winners to advices: 43%

Fee charged : £99 per month; £169 for two months; £229 for 3 months; £399 for 6 months; £699 for 12 months.

Costs incurred: £399.00

Net profit : £6888

This service provides almost daily forecasts of horses to be backed. No price proviso is included, except for a general claim that no odds-on advices are given, i.e. horses with a price of less than evens. The client is left, therefore, either to attempt to obtain the best price or else take the starting price. Where an advice actually does show as odds-on the client can either take it at starting price, or else avoid the bet altogether. For the purposes of the present assessment the return about winning bets is calculated at starting price. The returns are presented inclusive and exclusive of all odds-on starting prices. The results are presented before tax and also after tax (calculated at the ten per cent charged by the leading bookmakers at the time the advice was given, the eight per cent charged by some of the independents, and the six per cent charged by the on-course betting shops). Tax is calculated paid-on, i.e. paid in advance on the stake.

The profit (loss) to a level stake on each horse tipped is calculated, together with the standard deviation of the profit/loss.

The sample is of bets advised between 15th.April, 1995 and 29th.September, 1995.

Where horses are advised "each way" a bet is classified as winning if the return exceeds the stake. The findings are presented below, in table 7.8.

Table 7.8 Analysis of the performance of Service B - 15th.April, 1995 to 29th. September, 1995.

	Including odds-on	Excluding odds-on
Number of days service called and monitored	73	73
Total cost of subscription	£399	£399
Number of bets advised	107	89
Total stake at level unit stakes	107	89
Number of winning bets	44	34
Average number of winning bets	41.12%	38.20%
Standard deviation (winning bets)	0.55	0.55
Longest winning sequence	6	5
Longest losing sequence	7	6
Average starting price of advices	3.09	3.58
Standard deviation of starting price of advices	3.27	3.38
Profit to a level stake (pre-tax)	4.72 (4.41%)	7.98 (8.97%)
Standard deviation of profit (loss)	1.44	1.53
Profit to a level stake (10% tax)	-5.98 (-5.08%)	-0.92 (0.94%)
Profit to a level stake (8% tax)	-3.84 (-3.32%)	0.86 (0.89%)
Profit to a level stake (6 % tax)	-1.7 (-1.50%)	2.64 (2.80%)

Although a pre-tax profit could have been earned by backing all the Service B tips (inclusive and exclusive of odds-on starting prices) at level stakes, this was very small. Bets to a level stake yielded a loss at each of the levels of deduction operated and considered, except for the very small positive profit realizable from a strategy of betting exclusive of odds-on advices at deductions of eight per cent and below.

At the actual profit margin reported, to cover the cost of the subscription (£399) the required stake (inclusive of odds-on advices) before tax would have had to exceed £9047.62 in total or £84.56 per bet. At any conventional level of tax no unique level stake could have yielded a profit. Exclusive of odds-on advices, the required stake before tax would have had to exceed £4448.16 in total or £49.98 per bet. The corresponding amounts for deductions of eight per cent on stakes are £46395.34 and £521.30. For a six per cent deduction the figures are £15113.64 and £169.82.

The next step is to test whether the profits, where observed, are statistically significant. The results are presented below, in table 7.9.

Table 7.9 Testing the significance of the profitability of Service B.

	Including odds-on	Excluding odds-on
Expected profit (pre-tax), to level stakes	0.0441	0.0897
Standard deviation (actual profits)	1.44	1.53
Significance of expected profit (pre-tax)	t = 0.32	t = 0.55

Neither of the expected profits is significant at any conventional level of confidence.

7.2.2.1 Service B subscription telephone advisory service: Summary and conclusions

Service B operates on a subscription basis. The service offers (via a telephone answering service) almost daily recommendations of horses to bet upon to level stakes. An examination of the performance of these tips over a six-month period, covering 107 recommendations, reveals that a profit could have been made before tax, if ignoring odds-on advices, for a total stake exceeding £4448.16 (or about £50 per bet). Small profits after deduction of tax at eight per cent or below could also have been achieved, but only to a far greater average stake. Given the standard deviations of the returns, however, none of the profits reported above could be confirmed as significant at conventional levels of significance.

It was possible, therefore, to have earned a positive profit from Service B over the period examined, at least on-course and in betting shops levying deductions of eight per cent or less. However, given the size of the stakes required, and the variance of the returns, a strategy of following this service in the future is a risky one. Moreover, there is no convincing evidence that this additional risk will be compensated by additional positive net returns.

7.2.3 The performance of 'tipping' service C.

Service C was rated as the second best performer (in terms of profit in the year to the end of February, 1995) in the March 1995 edition of the Racing Information Database (R.I.D.). In this period it was out-performed by only one service, and that service subsequently dropped out of the top ten listed services. Service C has remained near the top in subsequent editions of the R.I.D., e.g. fifth in the year to the end of May.

This service is also very extensively promoted in the racing press and media, and makes repeated claims to be the most consistent performer of all tipping services.

As before, all profit/loss figures are calculated net of a ten per cent deduction to a win bet £100 level stake (£50 each way for each way advices). All figures are calculated at starting price unless a specific early price is advised, and this advised price must be available from Ladbrokes, Corals, William Hill or the Tote at the time the advice is proofed to the R.I.D.

The detailed findings are as follows, presented in table 7.10.

Table 7.10 Racing Information database analysis of the performance of tipping service C - for February, 1995, for January/February, 1995; and for January to May, 1995. (Source: R.I.D., March 1995 & June, 1995).

	February, 1995	To end Feb, 1995	To end May,1995
Number of advices given	40	55	120
Number of winners given	15	20	37
Proportion of winners to advices	38%	36%	31%
Fee charged (per year)	£945	£945	£945
Costs incurred	£79.00	£158.00	£394.00
Net profit	£831	£1656	£3728

This service provides almost daily forecasts of horses to be backed. No price proviso is normally included, although a minimum price stipulation is sometimes made. Bets are advised to varying stakes, from a quarter point to three points in the sample range thus far considered. The standard stake is half a point and this is taken as the unit stake for these purposes. A one point bet is, for example, taken as a two unit stake. For the purposes of the present assessment the return about winning bets is calculated at starting price for all bets advised. A separate analysis is conducted which eliminates those bets which do not meet the minimum price stipulation, where relevant. Where the minimum price is achieved, the return is calculated at the price at which the stipulated minimum price is first offered. The results are presented to the advised stakes and to level stakes. The results are calculated before tax and also after tax (calculated at the ten per cent charged by the leading bookmakers at the time the advice was given, the eight per cent charged by some of the independents, and the six per cent charged by the on-course betting shops). Tax is calculated paid-on, i.e. paid in advance on the stake.

The profit (loss) to a level stake on each horse tipped is calculated, together with the standard deviation of the profit/loss.

The sample is of bets advised between 4th.April, 1995 and 2nd.December, 1995.

Where horses are advised "each way" (i.e. to be placed), a bet is classified as winning if the return exceeds the stake.

The findings are presented below, in table 7.11.

Table 7.11 Analysis of the performance of Service C - 4th. April, 1995 to 2nd. December, 1995.

	At SP (advised stakes)	At SP (level stakes)
Number of bets advised	159	159
Total stake at advised stakes	188	159
Number of winning bets	46.5 (inc. a dead-heat)	46.5 (inc. a dead-heat)
Average number of winning bets	29.25%	29.25%
Standard deviation (winning bets)	0.53	0.53
Longest winning sequence	3	3
Longest losing sequence	6	9
Average starting price of advices	4.13	4.13
Standard deviation (SP of advices)	4.12	4.12
Profit (pre-tax)	41.79 (22.23%)	7.30 (4.59%)
Average profit per bet	0.26	0.046
Standard deviation of profit (loss)	3.19	1.99
Profit (10% tax)	22.99 (11.12%)	-8.6 (-4.92%)
Profit (8% tax)	26.75 (13.17%)	-5.42 (-3.16%)
Profit (6% tax)	30.51 (15.31%)	-2.24 (-1.33%)

Although a pre-tax profit could have been earned by backing all the Service C tips at level and advised stakes before tax (i.e. on-course), the post-tax return was only positive at advised stakes. However, the profit could not be confirmed at conventional levels of significance (t= 1.04).

At the actual profit margin reported, to cover the cost of a 6 months subscription (say £500) would have required a total stake of £2249 or £14.15 per bet. The required total stake net of ten per cent tax would have had to exceed £4496 in total, or £28.28 per bet.

It is possible that a higher rate of profit is realizable by betting only on the more heavily weighted advices. The advices at a lower staking level may be expected to simply pad out the genuine inside information, or else to reduce the variance of the return. To test for this the profitability of the Service C advices sub-aggregated into the various points categories is undertaken below. Table 7.12 examines the returns to the various points advices, ranging from horses about which the caller is advised to place a quarter point (half a unit stake), to those advices in the maximum three points category (6 unit stakes).

Table 7.12 Analysis of the performance of Service C sub-aggregated by 'points strength' of advice; 4th April 1995 to 2nd December 1995

Number of points advised	1/4	1/2	1	2	3
Number of bets advised	20	110	26	1	2
Total stake at level unit stakes	20	110	26	1	2
Number of winning bets	2	26	12	1	1
Profit to a level stake (pre-tax)	-9.5	-2.13	16.73	2.1	3.2

The set of all advices involving recommended stakes of one or above appears to perform better of expected profit than those below one point, the former showing a clear pre-tax and post-tax profit, the latter a loss.

In order to test whether a sub-aggregation of Service C tips into 'genuine' advices (1 point and above) and 'noise' advices (below 1 point) is capable of generating positive expected returns, all advices are aggregated into these categories and examined separately.[3]

The findings are presented in table 7.13.

Table 7.13 Analysis of the performance of service C's 'stronger advices' (1 point and above) and 'weaker advices' (below 1 point); 4th.April, 1995 to 2nd.December, 1995.

	1 points or above	**1 point or below**
Number of bets advised	29	130
Total stake at level unit stakes	29	130
Total stake at advised stakes	68	120
Number of winning bets	13	30
Average number of winning bets	44.83%	23.08%
Longest winning sequence	3	2
Longest losing sequence	3	11
Average starting price of advices	2.45	4.50
Standard deviation of starting prices of advices	1.15	4.43
Profit to a level stake (pre-tax)	18.93 (65.28%)	-11.63 (-8.95%)
Average profit to a bet (level stakes)	0.65	0.089
Standard deviation of profit (loss)	2.12	1.94
Profit to a level stake (10% tax)	16.03 (50.25%)	-24.63 (-17.22%)
Profit to a level stake (8% tax)	16.61 (53.03%)	-22.03 (-15.69%)
Profit to a level stake (6% tax)	17.19 (55.92%)	-19.43 (-14.10%)
Profit to advised stakes (pre-tax)	48.67 (71.57%)	-6.88 (-6.08%)
Average profit to a bet (advised stakes)	1.68	-0.057
Standard deviation of profit (loss)	6.15	1.85
Profit to advised stakes (10% tax)	41.87 (55.98%)	-18.88 (14.30%)
Profit to advised stakes (8% tax)	43.23 (58.86%)	-16.48 (-12.72%)
Profit to advised stakes (6% tax)	44.59 (61.86%)	-14.08 (-11.07%)

In summary, the more heavily weighted bets (1 point and above) produced positive profits, particularly when weighted by the recommended stake size. No such profit was available to the less heavily weighted bets (bets of less than one point). A significance test of the return to bets of one point and above was undertaken, and is presented below in table 7.14.

Table 7.14 Significance of the performance of Service C's 'stronger advices' (1 point and above); 4th.April, 1995 to 2nd.December, 1995.

<u>Bets of one point and above to level stakes.</u>

Average profit per bet = 0.65.

Standard deviation (profit/loss) = 2.12.

Number of bets = 29.

For the difference between 0.65 and 0, t = 1.65.

This is significant at the ten per cent level of confidence.

<u>Bets of one point and above to advised stakes.</u>

Average profit per bet = 1.68.

Standard deviation (profit/loss) = 6.15.

Number of bets = 29.

For the difference between 1.68 and 0, t = 1.47.

This difference could not be confirmed as significant at conventional levels of significance.

At the actual profit margin reported, to cover the cost of a 6 months subscription (say £500) would have required a total stake at level stakes of £766 or £26.41 per bet. The required total stake net of ten per cent tax would have had to exceed £995 in total, or £34.31 per bet. At advised stakes a total stake of £698.58 would have been required, or £24.09 per bet, or £10.27 per point. Net of ten per cent tax the required stake was £893.18, or £30.80 per bet, or £13.14 per point.

All the above calculations assume that every advice is taken at the starting price. Although this is the usual recommendation, sometimes a minimum price stipulation is included with the recommendation. When this is the case the horse is only to be backed at this price or greater. Below (table 7.15) the result is presented of ignoring all advices which did not meet this minimum price requirement. Where the price requirement is met the price is taken which first exceeds the stipulation during the development of the market. If the price subsequently lengthens this is ignored.

Table 7.15 Comparing the returns to all Service C advices at starting price with the return allowing for the advised minimum price stipulation.

<u>At starting price.</u>

At Level stakes : Profit = 7.30; average profit/bet = 0.046; standard deviation (profit/loss) = 1.99.

At Advised stakes : Profit = 41.79; average profit/bet = 0.26; standard deviation (profit/loss) = 3.19.

Net of minimum price provisos.

At Level stakes : Profit = 6.40; average profit/bet = 0.04; standard deviation (profit/loss) = 2.06.

At Advised stakes : Profit = 41.25; average profit/bet = 0.29; standard deviation (profit/loss) = 3.31.

The average return at both advised stakes is thus slightly improved by following the price advices, although not at level stakes. The differences in each case, however, are small.

7.2.3.1 The performance of Service C: Summary and conclusions

Service C operates on a subscription basis. The service offers (via a telephone answering service) almost daily recommendations of horses to bet upon to advised stakes. An examination of the performance of these tips over a six month period, covering 159 recommendations, reveals that a profit could have been earned both to advised and level stakes, although the post-tax profitability resulted solely from the performance of the more heavily advised recommendations (i.e. one point tips and above). For these bets (one point and above) the service was able to produce a profit net of subscription for bets averaging about thirty pounds. Although there was some evidence of significance in the profitability this was at a low level, perhaps because of the low number of more heavily weighted advices.

While, therefore, it is possible to confirm the opportunity to have earned a positive post-tax profit from Service C over the period examined, particularly by restricting the bets to a subset of all those advised, it is not possible from this sample to confirm whether the results are indicative of genuinely profitable information or simply good fortune.

7.2.4 The performance of 'tipping' service D

Service D is an example of direct-mail tipping agencies. Such agencies write directly to potentially interested parties gathered from mailing lists supplied by credit bookmakers. Service D charged £495 for advices in 1995, and published detailed claims to have made substantial profits in the previous year. However, the advices are not checked and verified by the Racing Information Database. The advices are supplied via an answering machine.

All recommendations are advised to a level unit stake, and many of these are advised "each way" (i.e. to be placed).

The service was monitored for these purposes between 11th. February 1995 and 9th.December 1995.

The results are published based upon the staking patterns advised for each bet (win or each way) and separately on the assumption that all recommendations bets should be staked as level unit win bets - see Table 7.16.

For the purposes of assessing the success rate of each way bets a bet is classified as a winning one if the return exceeds the stake.

Two bets were advised 'ante-post', i.e. prior to the day of the race. Such bets are calculated at the best price obtainable from the following bookmakers - Ladbrokes, William Hill, Corals and Tote - when the recommendation was made. All other bets are calculated at the starting price. The results are published pre-tax, inclusive of ten per cent tax, as levied at the time of the advice by the leading bookmakers, eight per cent tax and six per cent tax. The tax is calculated paid-on, i.e. paid in advance on the stake.

Table 7.16 Analysis of the performance of Service D - to advised staking patterns, i.e. win or "each way"; 11th.February, 1995 to 9th.December, 1995.

	All bets to advised staking patterns	**All bets placed "Win Only"**
Number of bets advised	139	139
Number of bets advised "each way"	77	-
Number of bets advised "win only"	62	-
Total stake	139	139
Number of successful bets	49	43
% "winning" bets of all bets	35.25%	30.94%
Longest winning sequence	6	6
Longest losing sequence	6	9
Average starting price of all advices	4.24	4.24
Standard deviation (SP of all advices)	11.89	11.89
Profit (pre-tax)	20.66 (14.86%)	42.86 (30.83%)
Average profit per bet	0.1486	0.3083
Standard deviation (profit/loss) per bet	1.78	2.54
Profit (10% tax)	6.76 (4.42%)	28.96 (18.94%)
Profit (8% tax)	9.54 (6.35%)	31.66 (21.09%)
Profit (6% tax)	12.32 (7.76%)	34.52 (23.43%)

A profit was realized to level stakes on all bets advised by Service D, although this profit was actually greater if all advices were placed as win bets rather than splitting the recommendations into each way and win bets along the lines advised by this service. The increased profit was made at the cost of slightly increased risk shown in the longest losing sequence (nine instead of six).

The next step is to test the profits for significance. The results are presented below, in table 7.17.

Table 7.17 Testing the significance of the profitability of Service D.

<u>For all bets to advised staking patterns.</u>

Average profit (pre-tax) per bet = 0.1486.

Standard deviation (profit/loss) = 1.78.

Number of bets = 139.

For the difference between 0.1486 and 0, t = 0.98.

<u>For all bets placed 'win only'.</u>

Average profit (pre-tax) per bet = 0.3083.

Standard deviation (profit/loss) = 2.54.

Number of bets = 139.

For the difference between 0.1486 and 0, t = 1.43.

Although the pre-tax and post-tax profits are positive in both cases, these profits could not be confirmed even pre-tax at any conventional level of significance.

In interpreting these results it is worth noting that two of the advices were made prior to the day of the race, i.e. ante-post. Both these advices won. The best ante-post price available form the leading bookmaking firms at the time of the recommendations was 11 to 1 and 14 to 1. These shortened to an eventual starting price of 7 to 1 and 5 to 1 respectively. These two advices are responsible for much of the net profit. To show this the pre and post-tax profits to advised and win-only betting patterns are presented below, first by substituting the starting prices for the ante-post prices (table 7.18), and secondly by excluding from the analysis these two ante-post recommendations (tables 7.19).

Table 7.18 Analysis of the performance of Service D - all returns taken at starting price; 11th February 1995 to 4th December 1995

	All bets to advised staking patterns, at starting price	All bets placed "win only" at starting price
Number of bets advised	139	139
Profit (pre-tax)	13.54	29.9
Expected profit (pre-tax) (%)	9.74	21.51
Profit (10% tax)	-0.36	16.00
Profit (8% tax)	2.42	18.78
Profit (6% tax)	5.2	21.56
Expected profit (10% tax)	-0.24	10.46
Expected profit (8% tax)	1.61	12.51
Expected profit (6% tax)	3.53	14.63

Table 7.19 Analysis of the performance of Service D - all returns taken at starting price (ante-post advices excluded from sample) - 11th February 1995 to 4th December 1995

	All bets to advised stakes, at starting price (excluding ante-post)	All bets placed "win only", at starting price (excluding ante-post)
Number of bets advised	137	137
Profit (pre-tax)	5.04	17.86
Expected profit (pre-tax) (%)	3.68	13.04
Profit (10% tax)	-8.67	4.16
Profit (8% tax)	-5.92	6.9
Profit (6% tax)	-3.18	9.64
Expected profit (10% tax)	-6.10	2.76
Expected profit (8% tax)	-4.00	4.66
Expected profit (6% tax)	-2.32	6.64

Although profits still remain to a policy of betting level stakes on all advised bets at starting price, and even (before tax) exclusive of the high-priced ante-post successes, the profits are much reduced and not significant at any conventional level of significance.

7.2.4.1 The performance of Service D: Summary and conclusions

Service D was able to show a clear post-tax profit to a policy of following all of its advices at level stakes, although this profit was greater if a standard policy of betting win only was substituted for the Service's own staking system of combining win and each-way recommendations. Much, though not all, of the profits could be explained by the success of two high-priced ante-post tips. Standard statistical tests applied to the positive profit failed, however, even in the best case, to confirm the significance of any of the profits shown at conventional levels of significance.

7.2.5 The performance of 'tipping' service E

Service E is regularly and prominently advertised in the racing trade newspapers, and makes detailed claims of high and consistent profits for its followers. In addition to a telephone pay-as-you call service, Service E operates a subscription service, at a cost of £125 per month, or £175 per month including a guarantee of a subscription refund should the months bets show a loss. Unlike most services the system adopted by this service requires the bettor to obtain a price equal to or higher than a minimum price stipulation applied to most bets. If this price is not obtained the bet is void.

In assessing the profitability of this service three possible approaches may be adopted. The first is to calculate the profit/loss to all bets at starting price regardless of the minimum price proviso. The advantage of this is that it allows a straight comparison with other services. This approach is applied by the Racing Information Database and results in a ranking for his pay-as-you-call service of fifth among telephone advisory services for February, 1995 and for the two months to the end of February, 1995. The figures for the subscription service are not presented.

The detailed findings are presented below, in table 7.20.

Table 7.20 Racing Information Database analysis of the performance of tipping service E - for February 1995 and for January/February 1995 (Source: R.I.D., March, 1995).

	For February 1995	For January/February 1995
Number of advices given	28	46
Number of winners given	10	15
Proportion of winners to advices	36%	33%
Average telephone call length	2 minutes 0 seconds	2 minutes 0 seconds
Average call cost	£0.98	£0.98
Cost incurred	£23.52	£47.00
Net profit	£1333	£1886.00

The second way of calculating the profits is to employ the best price that was available about the horse but only when this price exceeds the minimum price stipulation. This is the method of accounting adopted by the Service itself. This is clearly unfair, however, as the bettor does not know in advance of taking a price which way the odds will move. If, for example, the minimum price stipulation is 3 to 1 and a horse is currently showing at 3 to 1, the bettor does not know for sure which way the price will move if at all. If the price subsequently shortens then the bet is no longer applicable although it will be counted in the Service's accounts at 3 to 1. If it lengthens from 3 to 1 to 4 to 1, however, it is counted as a 4 to 1 bet. However, this is an accounting option not available in advance to the bettor. The third approach is to employ the price which first obtains which is equal to or greater than the minimum price stipulation. If no such price does obtain the bet is considered void for the purpose of an assessment of the returns.

For these purposes the first and third approaches are adopted.

Service E attaches different recommended stakes to different bets (from 1 to 9 in the sample considered). The returns to the recommended stakes and to level stakes are calculated employing both approaches.

For the purposes of the present assessment the return about winning bets is calculated at starting price. The results are presented before tax and also after tax (calculated at the ten per cent charged by the leading bookmakers, the eight per cent charged by some of the independents, and the six per cent charged by the on-course betting shops). Tax is calculated paid-on, i.e. paid in advance on the stake. The profit (loss) to a level stake on each horse tipped is calculated, together with the standard deviation of the profit/loss.

The sample is of bets advised between 1st.May, 1995 and 9th.December, 1995. Table 7.21 presents the results from the first approach, i.e. an analysis of all bets advised at the starting price, to level stakes and to advised stakes.

Table 7.22 presents the results from the other approach proposed, i.e. the returns to all bets for which the odds equalled or exceeded the minimum price stipulation. The price adopted is the first price which obtains in the market which was equal to or greater than the minimum price stipulation. The pre-tax profit for this set of runners at the starting price is also included at the bottom of the tables for comparison purposes.

On the very few occasions on which an each way bet is advised, the bet is classified as a winning one if the return exceeds the stake.

Table 7.21 Analysis of the performance of Service E - to level and advised stakes - at starting price - 1st May 1995 to 9th December 1995

	All bets to level stakes - at starting price	All bets to advised stakes - at starting price
Number of bets advised	190	190
Total stake	190	985
Number (%) of winning bets	63 (33.16)	63 (33.16)
Longest winning sequence	4	4
Longest losing sequence	12	12
Average starting price of all advices	3.24	3.24
Standard deviation (starting price) of all advices	2.27	2.27
Profit, pre-tax (expected profit)	20.99 (11.05%)	105.03 (10.66%)
Average profit per bet	0.1105	0.5526
Standard deviation (profit/ loss) per bet	1.81	9.54
Profit, including 10% tax (expected profit)	1.99 (0.95%)	6.5 (0.60%)
Profit, including 8% tax (expected profit)	5.79 (2.82%)	26.5 (24.9%)
Profit, including 6% tax (expected profit)	9.59 (4.76%)	45.93 (4.40%)

Table 7.22 Analysis of the performance of Service E - to level and advised stakes for bets meeting Service E's minimum price proviso; 1st May 1995 to 9th December 1995

	Bets meeting minimum price proviso - at level stakes	Bets meeting minimum price proviso - at advised stakes
Number of bets advised	153	153
Total stakes	153	780
Number (%) of winning bets	44 (28.76%)	44 (28.76%)
Longest winning sequence	4	4
Longest losing sequence	14	14
Average starting price of all advices	3.62	3.62
Standard deviation (starting price) of all advices	2.36	2.36
Profit, pre-tax (expected profit)	27.66 (18.08%)	98.77 (12.66%)
Average profit per bet	0.1808	0.6454
Standard deviation (profit/ loss) per bet	2.15	9.33
Profit, including 10% tax (expected profit)	12.36 (7.34%)	20.77 (2.42%)
Profit, including 8% tax (expected profit)	15.42 (9.33%)	36.37 (4.35%)
Profit, including 6% tax (expected profit)	18.48 (11.39%)	51.97 (6.29%)
Profit (pre-tax) at starting price	14.92 (9.75%)	57.18 (7.33%)
Average profit per bet at starting price	0.0975	0.0733
Standard deviation (profit/ loss) at starting price	1.92	8.73

A policy of betting level or advised stakes on all recommendations at starting price would have yielded post-tax profits. Post-tax profitability could also have been obtained by a policy of taking the first price offered in the market which matched or exceeded the minimum price stipulation. A policy of identifying the horses which met the minimum price proviso and taking the starting price about each of these, would have yielded lower profits than taking the price when it first obtained, albeit still a positive pre-tax profit.

The next step is to perform significance tests on the profit obtained by each of these approaches, the results of which are presented below (table 7.23).

Table 7.23 Testing the significance of the profitability of Service E - to level and advised stakes - at starting price; 1st. May, 1995 to 9th.December, 1995.

<u>For all bets at SP - to level stakes.</u>

Average profit per bet = 0.1105.

Standard deviation (profit/loss) = 1.81.

Number of bets = 190.

For the difference between 0.1105 and 0, t = 0.84.

<u>For all bets at SP - to advised stakes.</u>

Average profit per bet = 0.5526.

Standard deviation (profit/loss) = 9.54.

Number of bets = 190.

For the difference between 0.5526 and 0, t = 0.80.

For bets meeting minimum price proviso - to level stakes.

Average profit per bet = 0.1808.

Standard deviation (profit/loss) = 2.15.

Number of bets = 153.

For the difference between 0.1808 and 0, t = 1.04.

For bets meeting minimum price proviso - to advised stakes.

Average profit per bet = 0.6454.

Standard deviation (profit/loss) = 9.33.

Number of bets = 153.

For the difference between 0.1808 and 0, t = 0.86.

Although the pre- and post-tax profits are positive in both cases, these profits could not be confirmed at any conventional level of significance.

7.2.5.1 The performance of Service E: Summary and conclusions

Service E was able to show a clear post-tax profit to a policy of following all of its advices at level and advised stakes, although this profit was greater if bets were confined to the cases in which the price obtainable met or exceeded the minimum price stipulations normally built into the recommendation. In such cases the profit was greater from taking the first such price available than taking the starting price once the price had met the pre-defined minimum. Predictably, the profits are greater still if the longest price available is taken (Service E's mode of accounting), but there is no clear way of knowing this in advance. In any case, standard statistical tests applied to the positive profits failed, even in the best case, to confirm the significance of any of the profits at conventional levels of significance.

7.2.6 The performance of five professional forecasting services: Summary and conclusions

All the professional forecasting services examined produced a pre-tax profit to the advised staking and betting plans. However, in no case were these profits sufficiently large so as to permit confirmation of their significance at conventional levels. In some cases it was possible to improve upon the overall performance by ignoring the less strongly recommended tips. This may be because the genuine information is somewhat obscured by the 'noise' generated by some sort of implicit requirement for the agencies to issue a reasonable quantity as well as quality of tips. There is no unanimous evidence to support a view that variable stakes performed better than level stakes to all recommendations. Where minimum price stipulations were included there was some evidence that the expected profits increased if this price advice was followed, indicating that an assessment of performance at starting prices alone may be understating the performance of some of these agencies. The reason may be that the inside information may be gradually incorporated into the price during the development of the market.

Overall there is some evidence to indicate that professional forecasters can secure at least pre-tax and sometimes post-tax profits, and more surely to out-perform a policy of random selection.

However, the extent of these profits is not sufficient, within the existing individual samples, to confirm the possibility of making even a pre-tax profit at the 95 per cent level of confidence.

Although these findings are indicative of information inefficiency in the sense of an ability to earn above-average returns, there is little convincing evidence that they can be employed to earn abnormal returns, particularly net of transactions costs.

7.3 Testing for a 'gambler's fallacy'

The 'gambler's fallacy' is the proposition that bettors, instead of accepting an actual independence of successive outcomes, are influenced in their perceptions of the next possible outcome by the results of the preceding sequence of outcomes, e.g. throws of a die, or spins of a wheel. This idea was presented in a generalized form by Kahneman and Tversky (1974, 1982). Their idea of a 'winner's blessing and loser's curse' is a reported tendency for people, in revising their beliefs, to overweight new information and underweight older information. Such a hypothesis has been extensively tested in financial markets, e.g. De Bondt and Thaler (1985), Dark and Kato (1986), Dyl and Maxfield (1987), Brown and Harlow (1988), Stock (1990). Each of these studies found evidence of the existence of such an 'anomaly', and have yielded, therefore, the idea of trading upon a contrarian strategy. Insofar as this strategy is based on the historical pattern of past prices, it provides at least prima facie support for a hypothesis of weak form inefficiency in these markets. Since it also implies a failure by traders to allow for all public information, it is also indicative of a semi-strong form inefficiency. However, some authors explain any above-average or abnormal returns which can be elicited by acting on the basis of the above as fair compensation for additional risk or other factors.

Vermaelen and Verstringe (1986) and Chan (1986, 1987), for example, argue that a higher return is fair compensation for the greater risk incurred in buying losers compared with winners. The existence of a 'gambler's fallacy' has also been documented in lottery-type games (Clotfelter and Cook, 1991, 1993) and lotteries (Terrell, 1994). Two studies of a 'gambler's fallacy' in racetrack betting find the same effect - Metzger (1985) in U.S. horserace betting, and Terrell and Farmer (1996) in U.S. greyhound racing.

Metzger (1985) set out to test one prediction consistent with the concept of a gambler's fallacy, specifically that bettors will tend to overestimate the chances of favourites winning after a series of series of wins by longshots[4] compared to the situation after a series of wins by favourites. On the basis of an examination of a sample of U.S. horse races, Metzger concluded that there was indeed such a tendency shown by bettors in the aggregate. In particular, a series of wins by favourites (longshots) induced bettors to view less (more) favourably a bet on a favourite, which in turn produced underbetting (overbetting) of favourites.

Terrell and Farmer (1996) calculated the return to a strategy of betting the greyhound in the starting trap occupied by the previous winner. This yielded a 9 per cent profit, and as such constituted "... the only strategy earning positive profits" (p.864). The finding is consistent with the hypothesis that bettors were under-estimating the probability of a repeated outcome and that as such, they were victims of a gambler's fallacy. If confirmed, such trends or patterns could be exploited, at least in principle and perhaps in practice, to yield above-average or abnormal returns. If such configurations can be shown to exist in racetrack betting markets at an appropriate level of confidence and to constitute more than fair compensation for other factors, such as changes in the incidence of risk, then this constitutes potential evidence in contradiction of the existence of informationally efficient betting markets.

The new data set of British horse races collected and collated for this project is here employed to test for the existence of just such a gambler's fallacy.

For this purpose the data set for the period from Thursday, March 19, 1992 to Saturday May 16, 1992 inclusive is employed, i.e. period 1.

Four tables of results will be presented, each table presenting calculations of the expected return to level stakes bets placed on horses in specifically delineated races.

Table 7.24 presents the results for all races in the data set which are preceded by a race won by the favourite or joint favourite. This refers only to races occurring at the same meeting on the same day, and not to races which, although separated in time and place, are listed sequentially in the data set. The table presents the total number of observations in the data set, the total number of winners, the average return, the expected return, and the standard deviation of the actual returns. These values are presented for the aggregate of races, and separately at each of three odds classifications, i.e. starting prices less than 3.5 to 1, more than 3.5 and less than or equal to 8, and greater than 8 to 1.

Table 7.25 presents the same information, on the same basis, except that this table refers only to races preceded by TWO or more successive winning favourites/ joint-favourites.

Table 7.26 presents the same information, on the same basis, except that this table refer only to races preceded by two or more successive winning 'longshots' (defined for these purposes as horses NOT in the first TWO in the order of favouritism).

Table 7.27 again presents the same information, on the same basis, except that this table refer only to races preceded by two or more successive winning 'longshots' (defined NOW as horses NOT in the first FOUR in the order of favouritism).

Table 7.24 Expected return to level stakes, in various odds categories - for all races in Period 1 preceded by a race won by the favourite or joint-favourite

Starting price	≤3.5	> 3.5 ≤8	> 8	All
Field size	169	353	878	1400
Winners	53	40	29	122
Actual return	168.5	268.5	502	936
Expected return (%)	99.70	76.06	57.18	66.86
Standard deviation (Actual Returns)	1.54	2.17	3.45	2.99

Table 7.25 Expected return to level stakes, in various odds categories - for all races in Period 1 preceded by two or more successive races won by the favourite or joint-favourite

Starting price	≤3.5	> 3.5 ≤8	> 8	All
Field size	37	94	184	315
Winners	14	12	3	29
Actual return	39.11	81.5	45	165.61
Expected return (%)	105.70	86.70	24.46	52.57
Standard deviation (Actual Returns)	1.47	2.29	1.98	2.00

Table 7.26 Expected return to level stakes, in various odds categories - for all races in Period 1 preceded by two or more successive races won by a horse NOT in the first two in the order of favouritism

Starting price	≤3.5	> 3.5 ≤8	> 8	All
Field size	119	206	604	929
Winners	40	26	18	84
Actual return	117.22	186.5	296	599.72
Expected return (%)	98.50	90.53	49.01	64.56
Standard deviation (Actual Returns)	1.50	2.43	3.08	2.80

Table 7.27 Expected return to level stakes, in various odds categories - for all races in Period 1 preceded by two or more successive races won by a horse NOT in the first four in the order of favouritism

Starting price	≤3.5	> 3.5 ≤8	> 8	All
Field size	36	72	231	339
Winners	12	11	6	29
Actual return	31.71	84.5	96	212.21
Expected return (%)	88.08	117.36	41.56	62.60
Standard deviation (Actual Returns)	1.37	2.80	2.71	2.71

In all four samples the lowest expected return to a level stake was to bets placed in the highest odds category (greater than 8 to 1). The expected return was generally highest in the lowest odds category (less than or equal to 3.5), although in the case of the sample of races in Table 7.27, the middle odds category yielded a higher return than the lowest.

The object of the following analysis is to assess whether the return in the aggregate or in any of the odds categories is significantly greater in one particular sample of these races than another, i.e. whether the return is influenced by the immediate past history of winners expressed solely in terms of rank order of favouritism.

Table 7.28 presents the results of such comparisons for the two largest samples, i.e. tables 7.24 and 7.26. Table 7.24 is the category of all races preceded by a race won by the favourite or joint-favourite. Table 7.26 is the category of all races preceded by two or more successive races won by a horse not in the first two in the order of favouritism. Of the four categories (tables 7.24, 7.25, 7.26, 7.27), these two categories exhibit the weakest display of prior success by favourites and 'longshots' respectively.

Table 7.29 presents the results of such comparisons for the two smallest samples, i.e. tables 7.25 and 7.27. Table 7.25 is the category of all races preceded by two or more successive races won by the favourite or joint-favourite. Table 7.27 is the category of all races preceded by two or more successive races won by a horse not in the first four in the order of favouritism. Of the four categories, these exhibit the strongest display of prior success by favourites and 'longshots' respectively.

Table 7.28 Testing the significance of differences in the expected return to level stakes, for races preceded by a winning favourite/joint-favourite with that for races preceded by two or more successive winning 'longshots' - in various odds categories.

a. **For starting prices$\leq$3.5**

In Table 7.24:

Field size= 169; expected return (W1) = 99.70% (0.9970)

standard deviation (actual returns) = 1.54

In Table 7.26:

Field size= 119; expected return (W2) = 98.50 (0.9850)

standard deviation (actual returns) = 1.50

t-statistic for the difference between W1 and W2 = 0.07

It is not possible to confirm any difference in W1 and W2 at conventional levels of significance.

b. **For 3.5<starting price≤8**

In Table 7.24:

Field size= 353; expected return (W3) = 76.06% (0.7606)

standard deviation (actual returns) = 2.17

In Table 7.26 :

Field size= 206; expected return (W4) = 90.53% (0.9053)

standard deviation (actual returns) = 1.50

t-statistic for the difference between W3 and W4 = 0.71

It is not possible to confirm any difference in W3 and W4 at conventional levels of significance.

c. **For starting prices>8**

In Table 7.24:

Field size= 878; expected return (W5) = 57.18% (0.5718)

standard deviation (actual returns) = 3.45

In Table 7.26:

Field size= 604; expected return (W6) = 49.01% (0.4901)

standard deviation (actual returns) = 3.08

t-statistic for the difference between W5 and W6 = 0.48

It is not possible to confirm any difference in W5 and W6 at conventional levels of significance.

d. **For All odds.**

In Table 7.24:

Field size= 1400; expected return (W7) = 66.86% (0.6686)

standard deviation (actual returns) = 2.99

In Table 7.26:

Field size= 929; expected return (W8) = 64.56% (0.6456)

standard deviation (actual returns) = 2.80

t-statistic for the difference between W7 and W8 = 0.19

It is not possible to confirm any difference in W7 and W8 at conventional levels of significance.

Table 7.29 Testing the significance of differences in the expected return to level stakes, for races preceded by TWO or more successive winning favourites/joint-favourites with that for races preceded by TWO or more successive winning 'longshots.'

a. **For starting prices≤3.5.**

In Table 7.25:

Field size= 37; expected return (W9) = 105.70% (1.0570)

standard deviation (actual returns) = 1.47

In Table 7.27:

Field size= 36; expected return (W10) = 88.08% (0.8808)

standard deviation (actual returns) = 1.37

t-statistic for the difference between W9 and W10 = 0.53

It is not possible to confirm any difference in W9 and W10 at conventional levels of significance.

b. **For 3.5<SP≤8.**

In Table 7.25:

Field size= 94; expected return (W11) = 86.70% (0.8670)

standard deviation (actual returns) = 2.29

In Table 7.27:

Field size= 72; expected return (W12) = 117.36% (0.9053)

standard deviation (actual returns) = 2.80

t-statistic for the difference between W11 and W12 = 0.76

It is not possible to confirm any difference in W11 and W12 at conventional levels of significance.

c. **For SP>8.**

In Table 7.25:

Field size= 184; expected return (W13) = 24.46% (0.2446)

standard deviation (actual returns) = 1.98

In Table 7.27:

Field size= 231; expected return (W14) = 41.56% (0.4156)

standard deviation (actual returns) = 2.71

t-statistic for the difference between W13 and W14 = 0.74

It is not possible to confirm any difference in W13 and W14 at conventional levels of significance.

d.**All odds.**

In Table 7.25:

Field size= 315; expected return (W13) = 52.57% (0.5257)

standard deviation (actual returns) = 2.00

In Table 7.27:

Field size= 339; expected return (W14) = 62.60% (0.6260)

standard deviation (actual returns) = 2.71

t-statistic for the difference between W13 and W14 = 0.54

It is not possible to confirm any difference in W13 and W14 at conventional levels of significance.

7.3.1 Testing for a 'gambler's fallacy': Summary and conclusions

The expected return to bets placed at level stakes in races which follow a winning favourite was not different in this sample at conventional levels of significance to that on such bets placed following two or more successive winning 'longshots' (defined as horses ranked lower than SECOND in the order of favouritism). The same conclusion was reached when comparing races following two or more successive winning favourites with those following two or more successive 'longshots' (this time defined as horses ranked lower than FOURTH in the order of favouritism). Thus, on the basis of this evidence, there is no reason to accept a hypothesis that bettors alter their betting patterns in response to the rank order of favouritism displayed by the winners of preceding races. The hypothesis of a gambler's fallacy, or a winner's blessing/loser's curse is not confirmed, and cannot be employed in this sample as a benchmark to reject a postulate of semi-strong information efficiency in recent British racetrack betting markets.

7.4 Testing for a day-of-the-week 'calendar effect'

The apparent existence of a calendar effect in stock returns has been the subject of great attention in financial market analysis. Some studies, such as Bonin and Moses (1974), Tinic and West (1984), appear to suggest that stocks perform better at particular times of the year, while others, such as Gibbons and Hess (1981), Rogalski (1984) report a 'day-of-the-week' effect. Other writers report a 'turn-of-the-month' effect (e.g. Ariel, 1987), a 'weekend effect,' (e.g. Cross, 1973, Jaffe and Westerfield, 1985, Harris, 1986), and a 'holiday effect' (e.g. Lakonishok and Smidt, 1988, Wilson and Jones, 1993, Liano and White, 1994).

These effects are all adduced as prima facie evidence of semi-strong inefficiency in the markets of which they are characteristic.

The purpose of this section is to contribute to the debate by investigating the possible existence of a day-of-the-week effect in racetrack betting markets. There is already some evidence that stake size varies in different days of the week and at different times of day. In particular, Filby and Harvey (1988) report a larger average stake size on a Saturday, while Johnson and Bruce (1993) show that late race bettors (off-course at least) tend to bet a higher mean stake, back shorter-priced horses, and secure a higher expected return (and proportion of winning bets) than bettors on earlier races.

It is proposed in this chapter to offer a contribution to this issue by comparing the expected returns in two distinct odds groupings for individual days of the week. The reason for examining this particular potential calendar effect is the possibility it offers to explain in part the size of any favourite-longshot bias in terms of stake size. The idea behind such a formulation is that those betting small sums of money may be more willing to take a risk of losing the stake in order to obtain a positive, albeit small possibility of earning a large return. Those placing large bets on the other hand may well be more cautious, preferring the high probability of earning a small amount to the high probability of losing all their stake in pursuit of a low-probability large return. For example, if this notion is correct, bets placed on Saturdays (i.e. a lower proportion of small bets) might well contain an above-average proportion of favourites in them, and a below-average proportion of longshots. If so, Saturday bets would exhibit a smaller longshot bias than bets placed on other days.

Table 7.30 presents results for each day of the week, in period 1 (from Thursday, March 19, 1992 to Saturday May 16, 1992 inclusive) for the following variables: the number of observations; the number of winners; the average return, the expected return, and the standard deviation of actual returns in each of two odds categories.

The results are presented for each day of the week separately, from Monday to Saturday. There was no Sunday racing in the season in which the data set was collected.

Table 7.30 **A comparison of the expected return to level stakes, at starting prices < 8 and ≥8, on each day of the week, for all races in Period 1**

Day of the Week	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Field size (SP<8)	417	144	209	300	285	394
Winners (SP<8)	80	25	48	61	61	74
Actual Return (SP<8)	341.91	109.24	194.27	253.91	276.93	315.18
Expected Return (%), SP<8	81.99	75.86	92.95	84.64	97.17	80.00
Standard Deviation, Actual Returns, SP<8	1.84	2.24	1.88	1.94	2.06	1.86
Field size (SP≥ 8)	1064	370	539	643	671	812
Winners (SP≥ 8)	39	16	18	25	25	34
Actual Return (SP≥ 8)	650	255	252	367	403	477
Expected Return (%), SP≥ 8	61.09	68.92	46.75	57.08	60.06	58.74
Standard Deviation, Actual Returns, SP≥ 8	8.34	5.75	7.16	8.44	10.07	5.28

SP = Starting Price

No simple strategy based on betting systematically in either of these odds categories on a particular day of the week was capable of

No simple strategy based on betting systematically in either of these odds categories on a particular day of the week was capable of yielding a profit. Although there is no clear possibility for earning abnormal returns it is possible, however, that a day-of-the-week strategy might produce above-average returns compared to a random strategy.

To test this hypothesis, an analysis is undertaken to ascertain whether there exist any significant difference in the expected return in either of the odds categories chosen, as between particular days of the week.

To do this, tests of significance are applied to the largest differential in the expected returns at a given odds classification between particular days of the week.

The largest expected return in the odds category $SP<8$ is recorded on Fridays, i.e. 97.17 per cent, and the lowest on Tuesdays, i.e. 75.86 per cent. The results are presented in table 7.31.

The largest expected return in the odds category $SP\geq 8$ is recorded on Tuesdays, i.e. 68.92 per cent and the lowest on Wednesdays, i.e. 46.75 per cent. The results are presented in table 7.32.

Table 7.31 Testing the significance of differences in the expected return to level stakes, at starting prices<8, for all races on a Friday in period 1, with that for all races on a Tuesday - in period 1.

a. **For SP<8.**

On Fridays:

Field size= 285; expected return = 97.17%

standard deviation (actual returns) = 2.06

On Tuesdays:

Field size= 144; expected return = 75.86%

standard deviation (actual returns) = 2.24

t-statistic for the difference between 97.17% and 75.86%= 0.96

It is not possible to confirm any difference in the Friday and Tuesday expected returns in the odds classification SP≥8.

Table 7.32 Testing the significance of differences in the expected return to level stakes, at starting prices≥8, for all races on a Tuesday, with that for all races on a

Wednesday - in period 1.

a. **<u>For SP>=8.</u>**

On Tuesdays:

Field size= 370; expected return = 68.92%

standard deviation (actual returns) = 5.75

On Wednesdays:

Field size= 539; expected return = 46.75%

standard deviation (actual returns) = 7.16

t-statistic for the difference between 68.92% and 46.75%= 0.52

It is not possible to confirm any difference in the Tuesday and Wednesday expected returns in the odds classification SP>=8.

7.4.1 Testing for a day-of-the-week 'calendar effect': Summary and conclusions

No significant day-of-the-week effect has been identified in the sample of races taken from the 1992 British flat racing season. In particular, no pattern has been identified which would allow a bettor to secure profits by betting on a particular day on horses in particular odds groupings, nor was there was evidence that significant above-average returns could be made by betting on certain days of the week rather than others. Even so, an examination of variations in stake size at a more sub-aggregated level may well offer a useful avenue for future research.

7.5 Does favouritism matter?

Based on U.S. racetrack betting markets, Asch, Malkiel and Quandt (1982) found a close relationship between a horse's place in the order of favouritism and the likelihood of it winning. Furthermore, they reported evidence in their data set that whereas the objective probability of a horse winning was significantly greater (at the 5 per cent level of confidence) than the subjective probability for favourites, the bias was gradually reversed until the subjective probability was significantly greater than the objective probability for ninth place in the order of favouritism. They also found that the rates of return for bets at lower odds tended to be greater than those at higher odds. While these findings are consistent, no previous study (including that by Asch, Malkiel and Quandt) has sought to establish whether horses starting at given odds offer a different expected return depending on their rank order of favouritism. In particular, does favouritism in itself contribute any independent effect on expected returns? Do 4 to 1 favourites, for example, provide a different expected return from 4 to 1 non-favourites? Henery (1985) clearly implies that they do not:

"Firstly, and this advice goes back to Figgis (1951, p.112), do not bet on a horse simply because it is the favourite - much better to bet on horses which have very low odds. The distinction is that horses with very short SP odds are necessarily favourites also, but not vice-versa..." (p.345). However, he offers no evidence beyond an empirical link between higher odds and lower expected returns, and a theoretical explanation for same. It is the purpose of this chapter to explore this issue.

In order to standardize for the standard favourite-longshot bias (really a size of odds bias), the expected return is calculated in period 1 for the set of races grouped into a defined odds category, i.e. SP≤3.5. This is selected because it offers a reasonable sample size, and yet includes horses at three rank orders of favouritism. The expected return in this category is specified by the order of favouritism associated with each horse returning a given starting price. For example, the expected return for all favourites returning a starting price in this odds category is compared with that of second favourites which were returned in the same odds category.

Where two or more horses share a given starting price the mean position of these horses in the rank order of favouritism is taken. In a three horse race, for example, characterized by starting prices of 0.909, 1.5 and 1.5, the joint second favourites are accorded a rank order of (2+3)/2, i.e. 2.5. This is recorded in these tables as favourite = 2.5.

There are no observations in the data set with a starting price of less than or equal to 3.5 and a rank order of favouritism greater than 3.

Table 7.33 presents the results.

Table 7.33 Expected return to level stakes, at various rank orders of favouritism - for starting prices ≤ 3.5

Rank order of favouritism	**1**	**1.5**	**2**	**2.5**	**3**	**All**	**> 1**
Field size	392	54	199	20	36	701	309
Winners	146	10	52	5	8	220	74
Actual Return	376.06	38.21	192.25	21	29.42	656.94	280.87
Expected Return (%)	95.93	70.76	95.10	105	81.72	93.71	90.90
Standard Deviation (Actual Returns)	1.34	1.51	1.65	1.83	1.67	1.48	1.65

The expected return for clear first favourites (95.93 per cent) is greater than the expected return for clear second favourites (95.10 per cent), although not at any conventional level of significance (t=0.06). An overview of the expected returns in each of the rank orderings also provides no clear pattern relating expected return to rank order of favouritism. In table 7.34 the expected return (favourite = 1) is compared with the expected return (favourite > 1).

Table 7.34 Testing for the significance of differences in the expected return to level stakes, for clear favourites (favourite=1) and all other cases (favourite>1), for starting prices≤3.5.

a. **<u>For SP≤3.5 and favourite=1.</u>**

Field size= 392; expected return= 95.93%;

standard deviation (actual returns) = 1.34.

b. **<u>For SP≤3.5 and favourite>1.</u>**

Field size= 309; expected return= 90.90%;

standard deviation (actual returns) = 1.65.

The expected return (95.93%) is greater for clear favourites than for the rest of the observations (expected return = 90.90%), although not at any conventional level of significance (t= 0.43).

7.5.1 Does favouritism matter? : Summary and conclusions.

The expected return in the chosen data set, of horses starting at prices of 3.5 to 1 and less, varies from 105 per cent for horses starting as joint second favourite or equivalent (favourite = 2.5) to 70.76 per cent for horses starting as joint first favourites or equivalent (favourite= 1.5). However, both these samples contained few observations (20 and 54 respectively), and overall no consistent pattern is readily observable between order of favouritism and expected return, within this restricted odds grouping. A test on orderings characterized by sample sizes of 100 or more, i.e. clear favourites and clear second favourites reveals no difference at conventional levels of significance. A comparison was also made of the expected return to bets on clear first favourites with that to all other horses in this low odds category. Although the favourites did yield a higher expected return than the other horses, the difference was not significant at any conventional level.

7.6 Comparing expected return with the 'tote' and with bookmakers to test for semi-strong form efficiency

In this section, the expected return to bets 'on the tote' and with bookmakers about identical events are examined and compared with respect to the new data set compiled about the 1992 British racetrack betting market. Implications are drawn for the existence of semi-strong information efficiency in these markets.

7.6.1 Employing Tote/Starting Price differentials to test for semi-strong efficiency

Gabriel and Marsden (1990, 1991) tested for the existence of semi-strong form efficiency in the British racetrack betting market by comparing the returns to two types of betting, i.e. starting price betting made with bookmakers and bets laid with the Totalizator Board. Their data was drawn from the first 1427 flat races of the 1978 racing season in England, the year being chosen because the general absence of mechanical or electronic tote boards in that year limited the information available to bettors on betting patterns or likely final odds. The idea behind this is that in order to equalize the risk of betting as between the two betting mediums, bettors should be equally uncertain under either betting system of the exact odds until after the race starts.

"This placed the tote bettors in the position of having limited, if any, information on betting patterns or likely final odds for their betting choice." (Gabriel and Marsden, p.554).

In these circumstances, they argue, the two types of bet are equivalent options about the same item, and so if the market is efficient with respect to publicly available information (semi-strong form efficient), the odds in both markets should converge. Thus they compare the mean tote and starting price payoffs after races.

In this section, the mean tote and starting price payoffs for the new data set are compared at an aggregated level and at various levels of sub-aggregation, in order to examine whether any systematic differences can be identified, and if so whether these are systematically exploitable. Unlike in Gabriel and Marsden's study, bettors in these markets are served by electronic tote boards, which constantly update the pool payouts. Even so, these figures can only offer a rough indication of the post-race payouts to the different horses, and this uncertainty applies also to bets placed at starting price. Bets placed on course with bookmakers, however, are usually at agreed odds. Were it possible to arbitrage between this and the final tote payout, this would be of genuine value to these bettors.

Since a complete record of starting price and tote dividends is available in this data set for the period from Thursday, March 19, 1992 to Saturday May 16, 1992 inclusive, it is this period which is analyzed. In all 510 races are recorded, with 509 containing data on the differences between the starting price and tote dividend.[5] Since the tote dividend is declared inclusive of a unit stake, the tote odds are examined net of this, i.e. by subtracting 1 from the published dividend.

Table 7.35 compares the mean payout to a level stake on winning horses at tote prices and at starting prices, using a t-test for paired data and a Wilcoxon matched-pairs signed-rank test. Subsequent tables compare these returns at various sub-groupings of odds and other levels of sub-aggregation, using the same test.

Table 7.36 presents comparisons of the mean payout to a level stake on winning horses where the race winner was returned at tote odds of less than or equal to 8 to 1.

Table 7.37 presents comparisons of the mean payout to a level stake on winning horses where the race winner was returned at tote odds of greater than 8 to 1.

Table 7.38 presents comparisons of the mean payout to a level stake on winning horses where the race winner was returned at a starting price of less than or equal to 8 to 1.

Table 7.39 presents comparisons of the mean payout to a level stake on winning horses where the race winner was returned at a starting price of greater than 8 to 1.

All payouts are calculated exclusive of the stake.

Following Gabriel and Marsden (1990), the results of Wilcoxon matched-pairs signed-rank tests are reported, since this test makes no underlying assumptions about the shape of the distributions from which the data was drawn.

Table 7.35 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all races in period 1.

For All races.

Observations= 509

Mean tote payout= 7.86; standard deviation = 10.65;

Mean SP payout = 6.74; standard deviation = 6.93.

The null hypothesis of no difference is rejected at all conventional levels of significance (t=4.19).

Wilcoxon matched-pairs signed-ranks test: z = -1.82 (pr>|z|=0.0690), indicating rejection of the null hypothesis of no difference at the ten per cent level of significance.

Table 7.36 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all races in period 1 where win tote odds ≤8.

a. **For Tote ≤8.**

Observations=354

Mean tote payout = 3.19; standard deviation = 2.01;

Mean SP payout = 3.54; standard deviation = 2.51.

The null hypothesis of no difference is rejected at all conventional levels of significance (t=4.87).

Wilcoxon matched-pairs signed-ranks test: z = 4.87 (indicating rejection of the null hypothesis of no difference at all conventional levels of significance).

Table 7.37 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all races in period 1 where win tote odds >8.

<u>For Tote>8.</u>

Observations=155

Mean tote payout = 18.54; standard deviation = 14.13;

Mean SP payout = 14.05; standard deviation = 8.16.

The null hypothesis of no difference is rejected at all conventional levels of significance (t=5.59).

Wilcoxon matched-pairs signed-ranks test: z = 6.58 (indicating rejection of the null hypothesis of no difference at all conventional levels of significance).

Table 7.38 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all races in period 1 where starting prices ≤8.

For SP≤8.

Observations=371

Mean tote payout = 3.65; standard deviation = 2.76;

Mean SP payout = 3.55; standard deviation = 2.20.

The null hypothesis of no difference cannot be rejected at any conventional level of significance t = 1.46).

Wilcoxon matched-pairs signed-ranks test: z = 0.87 (indicating that the null hypothesis of no difference cannot be rejected at any conventional level of significance).

Table 7.39 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all races in period 1 where starting prices >8.

For SP>8.

Observations=138

Mean tote payout = 19.18; standard deviation = 14.93;

Mean SP payout = 15.31; standard deviation = 7.95.

The null hypothesis of no difference is rejected at all conventional levels of significance (t=4.14).

Wilcoxon matched-pairs signed-ranks test: z= 3.96 (indicating that the null hypothesis of no difference is rejected at all conventional levels of significance).

In summary, for the set of all races, using t-tests for paired data, the null hypothesis of no difference between the tote and starting price payouts about identical winning horses is rejected at all conventional levels of significance. The difference favours the tote. It is, however, only rejected at the ten per cent level, using the Wilcoxon matched-pairs signed rank test suggested by Gabriel and Marsden (1990). At tote odds of less than or equal to 8 to 1, the starting price payout is greater than the tote payout, at all levels of significance. The null hypothesis of no difference cannot be rejected, however, at any conventional level of significance, for the set of all horses with starting prices of less than or equal to 8 to 1. The results would indicate that the bias in favour of the tote is greater for horses less favoured in the market. At lower odds the advantage tends to disappear, and even to turn in favour of the starting price.

Attempts to implement a strategy based on observing likely tote returns are limited, however, by the difficulty of identifying the final return before the off. In principle, if bettors can identify horses likely to return low tote dividends it is generally better to bet with bookmakers. The observed tote superiority, particularly at higher odds, is common to this study and that of Gabriel and Marsden (1990, 1991). An explanation may lie in the reluctance of bookmakers to set large odds, even on horses whose estimated objective probability of winning is low, because of an adverse selection problem facing them in the context of insiders who know more than them, or who can even (in Shin's models, 1991, 1992, 1993) "observe" the outcome with certainty. In such an environment, bookmakers may 'squeeze' prices, particularly about a 'longshot,' even when no bets have been placed on it, because the 'insider' may pounce at any time. The 'insider' may, however, still prefer to use bookmakers to the tote. Schnytzer and Shilony (1995) offer one reason,

"Given that with bookmakers, the payout contingent on a win is known when the bet is placed and that inside information is likely to be more accurate as race time draws near, we should expect most 'insiders' to bet with bookmakers." (p.964).

If this is what is happening, a test which distinguishes the data by the likely presence of insider trading may be useful.

The data sample is thus sub-divided into a sample composed of races likely to be characterized by the absence (or a low incidence) of useful private information. If the observed superiority of tote returns at high odds is caused by the threat of insider trading, then the differential should be less (or disappear even) in samples characterized by low levels of insider activity.

In order to distinguish races by the incidence of insider trading, two methods are adopted. In the first, Crafts' (1985) handicap/non-handicap split is employed. In the second, a similar distinction, which allows for the fact that insiders may be able to use public information to improve their private information is employed, as in Vaughan Williams and Paton (1997a). Crafts (1985) suggested separating handicap races (where horses are allocated weights so as to equalize as far as possible their chances of winning) from non-handicaps. The idea behind this is that the past form of horses in handicaps is generally more established in the public forum than it is in non-handicaps. In the latter, therefore, there are likely to be very limited possibilities for betting on the basis of privately held information.

It is possible that insiders may use public information to improve their private information. For this reason it may be preferable to consider only higher grade handicaps as indicative of the absence of useful private information. The reason is that these race types (excluding both non-handicaps and handicaps rated below 100[6]) are the subject of particular media attention and might be expected to offer very little opportunity for non-disclosure of useful private information. It is assumed that the informational content of any private information available about these race types will be close to zero.

The results will contribute to the evidence about an adverse selection explanation of (at least part of) the favourite-longshot bias.

Standard t-tests of paired data are employed to test for significance differences in the two samples, as well as a Wilcoxon matched-pairs signed-rank test.

Table 7.40 compares the mean payout to a level stake on winning horses at tote prices and at starting prices, in handicap races, using a t-test for paired data and a Wilcoxon matched-pairs signed-rank test.

Table 7.41 presents comparisons of the mean payout to a level stake on winning horses at tote and at starting prices, in non-handicaps, using a t-test for paired data and a Wilcoxon matched-pairs signed-rank test.

Table 7.42 presents comparisons of the mean payout to a level stake on winning horses at tote and at starting prices, in higher grade handicaps, using a t-test for paired data and a Wilcoxon matched-pairs signed signed-rank test.

Table 7.43 presents comparisons of the mean payout to a level stake on winning horses at tote and at starting prices, in non-higher grade handicaps, using a t-test for paired data and a Wilcoxon matched-pairs signed-rank test.

All payouts are calculated exclusive of the stake.

Table 7.40 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all handicap races in period 1.

Handicaps.

Observations= 224

Mean tote payout= 9.73; standard deviation = 11.38;

Mean SP payout = 8.44; standard deviation = 7.85.

The null hypothesis of no difference is rejected at all conventional levels of significance (t=2.97).

Wilcoxon matched-pairs signed-ranks test: z= 1.34 (indicating that the null hypothesis cannot be rejected at any conventional level of significance).

Table 7.41 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all non-handicap races in period 1.

Non-Handicaps.

Observations= 285

Mean tote payout= 6.39; standard deviation = 9.81;

Mean SP payout = 5.40; standard deviation = 5.77.

The null hypothesis of no difference is rejected at all conventional levels of significance (t=2.95).

Wilcoxon matched-pairs signed-ranks test: z= 1.34 (indicating that the null hypothesis cannot be rejected at any conventional level of significance).

Table 7.42 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all 'higher grade'

handicap races in period 1.

<u>**Higher grade handicaps.**</u>

Observations= 64

Mean tote payout= 9.62; standard deviation = 12.49;

Mean SP payout = 8.36; standard deviation = 7.66.

The null hypothesis of no difference cannot be rejected at the ten per cent level of significance (t=1.67).

Wilcoxon matched-pairs signed-ranks test: z = 0.75 (indicating that the null hypothesis cannot be rejected at any conventional level of significance.

Table 7.43 Testing the significance of differences in the mean payout to a level stake on winning horses at tote prices and at starting prices - for all non-'higher grade' handicap races in period 1.

<u>Non-Higher grade handicaps</u>

Observations= 445

Mean tote payout= 7.61; standard deviation = 10.35.

Mean SP payout = 6.50; standard deviation = 6.79.

The null hypothesis of no difference is rejected at all conventional levels of significance (t=3.84).

Wilcoxon matched-pairs signed-ranks test: z = 1.72, indicating that the null hypothesis can be rejected at the ten per cent level of significance).

7.6.2 Employing Tote/Starting Price differentials to test for semi-strong efficiency: Summary and conclusions

Using a t-test for paired data, it was possible to reject a null hypothesis of no difference in tote and starting price payouts about identical winning horses in all races except higher-grade handicaps. The Wilcoxon matched-pairs signed-rank test was unable to distinguish these samples as clearly, the test indicating rejection of this null hypothesis only for non-higher grade handicaps. However, the z-value for higher grade handicaps was markedly lower than for any of the other samples.

There is some evidence in these results, therefore, that the differential between tote and starting price odds is less (or disappears) in samples which might be expected to contain low levels of insider activity, i.e. higher grade handicaps. Without stating the case too strongly, it may be reasonable to infer from this that the findings are at least indicative of some effect on starting prices resulting from bookmakers' response to potential insiders, and that this effect is not fully reflected in tote prices. Whatever conclusions one draws from this evidence, the findings reported in this section taken in aggregate are in any case consistent with the existence of semi-strong form inefficiency in these markets.

7.6.3 Employing Dual Forecast/Computer Straight Forecast differentials to test for semi-strong efficiency

The Dual Forecast bet (DF) operated by the tote requires bettors to forecast the first two finishers of a race in any order. The bookmaker-run Computer Straight Forecast (CSF) requires bettors to forecast the first two finishers of a race in the correct order. Since these two types of bet are comparable options about the same race it might be argued, along the same lines as Gabriel and Marsden (1990), that if the market is efficient with respect to publicly available information (semi-strong form efficient) the effective odds about a Dual Forecast and Computer Straight Forecast should, after adjusting for the difference in their requirements, tend to converge. Deductions from both pools are greater than for the win tote pool, and the CSF suffers from the added disadvantage that there is no ready public accessibility to the formula by which the payouts are calculated. In both cases, therefore, bettors have to take the ultimate payout on trust. The implications for equity and fairness of the higher deductions in these pools are not directly addressed here, but are treated at length in Dowie (1992a, 1992b), and the lack of ready access to the CSF formula is considered in Dowie, Coton and Miers (1991).

The Dual Forecast bet may be viewed as equivalent to two Computer Straight Forecast bets, one on two horses selected to finish first and second respectively, and the other on the same two horses to finish second and first respectively. If both horses have an equal chance of beating each other to the post, then a fixed £10 return to a £1 bet on a CSF has only half the chance of success as a fixed £10 return to a £1 bet on a Dual Forecast. In this sense, the return offered about a CSF should be twice that offered about an equivalent DF. However, if the probability of one of the two horses winning is significantly greater than the other, there may be an advantage in choosing the CSF even when the returns to a unit stake are less than twice as great as those to a DF bet. For example, say one horse has a zero probability of beating the other. In this case, a fixed £10 return to a £1 bet on a CSF has the same chance of success as a fixed £10 return to a £1 bet on a DF.

In practice, when the bets are struck, it is not possible to know with certainty the eventual return. The DF payout can and usually does fluctuate right up to the off, and the

CSF payout is calculated according to an esoteric formula which takes into account the ultimate starting prices and the number of runners. For simplicity, it is assumed here that the risk associated with both bets (in this form) is roughly equal.

In comparing the returns, it should be noted that the Computer Straight Forecast is an off-course bet, and the tax is already included in the payout. Tote dividends (of which the Dual Forecast are one) are subject off-course to betting tax, usually levied at ten per cent at the time the sample was collected. In order to compare the CSF with the Dual Forecast dividend, therefore, it would seem appropriate to deduct ten per cent from the Dual Forecast return. The calculations are also performed, however, on the gross payouts.

Since a complete record of Dual Forecast and Computer Straight Forecast dividends is available in this data set for period 1 (from Thursday, March 19, 1992 to Saturday May 16, 1992 inclusive), it is this period which is analyzed. In all 507 races are recorded and examined.

Table 7.44 presents comparisons of the payout to a unit stake (excluding the stake) from successful Dual Forecast bets with the corresponding Computer Straight Forecast bets. The Dual Forecast payouts are presented gross and net of a ten per cent deduction. Since the Dual Forecast is equivalent to two Computer Straight Forecasts, one on a given 1-2 and the other on the reverse, i.e. 2-1, the CSF returns are also presented to a half stake for purposes of comparability.

In each case t-tests for paired data, and the Wilcoxon matched-pairs signed-rank test are employed.

Table 7.44 Testing the significance of differences in the mean payout to a level stake on successful bets at Dual Forecast and Computer Straight Forecast odds.

Number of observations = 507

a. **Dual Forecast.**

Mean dual forecast payout = 52.65

Standard deviation = 91.61

b. **Computer Straight Forecast.**

Mean CSF payout = 57.18

Standard deviation = 75.72

c. **0.9 x Dual Forecast.**

Mean dual forecast payout = 47.39

Standard deviation = 82.45

d. **CSF / 2.**

Mean payout = 28.59

Standard deviation = 37.86

t-statistic for the null hypothesis of no difference between CSF/2 and 0.9 x Dual Forecast = 6.89

Wilcoxon matched-pairs signed-rank test: z= 3.21

Applying the t-test for paired data and the Wilcoxon matched-pairs rank-sum test, therefore, the null hypothesis of no difference between the Dual Forecast payout (gross or net) and the (adjusted) payout to a Computer Straight Forecast bet is rejected at all conventional levels of significance. The Dual Forecast offers a significantly superior return, even net of the tax paid on such bets off-course.

In order to test whether the superior return to a Dual Forecast bet can be explained by differences in the probability of correctly forecasting horses to finish 1-2 rather than 2-1, the sample is now restricted to those races in which the starting price of the second placed horse was less than that of the winner, i.e. the second placed horse was expected before the race in the bookmaker/bettor market to beat the actual winner. In these cases there is a less clear advantage in prior specification of the precise order of finishing, and here the Dual Forecast bet is more clearly established as similar to two Computer Straight Forecast bets about a given race.

The results are presented in Table 7.45, in the same form as for Table 7.44.

Table 7.45 Testing the significance of differences in the mean payout to a level stake on successful bets at Dual Forecast and Computer Straight Forecast odds - for races in which the starting price of the second placed horse was less than that of the winner.

Number of observations=247

a. **Dual Forecast.**

Mean dual forecast payout = 58.63

Standard deviation = 91.37

b. **Computer Straight Forecast.**

Mean CSF payout = 74.15

Standard deviation = 89.66

c. **0.9 x Dual Forecast.**

Mean dual forecast payout = 52.77

Standard deviation = 82.23

d. **CSF / 2.**

Mean payout = 37.08

Standard deviation = 44.82

t-statistic for the null hypothesis of no difference between CSF/2 and 0.9 x Dual Forecast = 4.54

Wilcoxon matched-pairs signed-rank test: z= 0.24 (pr>z|=0.8074).

Applying the t-test for paired data the null hypothesis of no difference between the net Dual Forecast payout and the (adjusted) payout to a Computer Straight Forecast bet is rejected at all conventional levels of significance. The Wilcoxon matched-pairs rank-sum test, however, fails to reject the null hypothesis at even an eighty per cent level of significance.

7.6.4 Employing Dual Forecast/Computer Straight Forecast differentials to test for semi-strong efficiency: Summary and conclusions

The odds implicit in a tote Dual Forecast bet, i.e. a bet on two horses to finish first and second in any order in a given race, are here compared directly with the odds offered in bookmaker-run Computer Straight Forecasts of the same two horses to finish in a particular pre-specified order. An alternative comparison is made by interpreting the Dual Forecast bet as two identical bets of a unit stake on the two horses to finish 1-2 and 2-1, and so declaring the Computer Straight Forecast returns to a half unit stake. The off-course tax levied on Dual Forecast bets is deducted to give a net figure with which to compare the Computer Straight Forecast payout (which already includes the tax deduction), but is not offered by on-course bookmakers. Two distinct tests are applied to the data, a t-test for paired data, and a Wilcoxon matched-pairs signed-rank test which makes no prior assumption about the shape of the distributions from which the data are drawn. Although the t-tests indicate rejection of a null hypothesis of no difference in net Dual Forecast and adjusted Computer Straight Forecast bets, the rank-sum test offers strong evidence against rejection when the sample is restricted to races in which the starting price of the second-placed horse was less than that of the first-placed horse. This is an interesting result, because it is only in this sample of races that allowance is made for any advantage bettors may have in prior specification of the 1-2 order of finishing. This provides the fairest comparison, therefore, of the actual (post-tax) payouts from Dual Forecast and Computer Straight Forecast bets with closely comparable probabilities of success. On the basis of this evidence in isolation it is not possible, therefore, to reject the hypothesis that the betting market under examination is semi-strong form efficient.

7.7 Does information efficiency require a perception of information inefficiency?

7.7.1 Introduction

In the U.K. a parimutuel (Tote) system operates alongside a fixed-odds system, and bettors can choose to bet about identical outcomes with either. If the market is informationally efficient, it has been argued (see Gabriel and Marsden, 1990) that these prices should converge.

In particular, differences in returns at Tote and bookmakers' prices should be eliminated by the end of trading as bettors exploit any arbitrage opportunities. Evidence presented both here, in Gabriel and Marsden (1990, 1991) and in Blackburn and Peirson (1995) indicate that they are not[7]. It is suggested here that the reason this does not occur is because bettors may be unaware that any information inefficiency exists, and so fail to act in a manner which would eliminate any Tote/bookmaker odds differential.[8] If this hypothesis is correct, then the differential should be less when bettors are more likely to be aware that a significant inefficiency exists.

Schnytzer and Shilony (1995) demonstrate that bettors act rationally in response to the signal given by a marked movement in bookmakers' odds. The transmission mechanism is through the perception of insider activity given by such odds changes. These insiders are trading upon information which is not publicly available, even to bookmakers. In this environment the odds set by bookmakers are unlikely to be viewed as reliable indicators of the objective probabilities, and bettors will be encouraged to seek opportunities to correct this "inefficiency". By arbitraging between the Tote and bookmakers, the difference between these two payouts should be reduced or even disappear at all odds levels.

7.7.2 Tests of market inefficiency

For market efficiency there should be no difference in the mean payout on the Tote and the mean payout at Starting Prices to identical winning outcomes. Gabriel and Marsden (1990) adapt work by de Leeuw and McKelvey (1984) and Zuber et al. (1985) and estimate the following model:

$$TOTE_i = \alpha_0 + \alpha_1 SP_i + u_i \tag{7.1}$$

where $TOTE_i$ is the Tote payout on race i; SP_i is the Starting Price[9] (payout) in race i; u_i is an error term.

The test for market efficiency is of the joint null hypothesis that $\alpha_0 = 0$ and $\alpha_1 = 1$ (Zuber et al., 1985, pp. 800-801). Whilst this test is certainly valid, care must be taken in estimating equation (7.1). The variance of the error term is likely to be related to the Starting Price. Put simply, the variation of Tote returns around any given Starting Price will be less, in absolute terms, at low odds such as 2 to 1 than at high odds such as 20 to 1. In this case, OLS estimates of (7.1) will suffer from heteroscedastic errors. Consequently, standard errors will be biased and the test of efficiency may be unreliable, unless appropriate adjustments to the standard errors are made.[10]

7.7.3 Data and Estimation

7.7.3.1 Data

The data set comprises all the Tote dividends and the Starting Prices of the winning horses in races for the period from Thursday, March 19, 1992 to Saturday May 16, 1992 inclusive. In all 510 races are recorded, with 509 containing data on the differences between the Starting Price and Tote dividend.[11] Since the Tote dividend is declared inclusive of a unit stake, the Tote odds are examined net of this, i.e. by subtracting 1 from the published dividend.

Mean levels of Tote and SP returns are given in Table 7.46 for five grouped odds categories. On average, Tote returns seem to be marginally inferior at lower odds levels and superior at higher levels. For the whole sample, Tote returns are about 17% higher than those at Starting Prices. The next step is to carry out formal tests on this data.

7.7.3.2 Empirical Tests of Inefficiency

Equation (7.1) is estimated using OLS. Results are reported in Table 7.47, column 1. The joint null hypothesis that $\alpha_0 = 0$ and $\alpha_1 = 1$ is clearly rejected at conventional significance levels. The negative value found here for α_0 suggests that, at low odds levels, Starting Price returns are higher, evidence which is consistent with the findings of Blackburn and Peirson (1995). As the odds levels increase, Tote returns become superior. However, the Cook-Weisberg test statistic (see Goldstein, 1992) suggests that, as expected, there is a serious problem of heteroscedasticity. Thus, in column 2, we report standard errors which are robust to heteroscedasticity. The null hypothesis of market efficiency is still clearly rejected (although the F-statistic is now much lower) but the estimate of α_0 is now not significantly less than zero[12].

There is strong evidence that Tote payouts differ significantly from Starting Price payouts and that the Tote is superior at higher odds. This is exactly what might be expected in a market in which bookmakers are posed an adverse selection problem in the context of an unknown proportion of "insiders" who may possess superior information. In this environment it has been argued (see Shin, 1993; Vaughan Williams and Paton, 1996, 1997a) that the optimal response by bookmakers is to artificially constrict odds, especially at higher levels where potential losses are greatest. Tote returns are not subject to any such supply side bias and so any differential will tend to be greater at higher odds. This effect will persist until and unless bettors act so as to arbitrage it away. We argue that they will only do so if some form of informational inefficiency is clearly signalled in the market.

7.7.4 Signals of Informational Inefficiency

Following the discussion in section 7.1, it is proposed here that a reasonable proxy for the existence of a ***perception*** of inefficiency by bettors is where the bookmakers' odds have moved markedly between Opening and Starting Prices. Races in which a marked difference exists between the Opening Price and the Starting Price of the winner are therefore separated out. The benchmark chosen is where either OP/SP or SP/OP is greater than 1.4[13]. There are 130 races in which the odds move markedly and 379 others. Equation (1) is estimated for each of the two sub-samples and results are reported in columns 3 and 4 of Table 7.47.

In the case of races in which there is no marked movement in the odds, the null hypothesis of efficiency is convincingly rejected. Indeed, the estimate of α_1 is much higher than for the whole sample. In the sample in which there is a marked movement in the odds, the joint null hypothesis of market efficiency cannot be rejected at any level. In other words, Tote returns converge to SP returns.

Since the hypothesis is that it is movements in the odds which provide a signal, no distinction is made between odds shortening or lengthening. The robustness of the results to this assumption is checked by splitting into two the sample of races in which there is a marked movement in odds: races in which the price has shortened from opening prices, and those in which the price has lengthened. In each case, the joint null hypothesis of market efficiency cannot be rejected at any conventional level[14]. This provides evidence that it is the movement of odds which is crucial for arbitraging to take place.

7.7.5. Does information efficiency require a perception of information inefficiency? Conclusions

The conclusion of this analysis is that the U.K. racetrack betting market does not always behave in a manner consistent with the incorporation of all publicly available information. In particular, significant differences sometimes exist in the expected returns to bets placed about identical outcomes in different sectors of the market. This may be because bettors are not always aware that any profitable unexploited opportunities for arbitrage exist. However, where a market signal (in the form of a marked movement in bookmakers' odds) causes bettors to abandon these perceptions, they act in such a way as to eliminate the information inefficiency.

TABLE 7.46: Mean SP and TOTE Returns to a Unit Stake

SP Range	Mean SP Return	Mean Tote Return	Number of Races
≤ 5	2.30	2.27	256
≤ 10	6.66	7.30	131
≤ 15	11.49	13.77	73
≤ 20	16.00	19.24	14
20 +	32.12	44.85	17
All	6.73	7.86	509

TABLE 7.47: Tests of Market Efficiency

	1 Whole Sample	2 Whole Sample	3 Movement in odds	4 No movement in odds
SP	1.300***	1.300***	0.973***	1.524***
	(0.036)	(0.1104)	(0.172)	(0.109)
Constant	-0.896**	-0.896	0.256	-1.866***
	(0.352)	(0.5532)	(1.018)	(0.472)
N	509	509	130	379
R^2	0.7149	0.7149	0.7593	0.7481
Heteroscedasticity	1609.23***	-	355.13***	847.32***
Omitted Variables	3.35**	3.35**	4.54**	9.60**
F-Test	43.69***	11.50***	0.10	13.96***

Notes:

1. The dependent variable is Tote.

2. Figures in brackets in 1 are standard errors. Those in 2-4 are adjusted for heteroscedasticity.

4. *** indicates significance at the 1% level; ** at the 5%; * at the 10%.

5. Heteroscedasticity statistic is the Cook-Weisberg one, distributed as $\chi^2(1)$ (see Goldstein, 1992).

6. Omitted variable test statistic is based on a Ramsey Reset test of powers up to the third degree, distributed as F(3, 504).

7. The F-test is of the joint null hypothesis that $\alpha_0 = 0$ and $\alpha_1 = 1$, distributed as F(2, n-k) where k is the number of regressors.

1. An each-way bet requires two stakes, one to win and the other to place. The meaning of a "place" depends on the nature of the race and the number of runners. Generally, there is no place payout for races with less than five runners. If there are five to seven runners, the payout is a quarter the odds for first or second. If there are eight to eleven runners, the payout is one fifth the odds to the first three. In handicaps, however, the payout is a quarter the odds to the first three in races of 12 to 15 runners, and a quarter the odds to the first four home in races with more than 15 runners.

2. This has now been reduced to 9 per cent, following a cut in the standard rate of betting tax of one per cent, to 6.75 per cent, advised in the November, 1995 Budget, and implemented in 1996.

3. In one case a horse was recommended to a half-point (one unit) stake "ante-post" and to a further half-point stake at starting price. The horse won at an ante-post price of 9 to 1 and a starting price of 15 to 2. For these purposes the bet was included as two advices of a half point, thus improving the returns to such advices at the expense of the return to one-point (two unit stake) advices.

4. A longshot is defined for these purposes as a horse which does not start as the first or second favourite in the race.

5. The missing item is the first race at Newcastle on 20th. April, 1992, which failed to produce a tote dividend due to a technical fault.

6. Flat race handicaps are normally rated between 0 to 60 for the lowest grade handicap races and 0 to 115 for the highest grade.

7. Bird and McCrae (1994) discuss somewhat contrasting findings for Australian data.

8. Alternatively they may not consider the inefficiency important enough to outweigh the costs of acting upon it.

9. The Starting Price is the general price at which a sizeable bet could be placed with bookmakers on the course at the start of the race.

10. Even and Noble (1992) suggest an additional problem with the linear model in the context of predicted and actual points spreads in American Football betting markets. As the expected Tote and SP returns will be equal under the hypothesis of efficiency, their criticism does not apply here.

11. The missing item is the first race at Newcastle on 20th. April, 1992, which failed to produce a tote dividend due to a technical fault.

12. The additional problem of omitted variables is diagnosed by the Ramsey Reset test. Inclusion of higher powers of SP improves the fit of the model only marginally and does not alter the conclusions.

13. The ratio of 1.4 equates to the difference between a number of common odds offers such as 7 to 2 and 5 to 2; 7 to 1 and 5 to 1; 14 to 1 and 10 to 1.

14. The F-statistic for the 42 races in which the odds of the winner moved in is $F(2,40) = 1.53$. That for the 88 races in which the odds of the winner moved out is $F(2,86) = 0.20$.

CHAPTER EIGHT

EMPLOYING A NEW DATA SET TO TEST FOR STRONG FORM EFFICIENCY IN BRITISH RACETRACK BETTING MARKETS

8.1 Introduction

It has been discussed already that a market can be considered strongly efficient if all information, including private information, is incorporated into prices, so that there exist no opportunities to trade upon it, so as to earn an above-average or abnormal return. There are clear possibilities inherent in an activity like horse racing for monopolistic access to private information, and consequent opportunities for insider trading. In this sense, racetrack betting markets are somewhat analogous to the operation of conventional financial markets, yet in some important respects any insider activity is easier to measure and assess.

Published investigations of this issue include contributions by Dowie (1976), Crafts (1985, 1994), Shin (1993), Hurley and McDonough (1995), Schnytzer and Shilony (1995), Terrell and Farmer (1996), and Vaughan Williams and Paton (1996, 1997a). Common to all these studies is the idea that there are 'insiders' who may exist, and who can predict race outcomes with more accuracy than the average of all bettors. Each of these studies adopts different approaches to test this hypothesis.

Section 8.2 identifies the approach used by Crafts (1985) to investigate the opportunity for profitable use of information not publicly available, and this is applied to the new data set. Section 8.3 applies Shin's approach (1993) to the new data set, and derives an estimate of the extent of insider trading in the market. New empirical tests are employed to examine whether this model actually does explain all the data, or whether the data is consistent with an explanation other than the presence of insider trading.

In section 8.4, a new model of bookmaker behaviour in the context of adverse selection is proposed, and tested against the data generated from this thesis.

8.2 Identifying insider trading in recent British racetrack betting markets - based on an approach used by Crafts (1985).

In this section, an approach proposed and applied by Crafts (1985) to identify the existence of insider trading in British racetrack betting markets, is adapted and applied to the data set compiled for this project.

Crafts (1985) argued that insiders in a bookmaking system may bet at odds greater than the starting price. He therefore proposed separating out cases where there existed a 'marked difference' between the forecast price and the starting price. He reasoned that a shortening (lengthening) of the odds available about a horse during the course of the market may indicate evidence of insiders who knew that the true probabilities of that horse winning were greater (less) than those implied in the odds offered early in the market, or even before the market was formed (forecast prices). He also identified cases of 'marked' movements in the prices actually offered on the course during the period of trading. In each case the data set was cleaned to allow for idiosyncrasies in the compilation of the forecast prices.

If insider trading is the cause of any such marked price movements, then where these occur the expected return at forecast prices should differ depending on the subsequent odds movements. In particular, the expected return to bets placed on horses which shortened markedly between forecast and starting price should be significantly higher at forecast prices than those whose odds lengthened. Moreover, the expected return to bets at forecast prices, or at early prices, might be expected to be especially strong in races which offer particular scope for insider trading. In order to isolate races which might be the target of above-average levels of insider activity, Crafts divided the sample into handicaps and non-handicaps. The idea behind this is that handicap races (which consist of horses with a specified degree of established public form) are less likely to offer as much scope for insiders to trade as non-handicap races, where the form need not be so exposed to public scrutiny.

Crafts' results suggest that horses displaying a marked shortening of the odds between the forecast and starting price stages are indeed characterized by an exceptionally high expected return at forecast prices. Moreover, this is particularly strong for non-handicap races.

An examination of the scope for profitable arbitrage during the formation of starting price odds in the actual on-course market produced similar findings, although splitting the sample as before (into handicaps and non-handicaps) failed to reproduce the earlier result.

On the basis of this evidence, Crafts concluded that British horse race betting markets do offer insiders the opportunity for profitable exploitation of information not publicly available. As such, it is not strongly efficient.

In this section, Crafts' methodology is applied to the new data set constructed for this thesis, with one additional distinction. Since it is possible that insiders may use public information to improve the value of their private information, an additional category is included, of higher grade handicaps (excluding both non-handicaps and handicaps rated below 100)[1]. As these races are the subject of particular public attention and scrutiny, it is unlikely that they offer much scope for the profitable use of private information, whatever the state of public information.

The data used in this study are cleaned somewhat along the lines suggested by Crafts. In particular, all horses for which neither the forecast or starting prices was equal to or less than 10 to 1 were excluded.

The results are presented in tables 8.1 and 8.2. For comparison purposes with Crafts' study, the cumulative returns to level stakes at starting price are shown in table 8.3.

Table 8.1: The Pattern of Marked Differences Between FP and SP Odds

(1) 1.5≤ FP.SP < 2.0 Average return/£1 bet (£)				(2) FP/SP ≥ 2.0 Average return/£1 bet (£)				(3) 1.5 ≤ SP/FP < 2.0 Average return/£1 bet (£)				(4) SP/FP ≥ 2.0 Average return/£1 bet (£)			
Won	**Lost**	**at FP**	**at SP**	**Won**	**Lost**	**at FP**	**at SP**	**Won**	**Lost**	**at FP**	**at SP**	**Won**	**Lost**	**at SP**	**at FP**
(a) 34	95	1.38	0.94	44	160	2.44	1.15	17	188	0.31	0.45	12	168	0.34	0.74
(b) 29	95	1.88	1.25	18	77	1.89	0.84	10	155	0.49	0.75	3	83	0.13	0.23
(c) 57	166	1.60	1.08	60	178	2.31	1.07	26	323	0.39	0.59	14	242	0.26	0.56
(d) 6	24	1.80	1.18	2	15	1.12	0.47	1	20	0.43	0.62	1	9	0.55	1.00
(e) 63	190	1.62	1.09	62	193	2.23	1.03	27	343	0.39	0.59	15	251	0.27	0.58

(a) Non-handicap races
(b) Handicap races
(c) Non-higher grade handicap races
(d) Higher grade handicap races
(e) All races

Table 8.2: The Pattern of Marked Differences Between LP and SP Odds For Horses Supported in the On-Course Market

	1.5 ≤ LP/SP < 2.0 Average return/£1 bet (£)				LP/SP ≥ 2.0 Average return/£1 bet (£)			
	Won	Lost	at LP	at SP	Won	Lost	at LP	at SP
(a)	29	75	1.68	1.14	9	30	2.67	1.37
(b)	10	56	1.04	0.69	5	25	2.3	1.14
(c)	38	120	1.47	1.00	14	49	2.75	1.39
(d)	1	11	0.92	0.63	0	6	0	0
(e)	39	131	1.43	0.97	14	55	2.51	1.27

(a) Non-handicap races
(b) Handicap races
(c) Non-high grade handicap races
(d) High grade handicap races
(e) All races

Table 8.3: Cumulative Returns to Level Stakes at SP

	All	Markedly in *	Markedly out +
Up to evens	0.93	0.92	0.39
Up to 5/1	0.89	0.98	0.62
Up to 10/1	0.84	0.56	0.56
Up to 16/1	0.77	1.12	0.63
All	0.67	1.09	0.42

* FP/SP ≥1.5
\+ SP/FP ≥ 1.5

These results are broadly in line with Crafts' own findings. For the sample of horses which showed the most marked shortening between forecast and starting prices (FP/SP$\geq$2) the expected return at forecast prices is particularly high for non-handicap and non-higher grade handicap races. Moreover, it is notably low for the sample of higher grade handicap races. The same pattern is found for the sample of races showing a marked shortening between the longest price available in the on-course market and the starting price.

It is true that at lower levels of shortening (1.5$\leq$FP/SP$\leq$2) the handicap/non-handicap split appears to work contrary to expectations, but the broad thrust of the findings are consistent with those reported by Crafts. To this extent, it is reasonable to suggest that there exists some evidence with which to reject a null hypothesis of strong efficiency in recent British racetrack betting markets.

8.3 Identifying insider trading in recent British racetrack betting markets - based on an approach used by Shin (1993).

Traditional explanations of the favourite-longshot bias have focused on the behaviour of bettors, and are as such demand-side in origin. Some recent explanations, traceable to Shin (1991, 1992, 1993) have focused on an alternative supply-side explanation, based on the behaviour of bookmakers who are faced by an unknown number of bettors who may possess information superior to themselves, and who may therefore be able to trade as "insiders." In this section Shin's explanation is examined, and new tests are proposed and applied which seek to distinguish this supply-side explanation from alternative explanations of the data consistent with a conventional demand-side approach.

8.3.1 Shin's supply-side explanation of a favourite-longshot bias in British racetrack betting markets

Shin modelled the British racetrack betting market as a process characterized by bookmakers who operate in the context of an adverse selection problem, posed by the existence of an unknown proportion of 'insiders', who are able to observe the outcome of a bet with certainty. In Shin's models, it is shown that so long as the incidence of insider trading is not larger when a favourite is tipped to win than when a longshot is tipped to win (or more precisely, so long as the ratio of insider trading to the winning probability falls as the probability of winning rises), then equilibrium prices will exhibit a favourite-longshot bias. The intuition lies in the comparative rates of change of costs and revenue faced by the bookmaker as odds increase. At a given stake, the bookmaker faces greater losses to insiders at greater odds, and so will tend to restrict disproportionately these odds relative to the perceived objective probabilities.

Shin (1993) provides an estimate of the extent of insider trading based on a proposed link between the size of the bid-ask spread in the market and the prevalence of insider trading. The key to his approach lies in the idea that the direct effect of insider trading on bookmakers' margins will tend to increase as the number of runners (and therefore the size of the odds) increases. Employing a small sample of 136 UK races, Shin found a strong positive correlation between the sum of bookmakers' prices and the number of runners. By isolating this effect from other influences on the bookmakers' margin, he found the incidence of insider trading in his data set to be about 2 per cent.

This interpretation of Shin's (1993) results depends on accepting the insider trading explanation of the observed correlation between the number of runners in a race and the sum of prices offered about runners in such races. It is argued here that Shin's observations could be consistent with explanations which do not rely on the presence (or even the perception) of insider trading. As a consequence, more discerning tests are required which are able to distinguish the incidence of insider trading in the data. This permits a more clear assessment of the effect of insider trading on the odds spreads offered by bookmakers in British horse race betting markets.

8.3.2 Alternative explanations of the correlation (observed in British racetrack betting markets) between the number of runners and the sum of prices

It is proposed to show here that the linear relationship between the number of horses and the sum of bookmakers' prices found in Shin (1993) can be explained not only by the presence of insider trading but also by bettors counting only a fixed fraction of their losses.

Assume that there are many, risk-neutral bookmakers who compete on price up to the point that they expect the subjective returns to the set of bettors on each horse to be zero. For simplicity, assume bookmakers have zero costs, that bettors are also risk-neutral, and that both bettors and bookmakers have equal access to all publicly available information.

In a field of n horses, let the objective probability of winning for horse i be given by p_i ($i = 1$ to n). We define the bookmaker's over-round (OR) as the sum of the winning probabilities implied by the odds of all horses in a race minus 1. If bettors count all their losses, the odds reflect the objective winning probabilities and the over-round is zero. However, if only a fixed fraction of losses, f, is counted by bettors (as in Henery, 1985), the subjective probability of losing on any horse is $f.q_i = f.(1-p_i)$, where q_i is the objective probability of losing for horse i. The winning probabilities implied by the equilibrium odds are now $1 - f.q_i$. OR may be derived as follows:

$$OR = \sum_{i=1}^{n} (1 - f.q_i) - 1$$

$$OR = n - f.\sum_{i=1}^{n} q_i - 1$$

$$OR = n - 1 - f.\sum_{i=1}^{n} (1 - p_i)$$

$$OR = (n-1) - f.(n-1)$$

$$OR = (n-1).(1-f)$$

The over-round is linearly related to the number of runners, which is precisely Shin's 1993 result. Thus insider trading is not the only theoretical explanation of a link between the sum of prices and number of runners.

One way to distinguish between these two alternative explanations is to isolate those races in which insider trading is likely to be more prevalent. In order to identify and distinguish such situations, higher grade handicap races are isolated from other races (see section 8.2).

The market may also reveal, ex-post, information as to whether bookmakers perceive certain races to be subject to insider trading through movement in the odds. At the opening of the market, bookmakers are unlikely to know on which horse private information is held. Thus they may post less favourable odds on all horses in races where they suspect insider trading will be prevalent. As the market progresses, bookmakers reduce the odds of those horses which are heavily backed. In races where insider trading is not suspected, bookmakers are less likely to adjust prices during the life of the market. Following suggestions in Crafts (1985, 1994) and Bird and McCrae (1987), those races are singled out in which the odds on the winning horse have decreased significantly over the course of the market.

8.3.3 Data and Estimation Results

8.3.3.1 Data

The data set consists of observations on 5903 horses running in 510 races in the 1992 UK flat racing season from 19th March to 18th May inclusive. Following Dowie (1976), Tuckwell (1983) and Crafts (1985, 1994), forecast prices provided about horses on the morning of the race day are used, and these are employed to proxy for earliest market prices. These prices are particularly useful in this regard since they are not susceptible to any pre-market influences on prices.[2]

Data on the forecast price (FP) and starting price (SP) is obtained from the *Sporting Life* newspaper, and include all races (except all-weather races) on standard race days (defined as days on which six or seven races took place), and in which no horse was withdrawn too late for a fresh book to be formed.

Two adjustments are made to the FP. First, from each FP is deducted an amount based on the implied probability in the forecast odds of any horse(s) quoted in the morning but withdrawn before the start of the race (variable "NV" in the data set). This is a similar deduction to that made by bookmakers, under Tattersall's Committee Rules on Betting 4(c), when horses withdraw after a market has been formed (variable "DE" in the data set). Second, it is common for a unique forecast price not to be quoted for some longshots. In these cases a price, known as the Bar Price, is allotted to these longshots. All horses for which only a Bar price is quoted are excluded. Crafts (1985) excludes some horses where the FP is less than the SP on the grounds that FPs are sometimes innately conservative. Since this is one of the phenomena being investigated, however, such a procedure is not justified here. This leaves a sample of 4689 horses running in 481 races.

Table 8.4 Distribution of Prices and Mean Returns to a Unit Stake

Odds Range	Freq (FP)	Mean Return	Freq (SP)	Mean Return
≤1	51	-0.0550	84	-0.0735
>1 & ≤5	1164	-0.1794	1045	-0.1126
>5 & ≤10	1663	-0.3086	1431	-0.2156
>10 & ≤15	1145	-0.3384	832	-0.3594
>15 & ≤20	541	-0.3754	716	-0.3575
>20 & ≤40	110	-0.4545	495	-0.6909
>40	15	-1	86	-0.4070
All Runners	**4689**	**-0.2944**	**4689**	**-0.2911**

Table 8.4 summarises the spread of odds and the mean return to a unit stake on all horses at both FP and SP. In both cases, the odds frequencies are skewed towards favourites, but there are a far greater number of longshots at SP than at FP. On average, returns are slightly lower at FP and in no odds range can bettors expect a positive return. At both FP and SP, returns decrease fairly consistently as the odds increase, suggesting the traditional longshot bias is present in this data. Table 8.5 provides a breakdown of the number of runners in each race.

Table 8.5 Distribution of Number of Runners

Number of Runners	Total Frequency
≤5	45
>5 & ≤10	196
>10 & ≤15	144
>15 & ≤20	78
>20 & ≤25	18
Total	**481**

8.3.3.2 Estimation Results

The first step is to replicate Shin's (1993) approach on this (much larger) data set, using the following equation taken from Shin (1993)[3]:

$$D = const + z(n-1) + \sum_{k=0}^{k=2} a_k n^k VarP + \sum_{k=0}^{k=2} b_k n^k (varP)^2$$

where D = sum of starting prices in each race - 1.

n = number of runners in each race.

VarP = vector of Shin's "variance" of winning probabilities in each race.

Shin interprets z as the incidence of insider trading which he finds to be about 0.02 (or 2 per cent). An initial estimate of z is obtained using the variance of prices as a proxy for variance of probabilities. This value of z is then used to estimate the variance of probabilities and the equation re-estimated. The process is repeated until convergence is achieved.

Table 8.6 OLS Estimates of Sum of Prices

	No Adjust	1st Adjust	2nd Adjust	3rd Adjust
n - 1	0.0203*** (0.0013)	0.0194*** (0.0012)	0.0194*** (0.0013)	0.0194*** (0.0013)
VarP	0.2951 (2.2955)	-0.7613 (1.8230)	-0.7685 (1.8149)	-0.7689 (1.814)
nVarP	-0.1358 (0.6540)	0.2202 (0.5000)	0.2230 (0.4986)	0.2231 (0.4984)
n^2VarP	0.0285 (0.0369)	0.0543** (0.0275)	0.0544** (0.0275)	0.0544** (0.02745)
$(\text{VarP})^2$	-5.0972 (2.189)	-26.748 (28.602)	-26.560 (28.525)	-26.549 (28.520)
$n(\text{VarP})^2$	4.6327 (19.112)	19.916* (12.103)	19.841 (12.089)	19.837 (12.089)
$n^2(\text{VarP})^2$	-0.8859 (2.1421)	-3.6077*** (1.2808)	-3.6039*** (1.2805)	-3.6036*** (1.2805)
Constant	0.0245 (0.0174)	0.0027 (0.0169)	0.0027 (0.0171)	0.0027 (0.0171)
R2(Adj)	0.6889	0.7138	0.7138	0.7138
n	481	481	481	481

Notes:
Standard errors are in brackets
* indicates significance at 10% level; ** at 5% level; *** at 1% level.

Table 8.6 presents the replication of Shin's approach.[4]

Convergence to five decimal places is achieved on the fourth iteration. The estimates of z reported here are slightly lower than those of Shin, although at convergence the difference is marginal (0.019 compared to 0.02). Like Shin, the inclusion of polynomials above degree two barely changes the estimate of z found here.

On the basis of Shin's theoretical model, this positive link between the sum of prices and the number of runners provides evidence of insider trading at work in the market. However, as already demonstrated, demand-side explanations of the longshot bias can also bring about this result.

In order to distinguish between these alternative explanations, a dummy variable is constructed, which takes the value of 1 if the race is a higher grade handicap, and zero otherwise. The link between the number of runners and the sum of prices is expected to be weaker in such cases if it is caused by the threat of insider trading. A dummy variable equal to 1 is also constructed for races in which the odds of the winning horse are significantly greater at FP than at SP. The benchmark chosen is where FP>1.4SP.[5] If caused by insider trading, the link between the number of runners and the sum of prices is again expected to be stronger in these races.

Table 8.7 OLS Estimates of Sum of Prices : controlled for insider trading proxies

	(1)	(2)
n - 1	0.0189*** (0.0012)	0.0203*** (0.0013)
VarP	-0.8939 (1.7432)	-0.9746 (1.7319)
Nvarp	0.0758 (0.4817)	0.2529 (0.4829)
n^2VarP	0.0482* (0.0269)	0.0334 (0.0272)
$(VarP)^2$	-22.411 (27.484)	-14.910 (27.441)
$n(VarP)^2$	17.555 (11.684)	13.6104 (11.697)
$n^2(VarP)^2$	-2.8593** (1.2487)	-2.5512** (1.2456)
Handicap	-0.0267*** (0.0088)	0.0262 (0.0214)
Odds move in	0.0177** (0.0070)	0.0180*** (.0070)
Handicap*(n-1)	-	-0.0050*** (0.0018)
Constant	0.0190 (0.0171)	0.0028 (0.0180)
R2(Adj) n	0.7276 481	0.7312 481

Notes: See Table 8.6.

The results of including these two dummies are shown in Table 8.7, column 1. The coefficient on the dummy for higher grade handicap races is negative, whilst on the dummy for horses whose odds move in it is positive. In order to allow z to vary across the race types, in column 2 an interaction term is included between high grade handicap races and (n - 1). This proves to be negative and significant at the 1 per cent level, whilst rendering the simple handicap dummy insignificant. This suggests that the link between the sum of prices and number of runners is weakened when the race is a higher grade handicap. The movement in the odds interaction term is not significant and is not reported here.

In the light of this result, the model is re-estimated separately for higher grade handicaps and other races, the results reported in Table 8.8. For the other races, convergence (to five decimal places) is achieved on the third iteration. Results of this last adjustment are reported in column 1. The measure of insider trading, z, is slightly higher than our original estimate and still strongly significant. Once again, the sum of prices is higher in races where the favourite's price has moved in. However, in the case of higher grade handicap races, z is never significant at the 5 per cent level and there is no convergence. Only the initial iteration is reported in column 2.

A Chow test for the stability of coefficients between columns 1 and 2 yields an F(9, 463) statistic of 2.17. The null hypothesis of constant coefficients is rejected at the 5 per cent level of significance. The Chow test, however, imposes the restriction that the error variance is constant across the two models. An F test rejects the hypothesis of constant variance at the 5 per cent level. Consequently, a $\chi^2(9)$ test is also applied, using a Wald statistic (see Greene, 1993, p.215) which does not require this assumption. The test statistic of 25.0 confirms rejection of the null hypothesis, now even at the 1 per cent level, and justifies the estimation of two separate models.

Thus the correlation between the number of runners and the sum of prices is restricted to races which do not involve a high grade of handicap. These results are consistent with Shin's 1993 claim that the link is due largely to insider trading.

Table 8.8 OLS Estimates of Sum of Prices by Higher Grade Handicap races

	(1) Other Races	**(2) High Grade Handicap Races**
n - 1	0.0208*** (0.0014)	0.0087* (0.0045)
VarP	-1.1210 (1.8187)	5.1390 (20.671)
Nvarp	0.3479 (0.5106)	-3.8277 (6.0893)
n^2VarP	0.0207 (0.0286)	0.5012 (0.3673)
$(VarP)^2$	-5.2470 (28.545)	-2552.0 (2813.9)
$n(VarP)^2$	8.6977 (12.1701)	945.64 (855.28)
$n^2(VarP)^2$	-1.9743 (1.3110)	-84.007 (57.382)
Odds move in	0.0156** (0.0074)	0.083*** (0.0260)
Constant	0.0013 (0.0191)	0.0678 (0.0512)
R2(Adj)	0.7291	0.7114
n	419	62

Notes: See Table 8.6.

8.3.4 Conclusions

The favourite-longshot bias, i.e. the tendency for the normalised prices available about short-odds events to understate the objective winning probabilities, has been demonstrated historically in British racetrack betting markets. Traditional explanations have emphasized the psychological profile and risk preferences of bettors, while some recent explanations have characterized this bias as a rational response by bookmakers facing some bettors who may possess and trade upon superior information.

In this section a new method has been identified for isolating data likely to be characterized by above-average and below-average amounts of insider activity. By using higher grade handicap races as a proxy for races which are less likely to provide scope for asymmetric information, support has been found for Shin's (1993) conclusion that insider trading is a likely explanation of the positive correlation between the sum of prices and the number of runners. Insider activity seems, therefore, to explain at least part of the favourite-longshot bias found in British racetrack betting markets.

8.4 Presenting a new model of horse race betting

In this section, a model is presented of behaviour in a betting market characterized by odds setters in which the threat of insider trading leads to a tendency for favourites to be under-bet and longshots over-bet. If the odds setters are able to adjust their prices in the light of new information revealed by the market, it is shown that the bias will be reduced over the course of the market. To test the model the large new set of data derived from the British racetrack betting market is employed.

The model of horse race betting market proposed here is based on similar assumptions about the nature and type of bettors to those of Hurley and McDonough (1995). For simplicity, the model is restricted to two horses - a favourite and a longshot. Two sets of winning probabilities are proposed for the two horses. First, there are the true probabilities of winning for the favorite and longshot respectively, namely [s; (1-s)]. Second, there are the subjective probabilities of the bookmakers, [a, (1-a)], all of whom are identical and have equal access to public information. When all information is available publicly, it is assumed that the subjective probabilities are equal to the true probabilities. In the presence of some inside information, it is assumed that the expected value of a is equal to the true winning probability, i.e. E(a) = s.

There is a distribution of risk-neutral bettors. Some of these may be informed about the true probabilities, while the rest are noise traders who cannot distinguish the value of a bet. Bookmakers are also risk-neutral and compete on price up to the point that they expect the subjective returns to the set of bettors on each horse to be zero. For simplicity, bookmakers are assumed to have zero costs. The set of odds for the favourite and longshot respectively are [O_f, O_l]. The expected return to a unit stake on the favourite is [E(a).O_f - (1-E(a))] and on the longshot, [1-E(a)].O_l - E(a)]. If all bettors are noise traders, an expected net payout of zero on each horse implies that the bookmaker sets odds of [(1-a)/a; a/(1-a)]. Where there is no privately held information, the set of odds are [(1-s)/s; s/(1-s)].

This framework is used to investigate the implications of potential insider trading. To do this, the assumption is made that bookmakers believe that a proportion of bettors (w) hold private information which is potentially superior to that publicly available. It is further assumed that in reality there is no private information and that all bettors are simply noise traders. This implies that the bookmakers' subjective probabilities are in fact equal to the true probabilities. Thus all that matters here is the threat of inside information and not its reality.

Two possibilities are considered: the bookmakers believe that they have either underestimated or overestimated the favorite's winning probability by a positive amount, α, (i.e. they believe $s = a + \alpha$ or $s = a - \alpha$). They do not know which is the actual case and assign an equal probability to each alternative.

The model employed here is of a market consisting of two stages, the first before trading has commenced and the second prior to the race being run.

STAGE 1:

At the start of the market, the bookmakers set odds on each horse to allow for the fact that they may have underestimated the winning probability. Their implied probabilities on the favorite (longshot) are increased by the positive amount B_f (B_l). Lower odds are now set on each horse, detailed in Stage 2.

$$\left[\frac{(1-a-B_f)}{(a+B_f)}\ ;\ \frac{(a-B_l)}{(1-a+B_l)}\right]$$

STAGE 2:

As the market progresses, the bookmakers have a chance to readjust their subjective probabilities in the light of information revealed by the market. If neither horse is more heavily backed than the other, then bookmakers assume that no insider trading is present and they therefore adjust their odds accordingly.[6]

In Stage 1, bookmakers believe that insiders will bet with certainty on the favorite if $s = a + \alpha$ and with certainty on the longshot if $s = a - \alpha$. Other bettors, not knowing which horse is underpriced, bet with equal probability on either horse.

The bookmakers expect the return to an insider from a unit bet to be:

$$\textit{Bet on favourite with certainty:}\quad \frac{(a+\alpha).(1-a-B_f)}{(a+B_f)} - (1-a-\alpha)$$

$$\textit{Bet on longshot with certainty:}\quad \frac{(1-a+\alpha).(a-B_l)}{(1-a+B_l)} - (a-\alpha)$$

The bookmakers expect the return to other bettors to be: (derived in Appendix 3)

$$\textit{Bet on favourite with pr=0.5:}\quad \frac{a.(1-a-B_f)}{(a+B_f)} - (1-a)$$

$$\textit{Bet on longshot with pr=0.5:}\quad \frac{(1-a).(a-B_l)}{(1-a+B_l)} - a$$

For equilibrium, the overall expected net payout to the set of all bettors must be zero. In Appendix 3 it is shown that, when the expected net payout to each horse is equal, the equilibrium values of both B_f and B_l are both equal to wα. Thus it is optimal for bookmakers to add the same amount to the implied probabilities of both longshots and favourites in the presence of perceived insider trading. The greater is the amount of private information which bookmakers believe to be held and the greater the estimated proportion of insiders, the greater will be the implied probabilities of winning for both horses. The implied equilibrium odds at Stage 1 are:

$$\left[\frac{(1-a-\alpha w)}{(a+\alpha w)};\ \frac{(a-\alpha w)}{(1-a+\alpha w)}\right].$$

Since the odds of both the longshot and the favorite are adjusted by the same amount, the longshot bias can be derived from the objective expected return. For any horse with objective probability of winning, s,

$$E(Ret) = s.(odds) - (1-s)$$
$$E(Ret) = \frac{s.(1-s-\alpha w)}{(s+\alpha w)} - (1-s)$$
$$E(Ret) = \frac{s}{(s+\alpha w)} - s - 1 + s$$
$$E(Ret) = \frac{s - s(s+\alpha w) - (s+\alpha w) + s(s+\alpha w)}{(s+\alpha w)}$$
$$E(Ret) = -\frac{\alpha w}{(s+\alpha w)}$$
$$E(Ret) = -\alpha w\left(\frac{1}{s+\alpha w} - \frac{s+\alpha w}{s+\alpha w}\right) - \alpha w$$
$$E(Ret) = -\alpha w\left(\frac{1-s-\alpha w}{s+\alpha w}\right) - \alpha w$$
$$E(Ret) = -\alpha w(odds + 1)$$

Thus as long as bookmakers believe there are some insiders with non-trivial private information, there is a longshot bias. The more inside knowledge that bookmakers believe to exist, the greater will be this bias.

In Stage 2 the odds are adjusted to the true probabilities, and the longshot bias is reduced. More realistically, the longshot bias decreases as the market progresses, a conclusion consistent with Kyle's (1985) modelling of insider trading as a process of information revealed over time.[7]

The model is now subjected to empirical tests.

8.4.1 Data and Estimation Results

8.4.1.1 Data

The data set consists of observations on 5903 horses running in 510 races in the 1992 UK flat racing season from March 19th to May 18th inclusive. Data on the forecast price (FP) and starting price (SP) for each horse were gathered as described in Section 2, and the same procedure used to clean the data. This leaves a sample of 4689 horses running in 481 races.

In order to estimate the expected return from a bet on each horse, odds categories are used, as in Henery (1985). Specifically FP and SP are divided into 25 and 26 categories respectively. The ex-post probabilities of a horse in a category winning are calculated as the number of winners divided by the total number in each category. The expected return to a unit stake in a particular category are then calculated as:

$$E(\text{return}) = \text{pr}(\text{winning}).\text{odds} - [(1 - \text{pr}(\text{winning})]$$

Grouping the observations in this way means that both the returns and odds variables are subject to some measurement error. The consequences of this for the regression estimates are discussed in Section 8.4.1.2 below.[8]

Table 8.9 summarizes the spread of odds and the expected return to a unit bet for both FP and SP. In both cases, the odds frequencies are skewed towards favourites, but there are a far greater number of outsiders at SP than at FP. Returns are lower on average at FP and in no odds category can bettors expect a positive return. At both FP and SP, returns decrease fairly consistently as the odds increase.

Table 8.9 Distribution of Prices and Mean Returns to a Unit Stake

Odds Range	Freq (FP)	Mean Return	Freq (SP)	Mean Return
≤1	51	-0.0550	84	-0.0735
>1 & ≤5	1164	-0.1794	1045	-0.1126
>5 & ≤10	1663	-0.3086	1431	-0.2156
>10 & ≤15	1145	-0.3384	832	-0.3594
>15 & ≤20	541	-0.3754	716	-0.3575
>20 & ≤40	110	-0.4545	495	-0.6909
>40	15	-1	86	-0.4070
All Runners	**4689**	**-0.2944**	**4689**	**-0.2911**

8.4.1.2 Estimation Results

The model predicts lower expected returns for higher priced horses under the threat of insider trading and, in particular, suggests a linear relationship between expected return and the odds plus 1. The following equation is estimated for horse i:

$$E(ret)_i = \beta_0 + \beta_1(odds_i + 1) + u_i$$

The intercept term (β_0) is included to take account of a premium which may be levied in practice by bookmakers to cover their costs, and is expected a priori to be negative. A longshot bias is present if β_1 is significantly negative. The model predicts that any longshot bias will be lower at starting prices than forecast prices, and so β_1 should be lower (i.e. more negative) for FPs than for SPs.

As the observations are clustered within the various odds categories, least squares is likely to underestimate the standard errors. White standard errors (White, 1980) are therefore employed, generalized to allow for cluster sampling. In addition, the fact that each of the odds categories contains a different number of observations may lead to increased heteroscedasticity problems. Thus weighted least squares estimates are presented, weighted by the odds frequency.[9]

The model is estimated for FP and SP, and the results reported in Table 8.10, columns 1 and 2 respectively. β_1 is significantly negative in both cases: at FP it is -0.0200 and at SP it is -0.0149. The measurement error in the independent variable implies that these estimates are biased towards zero. The estimated bias on β_1 in the FP equation (calculated as the ratio of the error variance to the total variance [see Greene, 1993, p.281]) is extremely small (1.3 per cent) whilst that for the SP equation is somewhat larger (9.6 per cent). However, after adjustment, the longshot bias is still greater at FP than at SP.

This is investigated further by pooling the data and including a dummy interaction term for the odds at forecast prices. The significant negative coefficient on this variable in column 3 suggests rejection of the null hypothesis that β_1 is constant across both samples, but only at the 10 per cent level of significance.[10] A link test (Pregibon 1980) suggests the inclusion of a quadratic odds term and this is reported in column 4. The fit of the model is improved slightly and the coefficient on the linear interaction term is now significant at the 1 per cent level, providing further evidence of the greater longshot bias at FP than at SP. The estimated bias from measurement error on the pooled odds is 8.4 per cent. This does alone seem capable alone of explaining the observed difference between the longshot bias at FP and SP.

8.4.2 Longshot Bias at Forecast Prices

The next step is to concentrate on the longshot bias at FP, incorporating some of the race information which is available for each horse.[11] First, it is hypothesized that bookmakers believe that there will be less insider trading in situations where more public information is available. In order to identify and distinguish such situations, Crafts (1985) suggests separating handicap races (where horses are allocated weights so as to equalize as far as possible their chances of winning) from non-handicaps. The idea behind this is that the past form of horses in handicaps is generally more established in the public forum than in non-handicaps. In the latter, therefore, bookmakers are likely to believe there to be greater possibilities for betting on the basis of privately held information.

It is possible that insiders may use public information to improve their private information. For this reason it may be better to consider only higher grade handicaps as indicative of the absence of useful private information (see discussion in Section 2). In order to measure the change in the longshot bias for higher grade handicap races compared to other races, the interaction term, between this dummy variable and forecast prices plus one, is included.

Second, the market may be able to provide, ex-post, information on which races bookmakers believe insider trading to be prevalent. The theoretical model proposed here suggests that bookmakers' odds will reflect the true probabilities throughout the market when there is no perceived insider trading. In this case, there is more likely to be no movement in the odds from FP to SP. Hence, a dummy variable is included for horses whose odds do not move as well as an interaction term with forecast prices plus one.

Finally, there is some evidence (Henery, 1985; Shin, 1993) that both the level of returns and the longshot bias may vary with the number of runners in a race. It is therefore proposed to include the number of runners and the interaction term between number of runners and odds as control variables.

Results of this specification are reported in Table 8.11, column 1. The two previous interaction terms are indeed positive, suggesting that the longshot bias is lower in higher grade handicap races and where the odds do not move. The straight dummy terms are both negative. This implies that the fixed premium levied by the bookmakers is greater in these cases.

A link test again suggests the inclusion of a quadratic term. This is done in column 2. Once again the results are robust to the different specification. Indeed, the odds variable as well as its interaction with the no movement dummy, are now even more significant. Interaction terms with the quadratic variable are never significant and are not reported here.

The model is replicated on starting prices in column 3. Although the longshot bias is again present, none of the variables which we use as proxies for perceived insider trading prove to be significant. This provides further evidence that the threat of insider trading has less of an effect on the odds as the market progresses.[12]

8.3.5 Conclusions

The results of the tests employed in this section are consistent with information on insider trading being revealed to bookmakers as the market progresses. In addition, proxy variables for the extent of perceived insider trading seem to be successful in explaining at least part of this bias. In particular, the bias at forecast prices is lower both for higher grade handicap races and where odds subsequently do not move. Thus the empirical work reported here is consistent with the proposition that the concept of insider trading can help to explain the longshot bias. If valid, this is evidence in contradiction of the proposition that racetrack betting markets, at least in the U.K., are strong-form efficient.

Table 8.10: WLS Estimates of Expected Returns

	(1) FP	(2) SP	(3) Pooled	(4) Pooled
Odds+1	-0.0200** (0.0087)	-.0149** (0.0062)	-0.0143*** (0.0043)	-0.0188*** (0.0059)
$(\text{Odds} + 1)^2$				7.63 e-5 9.94 e-5
FP dummy x Odds+1			-0.0093* (0.0049)	-0.0179*** (0.0076)
FP dummy x $(\text{Odds}+1)^2$				5.80 e-4 (4.08 e-4)
Constant	-0.1501** (0.0779)	-.09349** (0.0895)	-0.1034* (0.0556)	-0.0549 (0.0568)
R^2(adj) n	0.2517 4689	0.4164 4689	0.3529 9378	0.3735 9378

Notes:
White standard errors are in brackets
* indicates significance at 10% level, ** at 5% and *** at 1%

Table 8.11: WLS Estimates of Expected Returns At Forecast and Starting Prices

	(1) FP	**(2) FP**	**(3) SP**
Odds+1	-0.0144* (0.0079)	-0.0260** (0.0128)	-.0137*** (.0047)
$(\text{Odds} + 1)^2$		5.11 e-4 (3.79 e-4)	
Handicap Dummy	-0.0819** (0.0374)	-0.0609** (0.0296)	-0.0089 (0.0162)
Handicap x (Odds+1)	0.0079** (0.0036)	0.0060** (0.0029)	6.05 e-4 (9.42 e-4)
No Movement	-0.0593* (0.0316)	-0.0222* (0.0118)	0.0599** (0.0297)
No Movement x (Odds+1)	0.0055** (0.0026)	0.0023*** (8.61 e-4)	0.0023 (0.0047)
No. Runners	0.0040 (0.0056)	0.0050 (0.0053)	-0.0023 (0.0047)
No. Runners x (Odds+1)	-5.64 e-4 (3.93 e-4)	-5.79 e-4 (3.80 e-4)	1.51 e-4 2.35 e-4
Constant	-0.1810** (0.0784)	-0.1333 (0.0913)	-0.1078*** (0.0416)
R^2(adj)	0.2667	0.2845	.4018
n	4689	4689	4689

Notes:
As Table 8.10.

1. Flat race handicaps are given an official rating, normally between 0 and 60 for the lowest grade handicap races, and between 0 and 115 for the highest grade.

2. A minor exception is the small number of races about which ante-post prices (odds offered before the day of the race) are available.

3. Shin (1993) does not specify a constant term in his empirical model. However, the goodness of fit in this model is similar to that of Shin only if the constant term is included. It can be interpreted as a measure of bookmakers' costs. In any case, the estimates of z are not greatly dependent on the inclusion of a constant.

4. Jullien and Salanié (1994) describe alternative method of moments and nonlinear least squares estimators for z.

5. The ratio of 1.4 equates to the difference between a number of common odds offers such as 7 to 2 and 5 to 2; 7 to 1 and 5 to 1; 14 to 1 and 10 to 1.

6. If insider trading is actually present, one horse is more heavily backed. The bookmakers are thus informed of which horse they have underpriced and the same result follows.

7. In Kyle's (1985) model, all private information is incorporated into prices by the end of trading.

8. For more discussion of grouped odds, see Busche and Hall (1988); Woodland and Woodland (1994).

9. In fact, White's method does not require that the error terms are homoscedastic.

10. The straightforward FP dummy is not significant at any level and is not reported here. In other words, the hypothesis that β_0 is the same across the two models cannot be rejected.

11. In view of the small estimated bias on FP noted above, the issue is not pursued any further here.

12. Of course, these coefficients may be somewhat biased due measurement error on the odds variable, as discussed above.

CHAPTER NINE

EXPLAINING POSITIVE AND NEGATIVE FAVOURITE-LONGSHOT BIASES

9.1 Introduction

Most studies of both parimutuel and fixed-odds betting markets have shown a systematic tendency for the expected return to bets at lower odds to exceed those at higher odds. Some work, however, has revealed in certain markets the absence or even reversal of this bias. Chapter 9 presents a model which distinguishes two separate types of bettor to demonstrate how transactions costs, the extent of public information, and consumption benefits of betting can explain these disparities. Empirical evidence, taken from a fixed-odds market, is used to investigate this issue.

9.2 Background to the study

A number of studies have employed sports betting markets as a convenient perspective from which to improve our understanding of more sophisticated financial markets. These take advantage of the fact that betting markets are characterized by a well-defined termination point at which each asset (or bet) possesses a definite value. In consequence, it is much easier to use this particular context to formulate tests of information efficiency.

The most notable evidence of betting market inefficiencies derived from these examinations is the finding common to many of them that the expected return to a unit wager tends to vary significantly and systematically between odds classifications. This 'favourite-longshot bias' (or longshot bias) as it is commonly termed, is demonstrated in most studies as a tendency for the expected return at lower odds to exceed that at higher odds. A survey article of U.S. parimutuel markets by Thaler and Ziemba (1988) shows this as a common theme in the academic literature, and it is confirmed for fixed-odds markets in the UK (e.g. Dowie, 1976; Henery, 1985; Vaughan Williams and Paton, 1996, 1997a) and Australia (Bird, McCrae and Beggs, 1987; Bird and McCrae, 1994).

A notable exception to these findings is offered by Woodland and Woodland (1994), who found a small negative longshot bias in the US fixed-odds baseball betting market. The absence of a positive bias is also reported by Busche and Hall (1988) and Chapman (1994) for Hong Kong racetrack betting markets, by Busche (1994) for Hong Kong and Japanese racetrack betting and by Swidler and Shaw (1995) for a small US racetrack.

In this chapter, an attempt will be made to show that the differences in the findings of these investigations can be explained within a common theoretical framework. The central idea of this is the existence of costly and/or imperfect information.

9.3 Theoretical Discussion

An understanding of the potential inefficiencies in betting markets requires an appreciation of both the supply and demand sides of the market. Demand side features are common to both parimutuel and fixed odds systems. The supply side is only relevant when there are proactive odds setters or bookmakers.

A simple model is presented here of the demand side of a betting market characterized by the presence of money from both "uninformed" bettors (or noise traders) who cannot distinguish good bets and "informed" bettors who know the "true" probabilities of each horse winning. These are similar assumptions to those made by Hurley and McDonough (1995) and Terrell and Farmer (1996).[1]

Let the proportion of informed and uninformed money which is actually bet be given by x and 1-x respectively ($0 \leq x \leq 1$). For simplicity only the case of a two-horse race is considered, involving a favourite, defined as a horse with an objective winning probability of y, where y > 0.5, and a longshot which has an objective winning probability of (1-y).

The ex-post return (including the stake) to a unit bet on any horse is given by the inverse of the proportion bet on that horse, multiplied by (1 - t) where t is the ratio of transactions costs to the total amount bet.[2] t is related to the track take in a parimutuel system or the bookmakers' margin in a fixed odds system.[3]

The utility gained by informed traders from a unit bet on the favourite is defined as U_{IF}, and from a bet on the longshot as U_{IL} where:

$$U_{IF} = \text{Expected return - stake} + \alpha$$

$$U_{IL} = \text{Expected return - stake} + \beta$$

$$\text{where } \alpha = \beta + \gamma$$

In other words, all bets yield utility if there is a positive expected return, but there is also a consumption benefit equal to β. In addition, a bet on the favourite yields an extra consumption benefit, γ ($\gamma \geq 0$). The intuition behind this assumption is that informed bettors have superior information over noise traders in that they know which horse is more likely to win. Utility is gained from taking advantage of this superior knowledge. This may be the result of "...peer group esteem associated with perceived `skill'" (Bruce and Johnson, 1992, p.205) or simply the self-satisfaction gained from using this superior knowledge.

Finally, whereas noise traders (who act as if there is a fixed expected return to any strategy) bet a fixed amount of money, informed bettors vary the total amount staked according to the expected return, up to a pre-determined stake limit and subject to the constraint that utility from any bet is non-negative.[4]

The noise traders bet on the favourite and longshot with equal probability. That is, they bet (1-x)/2 on the favourite and (1-x)/2 on the longshot. Consequently, the noise traders underbet the favourite relative to the objective winning probabilities.

Informed traders can, in the first instance, make positive returns, on top of the consumption benefit, by betting on the favourite. In the case where all informed bets are placed on the favourite, it follows that the proportion of the total amount staked which is bet on the favourite is:

$$\frac{1-x}{2}+x=\frac{1+x}{2} \qquad (9.1)$$

The utility gained by informed traders from a bet on the favourite is:

$$U_{IF}=\frac{2}{1+x}.y+\alpha-1 \qquad (9.2)$$

where the expected return is calculated as the product of the objective winning probability and the return to a winning bet.

Should an informed trader decide to bet on the longshot, the utility for a marginal bet will be:

$$U_{IL}=\frac{2}{1-x}.(1-y)+\beta-1 \qquad (9.3)$$

As informed traders increase their stake on the favourite, U_{IF} decreases and U_{IL} from the marginal bet increases. If the proportion of bets placed on each horse reflects the probability of that horse winning, the expected return to each will be equal and there is no bias. The condition for this is:

$$\frac{1 + x}{2} = y \tag{9.4}$$

or:

$$x = 2y - 1 \tag{9.5}$$

A value of x greater or less than 2y - 1 implies that the favourite is overbet or underbet respectively.

Proposition 1: In the absence of transactions costs, the favourite will be overbet relative to its objective win probability.

To see this, consider the situation in which informed traders have increased the stake on the favourite so that expected returns to the two horses are equal.[5] At this point, the horses are backed according to their expected winning probabilities (i.e. x = 2y - 1). Because of the extra consumption benefit, a marginal bet on the favourite will still yield positive utility for the informed traders. Thus they continue to back the favourite beyond its expected winning probability and a reverse longshot bias is observed.[6]

Now consider the case in which there are transactions costs, which constitute a proportion t of the total amount bet. The utility gained from an informed bet on the favourite is now:

$$U_{IF} = \frac{2y}{1 + x}.(1 - t) + \alpha \tag{9.6}$$

Proposition 2: There will be some value of t for which the favourite is underbet relative to its objective winning probability.

The reason for this is that the constraint that U_{IF} be non-negative may now hold prior to the point at which the stake on the favourite is enough to equalise expected returns across the two horses. The constraint operates as follows:

$$\frac{2y}{1+x}.(1-t)+\alpha-1=0 \tag{9.7}$$

This reduces to:

$$x=\frac{2y.(1-t)}{(1-\alpha)}-1 \tag{9.8}$$

If $t = \alpha$, then the constraint holds at $x = 2y - 1$. The horses are bet precisely in line with their objective winning probabilities. For $t < \alpha$, the constraint will only bind when the stake on the favourite has been increased beyond this point and the favourite will be overbet as before. Lastly, for any $t > \alpha$ then the constraint holds at $x < 2y - 1$ and the favourite is underbet. In other words, if the value of transactions costs is high enough, the traditional longshot bias is present.

A further issue is the importance of the proportion of all bettors who are informed.

Proposition 3: As the proportion of bettors who are informed approaches unity, there will be neither a positive nor a negative longshot bias.

This follows from the assumption that the informed bettors gain an extra consumption benefit from taking advantage of their superior information. When all bettors are informed, there is no extra consumption benefit. The optimal strategy (within the constraint of non-negative utility) is to back each horse in line with its objective probability of winning.

An added complication in fixed odds markets is the supply side response of odds setters to asymmetric information. It has been shown (Shin, 1993; Vaughan Williams and Paton, 1997a) that the longshot bias can be the result of a rational odds-setting strategy by bookmakers who are faced by an unknown proportion of insiders who may hold superior, private information.[7]

Again, this additional source of bias is dependent on the extent of asymmetric information. The more information that is available in the public domain, the less scope there is for insider trading. In other words, betting markets in which all information is held publicly are unlikely to be subject to either a supply or demand side bias.

Fixed odds horse race betting markets provide the most general setting in which to test this framework. Included in such markets are a range of situations with differing amounts of publicly available information as well. In particular, Vaughan Williams and Paton (1997a) argue that higher-grade handicap races[8] provide a good example of a set of races in which nearly all information is available publicly. In these cases, neither insider trading, nor the presence of significant numbers of noise traders is likely. A further feature of fixed odds horse race betting markets is that, unlike both parimutuel and baseball betting markets, the bookmakers' margin (or over-round) varies over a wide range. This permits formal testing of the nature of any relationship between the longshot bias and the extent of transactions costs.

9.4 Data Description and Estimation Results

9.4.1 Data

The data set consists of observations on 5903 horses running in 510 races in the 1992 UK flat racing season. Data on the Starting Price (SP) for each horse as well as various pieces of race information were gathered from the racing daily, the *Sporting Life*.

The analysis of the variation of returns across high and low odds has been subject to much discussion. One common method (see, for example, Busche and Hall, 1988; Henery, 1985) has been to divide observations into odds groupings. Differences in the expected mean return, calculated from the number of winners in each odds category are then examined. Unfortunately such a procedure means that both returns and odds are subject to measurement error bias (for a discussion of this point, see Busche and Hall, 1988). Partly because of this bias, Woodland and Woodland (1994) argue that there are advantages to studying baseball betting markets where no such aggregation is necessary.

To get around this problem, a calculation is performed of the actual return to a unit stake on each horse. If the horse loses, the return is equal to -1. If the horse wins, the return is equal to SP. This eliminates both the need to group horses in artificial odds categories as well as some subsequent econometric difficulties. In effect, there is a mix of discrete and continuous distributions. In other words, what is partly being estimated is the probability of each horse losing. For this reason, OLS coefficients will be biased, and Tobit estimation is appropriate.[9]

The distribution of starting prices and the mean return based on this calculation is shown in Table 9.1. The mean return is negative in all the odds ranges and, in line with the traditional longshot bias literature, the mean return is markedly higher (i.e. losses are lower) in the lower odds ranges.

Table 9.1: Distribution of Prices and Mean Return to a Unit Stake

SP Range	**Freq**	**Mean Return**
≤1	84	-0.0735
>1 & ≤5	1053	-0.1099
>5 & ≤10	1463	-0.1958
>10 & ≤15	898	-0.3641
>15 & ≤20	928	-0.3362
>20 & ≤40	1085	-0.6221
>40	392	-0.6097
All Runners	**5903**	**-0.3323**

Table 9.2 summarizes the differences between higher-grade handicap races and others. On average, Starting Prices are about 11 per cent lower in higher-grade handicap races. The mean return, however, is about 45 per cent higher in the higher-grade handicap races.

Table 9.2: Distribution of Prices and Mean Return for Higher Grade Handicap and Other Races

	Higher Grade Handicap	**Others**	**All Races**
Frequency	743	5160	5903
Mean SP	15.36	17.34	17.09
Mean Return	-0.1934	-0.3523	-0.3323

Table 9.3 provides summary information on the number of runners and the "over-round" across

the 510 races. The over-round is calculated as the sum of bookmaker prices[10] and can be interpreted as the bookmaker's expected markup on each race. The markup ranges from about 10 to 45 per cent - significantly larger than the 2 to 3 per cent reported by Woodland and Woodland (1994) in their finding of a reverse longshot bias in baseball betting markets. In line with Shin (1993) and Vaughan Williams and Paton (1997a), the over-round tends to be higher the greater the number of runners there are in the race.

Table 9.3: Distribution of Number of Runners

Number of Runners	**Frequency**	**Mean Over-round** (%)
≤5	45	110.13
>5 & ≤10	198	116.95
>10 & ≤15	158	129.30
>15 & ≤20	90	137.74
>20 & ≤25	19	144.84
Total	**510**	**124.88**

9.4.2 **Estimation Results**

To establish the traditional longshot bias in this data set, the following model is estimated using the sample of 5903 horses:

$$Return_i = \beta_0 + \beta_1.SP_i + u_i \qquad (9.09)$$

A significantly negative value for β_1 would suggest that the return is lower at higher odds (that is, the longshot bias).

Column 1a of Table 9.4 presents results of Tobit estimates of this model and reveals precisely this picture. Both the constant term and the coefficient on SP are negative and significant at the 1 per cent level. In other words, the return to a unit stake can be expected to be negative at all levels, but the return decreases by a small (but significant) amount as the odds go up. It should be noted that the Tobit coefficients require adjusting before they can be interpreted as marginal effects. Estimates of the marginal effect are presented in column 1b. The figure of 0.0636 suggests that an increase in Starting Price of one unit increases the expected loss to a dollar bet by about six cents.

Table 9.4: Tobit Estimates of Returns

	(1a) Coefficient	(1b) Slope	(2a) Coefficient	(2b) Slope
SP	-.7356*** .0633	-.0636	.2896 .3850	.0250
Over-round*SP	-		-.0078*** .0030	-6.74 e-4
Constant	-15.84*** .9064		-15.93*** .9054	
N	5903		5903	
% Winning	8.64		8.64	
Log Like'd	-3141.4		-3137.5	
Pseudo R^2	.0440		.0451	
Std Error	17.20		17.15	

Notes:
(i) Dependent variable is return calculated as described in the text.
(ii) Standard errors are in brackets.
(iii) * indicates significance at the 10% level, ** at the 5 %, *** at the 1%.
(iv) Slopes are estimated following Greene (1993, p.695).

The role of commission is examined by including an interaction term between the SP for horse i, and the over-round in each race j. The model is now:

$$Return = \beta_0 + \beta_1.SP_i + \beta_2.SP_i * \text{Over-round}_j + u_i \tag{9.10}$$

The results of this specification are reported in column 2a of Table 9.4. The coefficient on the interaction term is negative and significant at the 5 per cent level, suggesting that the link between odds and the return (the longshot bias) is stronger when the over-round (bookmakers' commission) is higher. The residual effect of odds on the return (given by the coefficient on SP) is now positive, albeit not significant at conventional levels. Thus, as predicted by the theoretical model, there is, potentially at least, the possibility of a reverse bias at lower levels of transactions costs.

Again, the marginal effects are estimated and reported in column 2b. The critical value of the over-round at which the longshot bias becomes positive is just under 40 per cent. Given that this implies an over-round of less than 100 per cent (and thus negative transactions costs), these results suggest a positive longshot bias at all levels of transactions costs in this data.

The next step is to control for the amount of publicly available information in the market. Following the discussion in Section 9.3, the sample is split into the 743 horses which race in higher-grade handicaps and the 5160 others. The model is re-estimated on each sub-sample and results, with the results reported in Table 9.5.

Table 9.5: Tobit Estimates of Returns

	(1a) Higher Grade H'cap	(1b)	(2a) "Other" Races	(2b)
	Coefficient	**Slope**	**Coefficient**	**Slope**
SP	-.5457 1.453	-.0470	.2433 .4000	.0210
Over-round* SP	-.0084 .0113	-7.23 e-4	-.0076** .0030	-6.57 e-4
Constant	-22.94*** 3.404		-15.00*** .9259	
N % Winning	743 8.61		5903 8.64	
Log Like'd Pseudo R^2 Std Error	-416.7 .0168 20.50		-2716.7 .0502 16.63	

Notes:
See Table 9.4.

In the higher-grade handicap sample the signs on both the SP and interaction term coefficients are insignificant at any conventional level. There is little sign of any systematic link between the SPs and the expected return in these races. For the horses running in the "other" races, the coefficients and marginal effects are virtually unchanged from those in the full sample. The longshot bias seems to be present only in races where there is some scope for informational asymmetries.

9.5 Discussion and conclusions

The findings suggest that the longshot bias can be linked to the incidence of a commission levied on bettors in the aggregate, and also to the incidence of uninformed or noise traders in the betting market. When transactions costs are very low, a reverse bias may operate overall if informed bettors derive extra utility from betting on favourites. In situations where there is little scope for the holding of private information, there are no biases whatever the level of transactions costs. In a fixed-odds market the presence of insiders can also account for the existence of a positive longshot bias at any level of transactions costs.

The findings of Woodland and Woodland (1994) for the US baseball market are now readily explained. Since this particular market is characterized by a very low commission, one should expect the absence of a positive bias. Since some bettors may also derive extra utility from betting on favourites, it is not altogether surprising that a reverse bias is present (see Proposition 1).

The parimutuel betting markets examined by Busche and Hall (1988) and Busche (1994), are examples of a very large betting pool (in which information costs might be expected to be low relative to the size of the pool) and races which are almost exclusively higher-grade handicaps. In such an environment, one might expect the market to be dominated by informed bettors and thus the absence of a longshot bias (see Proposition 3). This is in line with the empirical conclusions of Busche and Hall (1988) and Busche (1994).

Swidler and Shaw's (1995) finding that the longshot bias is absent in an identified small racetrack market, populated by ostensibly "uninformed bettors," is at first sight more difficult to explain. A clue may lie in the following paragraph taken from their paper (p.312).

> "... in racetracks with small handles, relatively modest bets can shift the odds significantly. Thus, even with an unsophisticated betting public, all that is needed are one or two informed "arbitrageurs" to insure that subjective and objective odds are (nearly) equal."

At a very small racetrack, where the average bet of the uninformed bettor is also likely to be very small, the weight of purely profit-driven, informed (arbitrage) money is likely to dominate the market, turning the market (in terms of amounts staked) into an essentially informed one, where the absence of a longshot bias would be expected(see Proposition 3).

Exceptions to the traditional finding of a positive longshot bias can therefore be explained as special examples of the more general case. Fixed odds betting markets provide an ideal setting within which the conditions for such market anomalies can be investigated.

1. There will, in practice, be a link between the costs of becoming informed and the number of bettors who become informed. Terrell and Farmer (1996) for a formal modelling of this process.

2. This is precisely the way returns are calculated in a parimutuel system, but can be applied equally to a fixed odds system in which all bets are taken at Starting Prices and bookmakers set odds solely on the basis of amounts bet on each horse.

3. It can also be though of as incorporating costs of obtaining information.

4. This may be set by the bookmakers (see Shin, 1992) or by the bettors themselves.

5. Note that, at this point, U_{IF} is strictly positive.

6. Depending on the value of their stake limit, there may be a point at which the favourite is backed to such an extent that the utility gained from both bets is equal and both U_{IF} and U_{IL} are positive. Informed traders will then continue to wager on both horses until the stake limit is reached.

7. The process of odds setting by bookmakers over the life of the market is explored in Vaughan Williams and Paton (1997b).

8. In handicap races weights are allocated by the handicapper so as to equalize as far as possible each horse's winning chances. In higher-grade handicaps form is particularly well-established and accessible.

9. For details on the Tobit estimator see Greene (1993), pp. 691-701.

10. The bookmaker's price for each horse is given by 1/(1+SP). The over-round in an n horse race (in percentage terms) is then:

$$\sum_{i=1}^{n} \frac{1}{(1 + SP)}$$

CHAPTER TEN

SUMMARY AND CONCLUSIONS

10.1 Information Efficiency: Concepts, Definitions and Tests

A financial market is generally characterized as 'informationally efficient' if it incorporates all available information. The implication of the existence of such efficiency is that no actor within the market possesses information which is not already included in the prices obtaining about the various market assets. As a consequence, no-one is able to secure a superior return from trading on perceived differences in the actual price of an asset and its 'true' price, except chance or as compensation for accepting some other positive cost, such as above-average risk. The idea that ALL information, including privately or monopolistically held information, is instantaneously incorporated into market prices implies that no-one, not even 'insiders', can know more than is already revealed by the behaviour of market variables. This is usually referred to as 'strong form efficiency.' A less strict interpretation of strong form efficiency holds that although insiders may know more than the market instantaneously reveals, they are not able to make above-average returns from this extra information except as compensation for additional costs. The idea that all publicly available information, though not private information, is incorporated into market prices is usually termed 'semi-strong form efficiency.' 'Weak form efficiency' is the phrase normally used to indicate that all the information available from the history of relevant market prices and price movements is incorporated into present market prices. As above, a less strict interpretation of semi-strong and weak-form efficiency simply holds that it is not possible for actors within the market to secure above-average returns on the basis of semi-strong and weak-form information respectively, except as compensation for additional costs.

A paradox exists in the strictest interpretations of information efficiency. This arises from an examination of the process by which information efficiency comes to exist in the first place - through traders identifying disparities between the price of an asset and the 'value' of the asset and buying/selling such assets in order to realize profits. If traders know or believe that no such disparities exist - that the market already incorporates all information - then no trader will act so as to close the disparities. In such a situation the disparities will remain and the market will not be informationally efficient in the strictest sense. No such paradox exists if an above-average return is allowed to traders as compensation for their skill, knowledge and time. This can be viewed as a 'normal profit'.

Thus for given costs and risk, the expected return to all traders in a strictly informationally efficient market will be identical, and expected profits will be normal. The actual returns may vary and actual profit may be greater or less than this, but it will be due to chance. This is because all relevant information is already incorporated into the market. In a market characterized by information efficiency less strictly defined, there may be less than full incorporation of relevant information but this is only sufficient, for given risk, to permit fair compensation to traders for their skill, information and time. In the least strict sense of information efficiency, above-average returns can be made on the basis of existing information, but these are not sufficient to cover the transactions costs of acting upon the information.

Empirical tests of information efficiency can be broadly sub-divided into tests of the weak, semi-strong and strong form types of efficiency referred to above.

Tests of weak form efficiency in financial markets generally seek to identify whether any trends can be identified in the series of historical price data through time, and whether profits can be made from acting upon any such trends. Serial correlation tests of price dependence, variance ratio tests, cointegration analyses and most recently, rescaled-range analysis have been applied to this problem. Although the application of such tests has revealed some evidence of price dependence it is much less clear that such dependence can be translated into a profitable trading strategy.

Tests of semi-strong form efficiency in financial markets often take the form of a quest to identify systematic differences in returns to particular types of asset or to identical assets at particular times or in particular circumstances. Examples of 'market anomalies' thus far proposed are the 'January effect' (that shares in January demonstrate evidence of consistent above-average returns) and the 'small firm effect' (that shares of small firms tend to out-perform those of larger firms). Although there is some convincing evidence favouring the present or past existence of some such 'anomalies' it is far less clear that the above-average returns are not simply compensation for additional risk associated with the securities exhibiting such behaviour. Other tests of semi-strong form efficiency involve an analysis of the way in which new public information is incorporated into asset prices. While such studies tend to suggest that the market does not always adjust fully and instantaneously to new public information, there is far less evidence that it is possible to exploit this so as to earn abnormal returns.

Tests of strong form efficiency take two forms. The first is to assess the impact of identifiable monopolistic access to information and assess the impact of this insider knowledge on profitability. The second is to assess the performance of individuals and organisations, such as professional forecasting services, in order to assess whether they have access to private information not reflected in stock prices. The first sort of test is beset by the difficulties of identifying trading which can in certain contexts be illegal. There is some evidence to suggest that informational monopoly power can and has been used to make a systematic abnormal return, although there is less convincing evidence that outsiders in conventional financial markets can benefit in the same way from following the behaviour of insiders. The second sort of test is less clearly a test of strong form efficiency inasmuch as the information, although privileged and often costly to obtain, is publicly if not always widely disseminated. There is, therefore, a problem of distinguishing the relevant information domain and thus the form of efficiency which is being tested. Whichever, there is evidence that some professional forecasting services, at varying levels of public information disclosure, have significantly out-performed the market. However, there is less consensus on the reason for this performance, there being at least some indication that it results from the type of securities (usually more volatile) chosen.

10.2 Information efficiency in betting markets

In racetrack betting markets, tests of information efficiency have tended to concentrate on whether there exist significant differences in the expected return to wagers placed on those possible outcomes judged by the market to possess a high likelihood of occurring compared to those judged to possess a low likelihood of occurring. This provides a test of whether a simple trading rule exists, based on betting at particular identified odds, which would produce significantly higher expected returns than betting at other particular identified odds. There is substantial evidence in the literature for the existence of just such a rule in racetrack betting markets, at least for the U.K. and the U.S. In particular, betting level stakes in the lower range of odds tends in various studies to produce a significantly higher return than betting the same stakes in the higher ranges of odds. This phenomenon is usually known as the 'favourite-longshot bias.' This thesis generates evidence in support of the existence of this bias in expected returns, employing a new data set composed of recent British horse races.

Other studies have demonstrated variations in the bias at different points of the racing day, such as a strengthening of the favourite-longshot bias in later races. These findings are not confirmed in the present study. The general existence of a longshot bias at all, however, not only seems to violate one definition of weak form efficiency, but also requires explanation in terms of rational economic behaviour.

A number of explanations of the longshot bias adopt the idea of bettors as utility-maximizers rather than simple profit-maximizers, or as risk-lovers, while other proposals include the idea that bettors discount a fixed fraction of their losses. Recent contributions have divided the market into "uninformed" and "informed" bettors, who are distinguished by their ability to distinguish the true probabilities about the various outcomes. In the presence of positive transactions and information costs, optimizing and game-theoretic models have been proposed which can also, subject to certain assumptions, explain this bias. All of these are demand-side explanations of the bias. An alternative approach has been to explain the favourite-longshot bias instead as an optimal supply-side response by bookmakers in a market characterized by adverse selection, in which the bookmaker faces a number of bettors, of uncertain number and identity, who possess superior information. In the simplest modellings, such insiders are assumed to know the outcome in advance with certainty, and bookmakers shorten the odds of longshots because of the disproportionate expected losses at higher odds. One implication of this approach is a link between the number of runners in a race (and therefore the average size of the odds about the runners) and the bookmakers' margin (or over-round), as implied in the odds (or over-round). Because these regularities can be explained by both demand-side and supply-side explanations, however, in this study new tests are proposed to identify and distinguish the influence of insider trading. The results of these tests, which examine subsets of the data expected to be characterized by varying levels of insider activity, tend to support the view that the supply-side hypothesis can explain at least some of the bias observed in fixed-odds British racetrack betting markets. This thesis also proposes and tests a new model of betting market behaviour which assumes that insiders know the true objective probabilities of the various outcomes, but not the outcome itself. This model is capable of explaining the existence of a favourite-longshot bias in British fixed-odds racetrack betting markets, as well as interpreting the dynamics of the odds-setting process over the course of the market. A final model is proposed which seeks to reconcile the existence of the conventional bias in most betting markets with some studies which have either failed to reproduce its existence at all (e.g. studies of Hong Kong pool systems of betting), or else demonstrated evidence of a contrary bias (e.g. for U.S. fixed-odds baseball markets). The model includes an assumption that bettors can derive utility from betting, and also incorporates the influence, noted by recent writers, of information/transactions costs on the bias.

Other tests of weak-form efficiency include the various analyses of 'technical systems' of betting available in the academic literature. Such systems employ and utilize the information contained in current betting odds and the patterns in such odds. There is no clear evidence, however, that this information can be used to generate a strategy of win bets which can make significant abnormal profits. Although there is more evidence of success in generating such rules in the market for 'place' and 'show' bets in identified parimutuel markets, some of this has disappeared under close re-examination, and other systems are reliant on an ability to operationalize complex decision-making procedures in limited real time.

The existence of semi-strong form efficiency in a market would imply that no patterns in the returns should be identifiable which could indicate the existence of superior returns from betting in certain clearly defined circumstances. There should, for example, be no systematic tendency to over-react to recent information (a 'gambler's fallacy), or to earn superior returns on certain days of the week (a 'calendar effect'), or else any such patterns should tend to disappear as bettors identify and take advantage of such 'anomalies.'

The literature has generated some empirical support for the existence of a gambler's fallacy, at least in the sense, for example, that a series of wins by favourites (longshots) produces underbetting of favourites (longshots). However, such a hypothesis could not be confirmed in the present study. Nor could any evidence be found of a day-of-the-week 'calendar effect.' These findings of the present study could not therefore be employed as a basis upon which to reject a postulate of information efficiency in recent British racetrack betting markets. Nor was there any evidence in the present data set that rank order of favouritism provides any additional predictive or explanatory power to the odds alone.

Another sort of test involves an assessment as to whether there exists a differential expected return to different bets with identical probabilities of success. Evidence from past studies is mixed. This study looks at the returns to bets about similar or identical outcomes in different types of betting medium. The point is that in a semi-strong efficient market, the returns to identical outcomes in different types of betting medium, should be themselves identical. The co-existence of a bookmaker and tote (parimutuel) system in British on-course racetrack markets provides one such opportunity. Prior published evidence of a significantly superior payout by the tote at higher odds is supported in the present study. However, no systematic method of exploiting these divergences so as to make a profit appears to be possible. It is shown, however, that the divergence disappears when bookmakers' odds have moved significantly during the development of the on-course market. It is speculated that the perception of information inefficiency drawn by bettors from significant odds movements may be capable of explaining the ex-post information efficiency implied by the absence of a differential tote/starting price return.

Another approach is based on the idea that in a semi-strong form efficient market, all relevant publicly available information is impounded instantaneously into prices, and so it is not possible to use public information to earn above-average or abnormal returns. Although convincing evidence has been produced in existing studies that incorporating public information into forecasting models can significantly improve forecasting power, less evidence is available that such an improvement can be used to make abnormal returns. Where such evidence does exist the strength of the findings is either linked to the ability to operationalize a complex model in limited real time, or else subject to variations in the return which may permit extended short-term losses. There is also the problem confronting bettors seeking to operationalize a strategy that the findings will already have been incorporated into future odds.

Another test of semi-strong form efficiency, proposed and adopted in this study, was to compare the Tote Dual Forecast with the bookmaker-run Computer Straight Forecast on the basis of the new data set. The idea behind this test is that since these two types of bet are comparable options about a similar event, the actual and expected degree of convergence in the returns can be compared. A significant problem with this approach was that of comparing returns about predictive outcomes which are not, however, identical (predicting 1-2 in correct order with 1-2 in any order). This issue was addressed, and on the basis of the evidence available it was not possible to reject the hypothesis that the betting market was semi-strong form efficient.

As with stock market forecasting, the extent to which racetrack forecasting services can produce above-average returns is a test of information efficiency. The form of information which is being tested relies ultimately on the level of privilege at which the information can be accessed. Newspaper tips and price forecasts are clearly publicly available, and while there is evidence that genuine information is contained in forecast prices there is less evidence that such information is unincorporated into the payout odds available in the market. There is even less evidence that any such unincorporation can be systematically exploited so as to yield above-average or abnormal returns.

A more clear test of strong efficiency is whether bets placed at the best odds or earlier odds would yield abnormal returns. This is a way of inferring the existence of returns to insider trading. Whether outsiders can benefit from this information depends upon the possibilities for identifying such information and acting upon it before the information is itself incorporated into the odds. Although some evidence exists that this is possible in a limited number of cases in U.K. bookmaking markets, it is questionable whether the extra returns are sufficient to cover the extra costs incurred in obtaining and operationalizing the information. More promising indications have been reported for Australian horse race betting markets. In such markets there is some evidence that outsiders who identify insider knowledge through movements in bookmakers' odds can use this information to secure profits from arbitrage opportunities on the tote. If so, this would offer prima facie testimony to the existence of both strong and semi-strong form inefficiency in such markets. There has, however, been no evidence produced in the academic literature to support the proposition that significant abnormal returns can be earned from following the advice of professional forecasting services, and the limited evidence that abnormal returns can be gleaned from forecasting rules tends to disappear when tested on data sets covering periods subsequent to the publication of the rules. A detailed analysis undertaken in this study of five modern professional racetrack forecasting services does, on the other hand, reveal some evidence of profitability (at least pre-tax), although the profits could not be confirmed in any particular case at conventional levels of significance.

The case for information efficiency in racetrack betting markets has thus not been disproved, at least in the sense of an opportunity to earn systematic abnormal expected returns, except perhaps at the level of insiders acting upon monopolistically held private information.

10.2.1 Information efficiency in betting markets: conclusions

In racetrack betting markets, tests of weak form information efficiency, i.e. efficiency with respect to the set of historical and current prices, have tended to concentrate on whether there are significant differences in the expected return to wagers placed on those possible outcomes judged by the market to possess a high likelihood of occurring compared to those judged to possess a low likelihood of occurring. The evidence, at leat for the U.K. and U.S. tends to suggest that betting level stakes in the lower range of odds produces a significantly higher return than betting the same stakes in the higher ranges of odds. This phenomenon is usually known as the 'favourite-longshot bias.'

A number of explanations of the longshot bias adopt the idea of bettors as utility-maximizers rather than simple profit-maximizers, or as risk-lovers, while other proposals include the idea that bettors discount a fixed fraction of their losses. Recent contributions have explained the outcome as the equilibrium position of a game played according to identified rules. All of these are demand-side explanations of the bias. An alternative approach has been to explain the favourite-longshot bias instead as an optimal supply-side response by bookmakers in a market characterized by adverse selection, in which the bookmaker faces a number of bettors, of uncertain number and identity, who possess superior information. In the simplest modellings, such insiders are assumed to know the outcome in advance with certainty, and bookmakers shorten the odds of longshots because of the disproportionate expected losses at higher odds. One implication of this approach is a link between the number of runners in a race (and therefore the average size of the odds about the runners) and the bookmakers' margin (or over-round), as implied in the odds. In the absence of any convincing alternative account, however, there is little reason to believe that betting markets are weak form inefficient, at least in the sense of offering outsiders opportunities to earn abnormal returns. Rather, a number of explanations have been offered which can explain the observed biases in expected returns at different odds in the context of a market composed of rational players. Common to these explanations is a market which can efficiently process the information available to it.

The existence of semi-strong form efficiency in a market would imply that no patterns in the returns should be identifiable which could indicate the existence of superior returns from betting under certain clearly defined circumstances. One type of test is to assess whether there exists a differential expected return to identical outcomes in different identified types of betting medium, another whether there exists any 'anomalous' behaviour in the form of systematic exploitable patterns in the history of publicly available information. Although there is some evidence of unexploited opportunities for arbitrage between different sectors of the betting market, at least in the U.K., and evidence that bettors tend to under-estimate the likelihood of consecutive identical outcomes, there is much less (though there is some) evidence that the use of publicly available information, in whatever form, can be used to earn abnormal returns.

There is much stronger evidence for the suggestion that insiders possess valuable information unavailable to the public, which they can trade upon so as to earn abnormal returns, and to this extent the market may be considered informationally inefficient. In a sense, however, this is to confuse the ability of the market to process information efficiently with the withholding of information from the market itself. It is only in the latter sense that there is clear evidence of information inefficiency in betting markets.

The case for information efficiency in racetrack betting markets has thus not been disproved (although neither has it been proved), at least in the sense of an opportunity to earn systematic abnormal expected returns, except perhaps at the level of and in the sense of insiders acting upon monopolistically held private information.

Dowie (1976) termed "... a market as efficient to the extent that it passes the weak and semi-strong tests and equitable to the extent that it passes the strong test." (p.140).

In these terms, the weight of evidence supports the view that betting markets are inequitable, although it is not so clear that they are inefficient.

APPENDIX 1

Constructing a new data set of characteristics relevant to modern British racetrack betting markets.

1. The finishing position of the horse in each race covered.

Where two or more horses dead-heat, the order is taken as listed in *The Sporting Life Flat Results* (1992, 1993). No attempt is made to apply fractional positions, such as allocating a 2.5 placing to two horses dead-heating for second place. Given the small number of cases, the additional analytical complexities associated with such a procedure are not considered warranted. Where a horse is disqualified, it is placed lasat for the purposes of this study.

This variable is denoted by the column heading "PO".

2. A race number is allocated to each race covered.

Each race in the season is allocated a unique number in *The Sporting Life Flat Results* (1992) publication. This number is used to identify races in this data set. An additional 00 is added to the end of the race numbers found in *The Sporting Life Flat Results* (1993). Thus, 390 is the Burghclere stakes at Newbury (April 11th, 1992), a race designated race 390 in the 1992 publication. A race designated 390 in the 1993 publication is here allocated the number 39000. The additional 00 added to the end of the 1993 season's races ensures that no two races are ever allocated the same race number in this data set.

Denoted by "RC".

3. A number is allocated to each race to indicate the order of that race in the daily running order of all races at the racecourse in question.

The numbers run from 1 (first race of the day) to 6 (last race of the day) on some occasions, from 1 (first race of the day) to 7 (last race of the day) on other occasions. In unusual cases, less than six or more than seven races are run. Such circumstances are excluded.

Denoted by "NO".

4. The starting price of each horse in every race.

These are presented as the odds to 1 against, e.g. 3 indicates odds of 3 to 1 against, 0.5 indicates 1 to 2 against (2 to 1 on).

Denoted by "SP".

5. The opening price of each horse in every race.

This indicates the price about each horse at the first show in the market. This price is recorded in the *Sporting Life*, and is the general price at which it is estimated a sizeable bet could have been placed on a horse at the opening of a settled on-course market.

Presented as for starting prices. Denoted by "OP".

6. The best price available about each horse in every race. This is derived from an examination of all odds offered during the course of market trading, as reported in the *Sporting Life Flat Results* (1992, 1993). These "best prices" are not always generally available in the market. Presented as for starting prices.

Denoted by "BP".

7. The worst price available about each horse in every race. This is derived from an examination of all odds offered during the course of market trading, as reported in the *Sporting Life Flat Results* (1992, 1993). It is the lowest of the three prices shown above, i.e. the starting price, the opening price and the best price. Presented as for starting prices.

Denoted by "WP".

8. The forecast price of each horse in every race.

This is derived from the forecast odds provided in the *Sporting Life* newspaper every morning. Where not all horses are quoted, a single price is allotted to all horses which are not individually quoted. This is normally known as the "Bar Price." Presented as for starting prices.

Denoted by "FP".

9. The prize money associated with each race.

This is the prize money in pounds sterling allocated to the winner of the race.

Denoted by "PR".

10. The distance of each race.

This is quoted in furlongs. Where necessary it is rounded to the nearest furlong.

Denoted by "DI".

11. The ages of the horses competing in each particular race. Where all the horses are of the same age the number of years of age is quoted directly. Where horses of different ages are competing in the same race, a value of 0.5 is added to the minimum age requirement. A race for horses aged 3 and up, for example, is indicated by the number 3.5.

Denoted by "AG".

12. The number of runners in the race.

This covers all horses about which a starting price is reported, and includes, therefore, horses disqualified after officially starting, but excludes pre-race withdrawals.

Denoted by "RN".

13. The type of race.

All races are categorized into a class description, e.g. Listed races are categorized as 4, selling stakes as 23, handicaps as 50. For these purposes, the full list of classifications is represented as follows:

Number	Type of race
1	Group 1
2	Group 2
3	Group 3
4	Listed
10	Maiden Stakes
11	Maiden Auction Stakes
11.8	Maiden Auction Stakes (Fillies)
12	Maiden Claiming Stakes
13	Maiden Selling Stakes
17	Maiden Apprentice Stakes
18	Maiden Fillies Stakes
19	Maiden Amateur Stakes
20	Stakes

21	Auction Stakes
22	Claiming Stakes
23	Selling Stakes
24	Conditions Stakes
25	Graduation Stakes
26	Sweepstakes
27	Apprentice Stakes
27.2	Apprentice Claiming Stakes
28	Fillies Stakes
28.5	Fillies Graduation Stakes
29	Amateur Stakes
30	Nursery
31	Auction Graduation Stakes
35	Shield
36	Trophy
38	Nursery (Fillies)
39	Challlenge Whip
50	Handicap
53	Selling Handicap
Number	**Type of race**
53.3	Nursery Selling Handicap
56	Handicap Sweepstakes
57	Apprentice Handicap
57.3	Apprentice Selling Handicap
58	Fillies Handicap
59	Amateur Handicap

Denoted by "TY".

14. The month of the year in which the race took place.

Represented by order of month in the year,

e.g. January = 1 to December = 12.

Denoted by "MO".

15. The day of the month in which the race took place.

Represented by order of day in the week,

e.g. Monday = 1 to Sunday = 7.

Denoted by "DA".

16. The date of the month in which the race took place.

Represented by date of the month,

e.g. 21st. January = 21.

Denoted by "DT".

17. The official "going", or state of the ground, for the race.

These official classifications range in the U.K. from 1 (hard) to 7 (heavy). The full list is:

Number	**Official state of the "going"**
1	Hard
2	Firm
3	Good to Firm
4	Good
5	Good to Soft
6	Soft
7	Heavy

Denoted by "GO".

18. The forecast going, as listed in the Sporting Life on the morning of the race.

As above, ranges from 1 (hard) to 7 (heavy).

Denoted by "FG".

19. The racecourse.

Each racecourse is allocated a four-letter code, comprising the first four letters of its full name, e.g. donc for Doncaster. The only exception is Ayr, the code for which is, naturally, ayr.

Racecourse	Code
Ascot	asco
Ayr	ayr
Bath	bath
Beverley	beve
Brighton	brig
Carlisle	carl
Catterick	catt
Chester	ches
Doncaster	donc
Epsom	epso
Folkestone	folk
Goodwood	good
Hamilton	hami
Haydock	hayd
Kempton	kemp
Leicester	leic
Lingfield	ling

Newbury	newb
Newcastle	newc
Newmarket	newm
Nottingham	nott
Pontefract	pont
Redcar	redc
Ripon	ripo
Thirsk	thir
Salisbury	sali
Sandown	sand
Warwick	warw
Windsor	wind
York	york

Denoted by "CR".

20. The over-round.

The over-round is the proportion of a given sum over and above that sum which it would be necessary, given the prevailing odds, to bet in order to guarantee the return of that sum. If, for example, the odds offered about five horses are each 2 to 1, then £1-60 would have to be wagered in order to guarantee a return of £1, i.e. an over-round of 60 per cent (also referred to as an over-round of 160 per cent). Given that it is necessary to bet £1/(1+X) to obtain a return of £1 at odds of X, the over-round is equal to $\Sigma 1/(1+X)-1$. The latter convention, i.e. including an initial 100 per cent is used for these purposes).

Denoted by "OR".

21. The grade of handicap, where applicable.

Racehorses are allocated a rating by the official handicapper which varies according to performance. A handicap for horses rated 0 to 70 for instance restricts the race to horses with a mediocre to poor performance record. Such a handicap is coded for these purposes as 70. The range in this data set extends from 60 at the lower end to 115 at the upper end. Where the handicap is not classified numerically but rather graded by letter, the following convention is adopted for these purposes. Grade A is coded as 111, grade B as 101, and grade C as 91. The prestigious Lincoln Handicap run at Doncaster, and the European Free Handicap run at Newmarket are not classified conventionally and are accorded here a coding of 121. A coding of 999 signifies that the race in question is not a handicap.

Denoted by "HC".

22. The "going correction", as listed in the Sporting Life Flat Results (1992, 1993).

This is an estimate of the ground condition, assessed in seconds per furlong, which takes account of race times compared with normal or standard times, wind and other conditions. Ground conditions which favour slow running times will thus be corrected by a positive going correction, e.g. 0.8 seconds per furlong. A correction of, say, minus 0.5 seconds per furlong would indicate the opposite.

Denoted by "GA".

23. The tote odds returned to a bet placed on the winning horse.

The return is quoted inclusive of a unit stake. A tote payout of 3.5, for example, indicates effective odds of 2.5 to 1 against.

Denoted by "TO".

24. The order of the race on the racecard, "f" indicating the first race on the card, "p" the penultimate race, "l" the last race, and "m" all other races.

This method of classification permits direct comparison of studies from race meetings characterized by a varying number of races on the card.

Denoted by "TI".

25. The finishing position of the favourite, i.e. the horse or horse with the shortest starting price(s).

Where two or more horses shared favouritism, an average of the finishing positions is taken.

Denoted by "FA".

26. The finishing position of the favourite at the opening show, as indicated by the shortest opening price in the market.

Where two or more horses shared opening favouritism, an average of the finishing positions is taken.

Denoted by "OF".

27. The early odds offered about each horse running on a particular day (where applicable) by the William Hill bookmaking organisation, as reported in the Sporting Life on the day of the race.

Where no odds are offered about a particular horse the code 999 is employed.

Presented as for starting prices. Denoted by "WH".

28. The early odds offered about each horse running on a particular day (where applicable) by the Corals bookmaking organisation, as reported in the Sporting Life on the day of the race.

Where no odds are offered about a particular horse the code 999 is employed.

Presented as for starting prices. Denoted by "CO".

29. The early odds offered about each horse running on a particular day (where applicable) by the Ladbrokes bookmaking organisation, as reported in the Sporting Life on the day of the race.

Where no odds are offered about a particular horse the code 999 is employed.

Presented as for starting prices. Denoted by "LA".

30. The starting price of the favourite in each race. Denoted by "PF".

31. The starting price of the second favourite in each race. Denoted by "PS".

32. The finishing position of the forecast favourite, as indicated by the shortest forecast starting price offered in the Sporting Life on the day of the race.

Where there are two or more forecast favourites, an average of their finishing positions is taken.

Denoted by "FF".

33. The finishing position of the bookmakers' favourite, as indicated by the early odds offered about each horse in a race by the three major bookmaking organisations, i.e. William Hills, Corals and Ladbrokes.

These are as published in the Sporting Life on the day of the race. Where there is a conflict, the horse offered by the majority is taken to be the bookmakers' favourite. Where there is no majority, an average of the finishing positions of the various horses favoured by the bookmakers is taken.

Where this information is applicable because no early odds are offered about a particular horse the code 99 is employed.

Denoted by "BF".

34. Where not all horses are quoted in the Sporting Life forecasts of starting prices, a general price is allotted race by race to horses not so quoted.

This is usually known as the Bar Price. This price is recorded here.

Presented as for starting prices. Denoted by "BP".

35. The position of the race winner in order of favouritism. If the race winner started the race as third favourite, for example, this would be presented as 3, if as fifth favourite, as 5. If the winner shared some position of favouritism, e.g. two horses sharing third favouritism, this is presented as the mean of the adjoining positions, i.e. 3.5 in this case.

Denoted by "WI".

36. The proportion of the returns to a winning bet deducted by bookmakers when a horse or horses does not run, but where bets have been placed in advance of the start on the basis that they are runners.

This is determined by Tattersall's Committee Rule 4(c). A deduction of 0.2 implies, for example, that 20 per cent of the net returns calculated at the taken or starting price are retained by the bookmaker. 4(C) of Tattersall's Committee Rules on Betting states:-

In the case of bets made at a price on the day of the race before it has been officially notified that a horse has been withdrawn before coming under Starter's Order, the liability of a layer against any horse remaining in the race, win or place, will be reduced in accordance with the following scale depending on the odds current against the withdrawn horse at the time of such official notification.

(a). if the current odds are 30/100 or longer odds on, by 75p. in the £.

(b). if shorter odds than 30/100 up to and including 2/5, by 70p. in the £.

(c). if shorter odds than 2/5 up to and including 8/15, by 65p. in the £.

(d). if shorter odds than 8/15 up to and including 8/13, by 60p. in the £.

(e). if shorter odds than 8/13 up to and including 4/5, by 55p. in the £.

(f). if shorter odds than 4/5 up to and including 20/21, by 50p. in the £.

(g). if shorter odds than 20/21 up to and including 6/5, by 45p. in the £.

(h). if over 6/5 up to and including 6/4, by 40p. in the £.

(i). if over 6/4 up to and including 7/4, by 35p. in the £.

(j). if over 7/4 up to and including 9/4, by 30p. in the £.

(k). if over 9/4 up to and including 3/1, by 25p. in the £.

(l). if over 3/1 up to and including 4/1, by 20p. in the £.

(m). if over 4/1 up to and including 11/2, by 15p. in the £.

(n). if over 11/2 up to and including 9/1, by 10p. in the £.

(o). if over 9/1 up to and including 14/1, by 5p. in the £.

(p). if over 14/1 the liability would be unchanged.

(q). in the case of two or more horses being withdrawn the total deduction shall not exceed 75p. in the £.

Ante-post bets are not affected and S.P. bets are also not affected, except in cases where insufficient time arises for a fresh market to be formed, when the above named scale of reductions will apply.

Denoted by "DE".

37. An alternative measure of the returns to a winning bet which should be deducted by bookmakers when a horse or horses does not run, but where bets have been placed in advance of the start on the basis that they are runners.

This measure is based on the likelihood that the withdrawn horse(s) would win based on the implied probabilities contained in the odds.

Denoted by "NV".

38. The Dual Forecast odds, i.e. the odds returned by the 'tote' for predicting the first and second placed horses in a given race, in any order.

The return is quoted inclusive of a unit stake. A payout of 12.5, for example, indicates effective odds of 11.5 to 1 against.

Where no dual forecast is available the code 999999 is used.

Denoted by "DF".

39. The Computer Straight Forecast, i.e. the odds returned by bookmakers for predicting the first and second placed horses in a given race in the correct order.

Calculated according to a pre-determined formula.

Denoted by "CS".

40. The place odds returned by the 'tote' about the winning horse.

Where no place odds are available the code 99 is used.

The return is quoted inclusive of a unit stake.

Denoted by "TA".

41. The place odds returned by the 'tote' about the horse finishing second, where applicable.

The return is quoted inclusive of a unit stake.

Where no place odds are available the code 99 is used.

Denoted by "TB".

42. The place odds returned by the 'tote' about the horse finishing third, where applicable.

The return is quoted inclusive of a unit stake.

Where no place odds are available the code 99 is used.

Denoted by "TC".

43. The place odds returned by the 'tote' about the horse finishing fourth, where applicable.

The return is quoted inclusive of a unit stake.

Denoted by "TD".

44. The Tricast odds returned by the 'tote'.

The Tricast odds are the odds returned by the 'tote' for predicting the first, second and third placed horses in a given race, in any order.

The return is quoted inclusive of a unit stake.

Where no tricast odds are available the code 999999 is used.

Denoted by "TR".

45. The number of horses contained in the Sporting Life 'Bar' about each particular race,

i.e. the number of horses about which no individual odds are forecast in that race.

Denoted by "BA".

46. The draw of the horse.

This is the number of the randomly selected starting stall, where applicable, from which a horse must start the race.

Denoted by "DR".

47. The jockey.

Normally indicated by the first four letters of the surname, e.g. dett for dettori. Where there is a possible confusion the first letter of the first name is added on to the end. Pat Eddery (eddp) is distinguished from Paul Eddery (eddpa) as shown in the parentheses.

Denoted by "JO".

48. The trainer.

The coding is as for variable 47 (above).

Denoted by "TN".

49. The age of the horse.

This is the official age of the horse for racing purposes (all racehorses are deemed to be born on January 1 of the year in which they are born).

Denoted by "OL".

50. The weight carried by the jockey.

This is coded from 700 (7 stones) to 1000 (10 stones). Other weights are described by the first number (number of stones) followed by the number of pounds additional to this, e.g. 707 indicates 7 stones 7 pounds, 812 indicates 8 stones 12 pounds.

Denoted by "WT".

51. The number of pounds the jockey is allowed to deduct from the weight allocated to the horse, i.e. the jockey's claim.

This varies from 7 to 5 to 3 to 0, according to the number of wins a jockey has secured in specified types of horse race.

Denoted by "CL".

52. The number of pounds a horse is carrying excess of the weight allocated to it, where applicable.

Denoted by "OW".

53. A measure of the movement in the betting odds, as reported in The Sporting Life Flat Results (1992).

This is coded from 1 when the odds uniquely lengthen

from the opening show to 6 where the odds movement changes direction from the opening show more than once.

"1" - indicates a simple lengthening of the odds from opening to starting price.

"2" - indicates a simple shortening of the odds from the opening to the starting price.

"3" - indicates a lengthening followed by a shortening of the odds from the opening to the starting price.

"4" - indicates a shortening followed by a lengthening of the odds from the opening to the starting price.

"5" - indicates no movement in the odds between the first show and the returned starting price.

"6" - indicates a change in the direction of movement on more than one occasion from the opening to the starting price.

Denoted by "MV".

54. A measure of the movement in the televised betting odds.

Coded as for variable 53 (above).

Where no information is collected or available the code 99 is used.

Denoted by "TM".

55. The opening price quoted about the opening favourite, i.e. the horse with the shortest opening price.

Presented as for starting price.

Denoted by "YF".

56. The opening price of the second favourite in each race. Presented as for starting price.

Denoted by "YS".

57. The forecast price quoted about the forecast favourite.

i.e. the horse with the shortest forecast starting price (as quoted in the Sporting Life).

Presented as for starting price.

Denoted by "ZF".

58. The Forecast Price of the forecast second favourite in each race.

Presented as for starting price.

Denoted by "ZS".

APPENDIX 2

TABLE 2A Returns from bets at given SP

(For all races in periods 1, 2 and 3 combined)

SP	Win ratio	Rate of Return (%)	Subjective lose Probability (Q)	Objective Lose Probability (q)
0.111	1 out of 1	111.1	0.1	0
0.143	1 out of 1	114.3	0.125	0
0.167	0 out of 1	0	0.143	1
0.20	4 out of 4	120	0.167	0
0.222	2 out of 2	122.2	0.182	0
0.25	1 out of 2	62.5	0.2	0.5
0.286	2 out of 2	128.6	0.222	0
0.3	3 out of 3	130	0.231	0
0.333	3 out of 7	57.13	0.25	0.571
0.364	4 out of 5	109.12	0.267	0.2
0.4	11 out of 12	128.33	0.286	0.083
0.444	8 out of 10	115.52	0.307	0.2
0.5	7 out of 10	105	0.333	0.3
0.533	0 out of 3	0	0.348	1
0.571	12 out of 17	110.89	0.363	0.412
0.615	12 out of 19	102	0.381	0.368
0.667	14 out of 25	93.35	0.4	0.44
0.727	12 out of 33	62.8	0.421	0.636
0.8	16 out of 23	125.22	0.444	0.304
0.833	8 out of 13	112.8	0.454	0.385
0.909	15 out of 30	193.33	0.476	0.5
1	29 out of 59	48.53	0.5	0.508
1.05	0 out of 1	0	0.512	1
1.1	18 out of 40	94.5	0.524	0.55
1.2	7 out of 16	96.25	0.545	0.563

1.25	15 out of 45	75	0.556	0.667

SP	**Win ratio**	**Rate of Return (%)**	**Subjective lose Probability (Q)**	**Objective Lose Probability (q)**
1.375	24 out of 59	96.61	0.579	0.593
1.5	30 out of 80	93.75	0.6	0.625
1.625	30 out of 61	129.10	0.619	0.508
1.75	32 out of 93	94.62	0.636	0.656
1.875	17 out of 48	101.83	0.652	0.646
2	54 out of 153	105.88	0.667	0.647
2.125	0 out of 3	0	0.68	1
2.25	53 out of 173	99.57	0.692	0.694
2.5	33 out of 174	66.38	0.714	0.81
2.75	40 out of 155	96.77	0.733	0.742
3	62 out of 263	94.30	0.75	0.764
3.333	26 out of 103	109.38	0.769	0.748
3.5	57 out of 338	75.89	0.778	0.831
4	65 out of 387	83.98	0.8	0.832
4.5	64 out of 364	96.70	0.818	0.824
5	89 out of 545	97.98	0.833	0.837
5.5	33 out of 323	66.41	0.846	0.898
6	59 out of 509	81.14	0.857	0.884
6.5	31 out of 256	90.82	0.867	0.879
7	61 out of 553	88.25	0.875	0.890
7.5	10 out of 143	59.44	0.882	0.930
8	71 out of 773	82.66	0.889	0.908
8.5	2 out of 49	38.78	0.895	0.959
9	48 out of 514	90.27	0.9	0.907
10	70 out of 904	85.18	0.909	0.923
11	14 out of 297	56.57	0.917	0.953
12	46 out of 947	63.15	0.923	0.951

14	38 out of 1092	52.20	0.933	0.965
16	44 out of 1080	69.26	0.941	0.959
SP	**Win ratio**	**Rate of Return (%)**	**Subjective lose Probability (Q)**	**Objective Lose Probability (q)**
20	34 out of 1240	57.58	0.952	0.973
22	0 out of 9	0	0.957	1
25	22 out of 1001	57.14	0.962	0.978
28	0 out of 6	0	0.966	1
33	9 out of 1572	19.47	0.971	0.994
40	2 out of 99	82.83	0.976	0.980
50	5 out of 792	32.30	0.980	0.994
66	2 out of 228	58.77	0.985	0.991
80	0 out of 2	0	0.988	1
100	0 out of 153	0	0.990	1
125	0 out of 1	0	0.992	1
150	0 out of 8	0	0.993	1
200	0 out of 20	0	0.995	1
250	0 out of 2	0	0.996	1
400	0 out of 0	0	0.998	-
500	0 out of 1	0	0.998	1

Table 2B **Returns from bets at given SP**
(For all races in Period 1)

SP	Win ratio	Rate of return (%)	Subjective lose Probability (Q)	Objective lose Probability (q)
0.111	0 out of 0	-	0.1	-
0.143	0 out of 0	-	0.125	-
0.167	0 out of 0	-	0.143	-
0.20	2 out of 2	120	0.167	0
0.222	0 out of 0	-	0.182	-
0.25	0 out of 0	-	0.2	-
0.286	0 out of 0	-	0.222	-
0.3	1 out of 1	130	0.231	0
0.333	3 out of 4	133.33	0.250	0.25
0.364	1 out of 1	136.4	0.267	0
0.4	3 out of 3	140	0.286	0
0.444	0 out of 1	0	0.307	1
0.5	2 out of 3	100	0.333	0.333
0.533	0 out of 0	-	0.348	-
0.571	3 out of 3	157.1	0.363	0
0.615	5 out of 8	100.94	0.381	0.375
0.667	3 out of 5	100	0.4	0.4
0.727	5 out of 10	86.35	0.421	0.5
0.8	6 out of 9	120	0.444	0.333
0.833	3 out of 4	137.5	0.454	0.25
0.909	4 out of 12	63.63	0.476	0.667
1	5 out of 17	58.82	0.5	0.706
1.05	0 out of 0	-	0.512	-
1.1	2 out of 8	52.5	0.524	0.75
1.2	1 out of 3	73.33	0.545	0.667
1.25	9 out of 16	126.56	0.556	0.438
1.375	10 out of 21	113.10	0.579	0.524

1.5	11 out of 32	85.94	0.6	0.656
1.625	11 out of 22	131.25	0.619	0.5
SP	**Win ratio**	**Rate of return (%)**	**Subjective lose Probability (Q)**	**Objective lose Probability (q)**
1.75	8 out of 23	95.65	0.636	0.652
1.875	6 out of 23	75	0.652	0.739
2	18 out of 50	108	0.667	0.64
2.125	0 out of 0	-	0.68	-
2.25	19 out of 57	108.33	0.692	0.667
2.5	15 out of 78	67.31	0.714	0.808
2.75	12 out of 57	78.95	0.733	0.789
3	20 out of 85	94.12	0.75	0.765
3.333	11 out of 31	153.75	0.769	0.645
3.5	20 out of 111	81.08	0.778	0.820
4	18 out of 128	70.31	0.8	0.859
4.5	18 out of 123	80.49	0.818	0.854
5	28 out of 180	93.33	0.833	0.844
5.5	10 out of 106	61.32	0.846	0.906
6	24 out of 197	85.28	0.857	0.878
6.5	9 out of 95	71.05	0.867	0.905
7	20 out of 174	91.95	0.875	0.885
7.5	2 out of 45	37.78	0.882	0.956
8	22 out of 274	72.26	0.889	0.920
8.5	0 out of 17	0	0.895	1
9	16 out of 189	84.66	0.9	0.915
10	30 out of 352	93.75	0.909	0.803
11	3 out of 79	45.57	0.917	0.962
12	24 out of 378	82.54	0.923	0.937
14	14 out of 434	48.39	0.933	0.968
16	14 out of 405	58.77	0.941	0.965

20	17 out of 505	70.69	0.952	0.966
22	0 out of 4	0	0.957	1
25	10 out of 415	62.65	0.962	0.976
28	0 out of 2	0	0.966	1
33	2 out of 628	10.83	0.971	0.997

SP	**Win ratio**	**Rate of return (%)**	**Subjective lose Probability (Q)**	**Objective lose Probability (q)**
40	2 out of 25	328	0.976	0.92
50	3 out of 282	54.26	0.980	0.989
66	0 out of 54	0	0.985	1
80	0 out of 1	0	0.988	1
100	0 out of 40	0	0.990	1
125	0 out of 0	-	0.992	-
150	0 out of 4	0	0.993	1
200	0 out of 4	0	0.995	1
250	0 out of 2	0	0.996	1
400	0 out of 0	-	0.998	-
500	0 out of 0	-	0.998	-

Table 2C **Returns from bets at given SP**
(For all races in Period 3)

SP odds	**No of bets winning at these SP odds**	**Rate of return (%)**	**Subjective lose Probibility (Q)**	**Objective lose Probability (q)**
0.111	0 out of 0	-	0.1	-
0.143	0 out of 0	-	0.125	-
0.167	0 out of 1	0	0.143	1
0.20	2 out of 2	120	0.167	0
0.222	0 out of 0	-	0.182	-
0.25	0 out of 0	-	0.2	-
0.286	0 out of 0	-	0.222	-
0.3	1 out of 1	130	0.231	0
0.333	0 out of 1	0	0.250	1
0.364	1 out of 2	136.4	0.267	0.5
0.4	4 out of 4	140	0.286	0
0.444	2 out of 3	96.27	0.307	0.333
0.5	2 out of 3	100	0.353	0.333
0.533	0 out of 1	0	0.348	1
0.571	6 out of 7	134.66	0.363	0.143
0.615	5 out of 6	134.58	0.381	0.167
0.667	1 out of 3	55.57	0.4	0.667
0.727	3 out of 9	57.57	0.421	0.667
0.8	5 out of 6	150	0.444	0.167
0.833	2 out of 3	122.2	0.454	0.333
0.909	3 out of 5	144.54	0.476	0.4
1	8 out of 14	114.29	0.5	0.429
1.05	0 out of 0	-	0.512	-
1.1	5 out of 13	80.77	0.524	0.615
1.2	3 out of 5	132	0.545	0.4
1.25	2 out of 10	45	0.556	0.8
1.375	5 out of 13	91.35	0.579	0.615

1.5	7 out of 16	109.38 0.6		0.563
1.625	3 out of 13	60.58	0.619	0.769
SP odds	**No of bets winning at these SP odds**	**Rate of return (%)**	**Subjective lose Probibility (Q)**	**Objective lose Probability (q)**
1.75	11 out of 27	112.04	0.636	0.593
1.875	4 out of 12	95.83	0.652	0.667
2	11 out of 36	91.67	0.667	0.694
2.125	0 out of 2	0	0.68	-
2.25	14 out of 37	122.97	0.692	0.622
2.5	7 out of 37	66.22	0.714	0.811
2.75	11 out of 33	126.52	0.733	0.667
3	15 out of 65	92.31	0.75	0.769
3.333	8 out of 32	108.33	0.769	0.75
3.5	11 out of 85	58.24	0.778	0.871
4	18 out of 98	91.84	0.8	0.816
4.5	16 out of 93	94.62	0.818	0.828
5	24 out of 137	105.11	0.833	0.825
5.5	5 out of 89	36.52	0.846	0.944
6	23 out of 142	113.38	0.857	0.838
6.5	12 out of 78	115.38	0.867	0.846
7	18 out of 159	90.57	0.875	0.887
7.5	3 out of 42	60.71	0.882	0.929
8	19 out of 239	71.55	0.889	0.921
8.5	0 out of 16	0	0.895	1
9	16 out of 163	88.34	0.9	0.902
10	21 out of 304	75.99	0.909	0.931
11	9 out of 125	86.4	0.917	0.928
12	15 out of 312	62.5	0.923	0.952
14	13 out of 368	52.99	0.933	0.947
16	16 out of 401	67.83	0.941	0.960
20	7 out of 422	34.83	0.952	0.983

22	0 out of 4	0	0.957	1
25	9 out of 332	70.48	0.962	0.973
28	0 out of 4	0	0.966	1
SP odds	**No of bets winning at these SP odds**	**Rate of return (%)**	**Subjective lose Probibility (Q)**	**Objective lose Probability (q)**
33	6 out of 486	41.98	0.971	0.988
40	0 out of 34	0	0.976	1
50	0 out of 260	0	0.980	1
66	1 out of 97	69.07	0.985	0.990
80	0 out of 0	-	0.988	-
100	0 out of 64	0	0.990	1
125	0 out of 0	-	0.992	-
150	0 out of 3	0	0.993	1
200	0 out of 11	0	0.995	1
250	0 out of 0	-	0.996	-
400	0 out of 0	-	0.998	-
500	0 out of 1	0	0.998	1

Table 2D **Average return from a unit bet on horses with SP odds in the given classes (races in periods 1, 2, 3 (combined) for which prize money exceeds the median prize money in each individual period)**

SP	Win ratio	Rate of Return (%)	Subjective lose Probability (Q)	Objective lose Probability (q)
0.111	0 out of 0	-	0.1	-
0.143	0 out of 0	-	0.125	-
0.167	0 out of 0	-	0.143	-
0.20	0 out of 0	-	0.167	-
0.222	0 out of 0	-	0.182	-
0.25	0 out of 1	0	0.2	1
0.286	0 out of 0	-	0.222	-
0.3	0 out of 0	-	0.231	-
0.333	1 out of 1	133.33	0.25	0
0.364	1 out of 1	136.4	0.267	0
0.4	2 out of 2	140	0.286	0
0.444	2 out of 2	144,44	0.307	0
0.5	2 out of 4	112.5	0.333	0.5
0.533	0 out of 2	0	0.348	1
0.571	3 out of 5	94.26	0.363	0.4
0.615	7 out of 12	77.54	0.381	0.417
0.667	7 out of 14	83.35	0.4	0.5
0.727	5 out of 13	66.42	0.421	0.615
0.8	5 out of 7	128.57	0.444	0.286
0.833	4 out of 6	122.2	0.454	0.333
0.909	5 out of 10	75.45	0.476	0.5
1	11 out of 23	96.65	0.5	0.522
1.05	0 out of 1	0	0.512	1
1.1	10 out of 21	100	0.524	0.524
1.2	1 out of 7	31.43	0.545	0.857
1.25	7 out of 17	92.65	0.556	0.588
1.375	10 out of 28	84.82	0.579	0.643

1.5	9 out of 35	64.29	0.6	0.743
SP	**Win ratio**	**Rate of Return (%)**	**Subjective lose Probability (Q)**	**Objective lose Probability (q)**
1.625	17 out of 26	163.96	0.619	0.346
1.75	12 out of 36	91.67	0.636	0.667
1.875	10 out of 24	119.79	0.652	0.583
2	22 out of 67	98.51	0.667	0.672
2.125	0 out of 0	-	0.68	-
2.25	21 out of 83	82.23	0.692	0.747
2.5	16 out of 78	71.79	0.714	0.795
2.75	22 out of 82	99.40	0.733	0.735
3	26 out of 126	82.54	0.75	0.794
3.333	18 out of 66	112.11	0.769	0.727
3.5	28 out of 179	70.39	0.778	0.844
4	33 out of 192	85.93	0.8	0.828
4.5	38 out of 199	105.03	0.818	0.809
5	41 out of 294	83.67	0.833	0.861
5.5	21 out of 194	70.36	0.846	0.892
6	34 out of 283	84.10	0.857	0.880
6.5	19 out of 135	105.56	0.867	0.859
7	36 out of 287	100.35	0.875	0.875
7.5	6 out of 90	56.67	0.882	0.933
8	33 out of 410	72.44	0.889	0.920
8.5	2 out of 28	67.86	0.895	0.929
9	33 out of 285	115.79	0.9	0.884
10	42 out of 488	94.67	0.909	0.914
11	8 out of 173	55.49	0.917	0.954
12	26 out of 517	65.38	0.923	0.950
14	23 out of 549	62.84	0.933	0.958
16	24 out of 573	71.20	0.941	0.958
20	17 out of 594	60.10	0.952	0.971

22	0 out of 9	0	0.957	1
25	11 out of 468	61.11	0.962	0.976
28	0 out of 4	0	0.966	1

SP	**Win ratio**	**Rate of Return (%)**	**Subjective lose Probability (Q)**	**Objective lose Probability (q)**
33	5 out of 657	25.88	0.971	0.992
40	1 out of 37	110.81	0.976	0.973
50	2 out of 259	39.38	0.980	0.992
66	0 out of 61	0	0.985	1
80	0 out of 1	0	0.988	1
100	0 out of 51	0	0.990	1
125	0 out of 0	-	0.992	-
150	0 out of 3	0	0.993	1
200	0 out of 9	0	0.995	1
250	0 out of 2	0	0.996	1
400	0 out of 0	-	0.998	-
500	0 out of 0	-	0.998	-

APPENDIX 3

A model of horse race betting: Deriving some equilibrium conditions

The bookmakers expect the return to a noise trader from a unit bet on the favorite in the presence of suspected insider trading to be:

$$E(ret_f) = \frac{(1-a-B_f)}{(a+B_f)}.s - (1-s)$$

The bookmakers believe that s = a + α with pr(0.5) and s = a - α with pr(0.5):

$$E(ret_f) = 0.5[\frac{(1-a-B_f).(a+\alpha)}{(a+B_f)} - (1-a-\alpha)] + 0.5[\frac{(1-a-B_f).(a-\alpha)}{(a+B_f)} - (1-a+\alpha)]$$

$$E(ret_f) = 0.5.[\frac{2a.(1-a-B_f)}{(a+B_f)} - 2 + 2a]$$

$$E(ret_f) = \frac{a(1-a-B_f)}{(a+B_f)} - (1-a)$$

Similarly, expected return on the longshot is:

$$E(ret_l) = \frac{(a-B_l).(1-s)}{(1-a+B_l)} - s$$

With s = a+α with pr(0.5) and a-α with pr(0.5), this reduces to

$$E(ret_l) = \frac{(a-B_l).(1-a)}{(1-a+B_l)} - a$$

Total expected net payout (TEP) on each horse is given by:

TEP = w.E(ret to insider) + (1-w).E(ret to noise trader)

We require TEP to be zero for both horses:

TEP(favorite) = 0

$$w.[\frac{(a+\alpha).(1-a-B_f)}{(a+B_f)} - (1-a-\alpha)] + (1-w).[\frac{a(1-a-B_f)}{(a+B_f)} - (1-a)] = 0$$

which reduces to: $\alpha w - B_f = 0$

$B_f = \alpha w$

TEP(longshot) = 0

$$w[\frac{(1-a+\alpha).(a-B_l)}{(1-a+B_l)} - (a-\alpha)] + (1-w)[\frac{(a-B_l).(1-a)}{(1-a+B_l)} - a] = 0$$

which reduces to $\alpha w - B_l = 0$

$B_l = \alpha w$

Thus under the threat of insider trading, bookmakers increase their implied probabilities by an equal amount for both the favorite and longshot.

BIBLIOGRAPHY

Abarbanell, J.S. and Bernard, V.L. (1992) 'Test of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior', *Journal of Finance*, 47, pp. 1181-1207.

Affleck-Graves, J, and Mendenhall, R.R. (1992) 'The Relation Between the Value Line Enigma and Post-Earnings-Announcement Drift', *Journal of Financial Economics*, 31, pp. 75-96.

Ainslie, T. (1968), 'Ainslie's Complete Guide to Thoroughbred Racing', New York: Trident Press.

Alexakis, P. and Xanthakis, M. (1995) 'Day of the Week Effect on the Greek Stock Market', *Applied Financial Economics*, 5, pp. 43-50.

Alexander, S.S. (1961) 'Price Movements in Speculative Markets: Trends or Random Walks?', *Industrial Management Review*, 2, May, pp. 7-26. Also reprinted in : Cootner, P.(ed.) (1964) : *The Random Character of Stock Market Prices*, pp. 199-218.
Cambridge: M.I.T.

Alexander, S.S. (1964) 'Price Movements in Speculative Markets : Trends or Random Walks? No.2', *Industrial Management Review*, 5, Spring. Also reprinted in : Cootner, P. (ed.) (1964) : *The Random Character of Stock Market Prices*, pp. 338-372.

Alexander, C.O. and Johnson, A. (1992) 'Are Foreign Exchange Markets Really Efficient?', *Economics Letters*, 40(4), pp. 449-453.

Ali, M.M. (1977) 'Probability and Utility Estimates for Racetrack Bettors,' *Journal of Political Economy*, 83, pp. 803-815.

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

Allais, M. (1953) 'L'extension des theories de l'equilibre economique general et du rendement social au cas du risque', *Econometrica*, 21, April, pp. 269-290.

Allen, S. and Ramanan, R. (1995) 'Insider Trading, Earnings Changes and Stock Prices', *Management Science*, 41(4), pp. 653-668.

Allen, H.C. and Taylor, M.P. (1989) 'Chartists and Fundamentals in the Foreign Exchange Market', *Bank of England Discussion Paper*, no.40. Bank of England, Threadneedle St., London EC2R 8AH.

Allen, H.C. and Taylor, M.P. (1990) 'Charts, Noise and Fundamentals in the London Foreign Exchange Market', *Economic Journal*, 100, pp. 49-59.

Ambachtsheer, K.P. (1972) 'Portfolio Theory and the Security Analyst', *Financial Analysts Journal*, 28, November-December, pp. 53-57.

Ambachtsheer, K.P. (1974) 'Profit Potential in an "Almost Efficient" Market', *Journal of Portfolio Management*, 1, Fall, pp. 84-87.

Ambachtsheer, K.P. and Farrell, J.L. (1979) 'Can Active Management Add Value?', *Financial Analysts Journal*, 35, November-December, pp. 39-48.

Ambrose, B.W., Ancel, E. and Griffiths, M.D. (1992) 'The Fractal Structure of Real Investment Trust Returns: The Search for Evidence of Market Segmentation and Nonlinear Dependency', *American Real Estate and Urban Economics Association Journal*, 20 (1), Spring, pp. 25-54.

Anderson, D., Clarke R. and Ziegler, P. (1985), 'Information, Equilibrium, Efficiency in Betting Markets,' University of Queensland Working Paper.

Antoniou, A. and Holmes, P. (1996) 'Futures Market Efficiency, the Unbiasedness Hypothesis and Variance-Bounds tests: the Case of the FTSE_100 Futures Contract', *Bulletin of Economic Research*, 48(2), pp. 115-128.

Arak, M. and Taylor, D. (1996) 'Risk and Return in Trading Closed-End Country Funds: Can Trading Beat Holding Foreign Stocks?' *Quarterly Review of Economics and Finance*, 36(2), Summer, pp. 219-231.

Ariel, R.A. (1987) 'A Monthly Effect in Stock Returns', *Journal of Financial Economics*, 18 (1), pp. 161-174.

Ariel, R.A. (1990) 'High Stock Returns Before Holidays: Existence and Evidence on Possible Causes', *Journal of Finance*, 45, pp. 1611-1626.

Arrow, K. (1953) 'The Role of Securities in the Optimal Allocation of Risk-Bearing', *Review of Economic Studies*, 31, April 1964, pp. 91-96.

Asch, P. and Quandt, R.E. (1986), *The Professor's Guide to Strategies*, Dover, MA: Auburn House.

Asch, P. and Quandt, R.E. (1987), 'Efficiency and Profitability in Exotic Bets', *Economica*, 59, pp. 278-298.

Asch, P. and Quandt, R.E. (1988), 'Betting Bias in Exotic Bets', *Economic Letters*, 28, pp. 215-219.

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

Asquith, P. (1983) 'Merger Bids, Uncertainty and Stock Holder Returns', *Journal of Financial Economics*, 11, pp. 51-83.

Atkins, A.B. and Basu, S. (1991) 'The Impact of Public Announcements Made After the Stock Market Closes', Unpublished Manuscript, University of Arizona and University of Denver.

Attanasio, O. and Wadhwani, S. (1990) 'Does the CAPM Explain Why the Dividend Yield Helps Predict Returns?' LSE Financial Markets Group Discussion Paper.

Ayadi, O.F. and Pyun, C.S. (1994) 'An Application of Variance Ratio Tests to the Korean Securities Market', *Journal of Banking and Finance*, 18(4), pp. 643-658.

Bachelier, L. (1900) '*Theorie de la speculation*'. Paris:Gautier-Villars. Reprinted in English in: Cootner, P.(ed.)(1964), *The Random Character of Stock Market Prices*.
Cambridge : M.I.T.

Ball, R. (1972) 'Changes in Accounting Techniques and Stock Prices, Empirical Research in Accounting: Selected Studies', *Journal of Accounting Research*, supplement, 10, pp. 1-38.

Ball, R. and Brown, P. (1968) 'An Empirical Evaluation of Accounting Income Numbers', *Journal of Accounting Research*, 6, Autumn, pp. 159-78.

Ball, R. and Kothari, S.P. (1989) 'Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlation in Returns', *Journal of Financial Economics*, 25, pp. 51-74.

Bank of England Quarterly Bulletin (1989), 'Chart Analysis and the Foreign Exchange Market', November, pp. 548-551.

Banz, R.W. (1981), 'The Relationship Between Return and Market Value of Common Stocks', *Journal of Financial Economics*, March, 9 (1), pp. 3-18.

Basu, S. (1983) 'The Relationship Between Earnings' Yield, Market Value and Returns for NYSE Common Stocks: Further Evidence', *Journal of Financial Economics*, June, 12 (1), pp. 129-156.

Baumol, W.J. and Benhabib, J. (1989) 'Chaos: Significance, Mechanism and Economic Applications', *Journal of Economic Perspectives*, Winter, 3(1), pp. 77-105.

Beck, S.E. (1994) 'Cointegration and Market Efficiency in Commodities Futures Markets', *Applied Economics*, 26(3), pp. 249-257.

Becker, K.G., Finnerty, J.E. and Gupta, M. (1990) 'The Intertemporal Relation Between the U.S. and Japanese Stock Markets', *Journal of Finance*, 45 (4), September 1990, pp. 1297-1306.

Becker, K.G., Finnerty, J.E. and Tucker, A.L. (1992) 'The Intraday Interdependence Structure between U.S. and Japanese Equity Markets', *Journal of Financial Research*, 15 (1), Spring, pp. 27-37.

Benter, W. (1994), Computer Based Horse race Handicapping and Wagering Systems: A Report', in *Efficiency of Racetrack Betting Markets* (Eds.) Donald B. Hausch, Victor S. Y. Lo and William T. Ziemba, London: Academic Press, pp. 183-198.

Bernard, A. (1984) '*How to Use the Value Line Investment Survey: A Subscriber's Guide*', Value Line : New York.

Bernard, V.L. and Thomas, J.K. (1989) 'Post-Earnings Announcement Drift: Delayed Price Response or Risk Premium?', *Journal of Accounting Research*, 27 (Supplement), pp. 1-36.

Bernard, V.L. and Thomas, J.K. (1990) 'Evidence that Stock Prices Do Not Fully Reflect the Implications of Current Earnings for Future Earnings', *Journal of Accounting and Economics*, 13 (4), December, pp. 305-340.

Berry, T.D. and Howe, K.M. (1994) 'Public Information Arrival', *Journal of Finance*, 49 (4), pp. 1331-1346.

Betton, S. (1994), Post Position Bias: An Econometric Analysis of the 1987 Season at Exhibition Park', in *Efficiency of Racetrack Betting Markets* (Eds.) Donald B. Hausch, Victor S. Y. Lo and William T. Ziemba, London: Academic Press, pp. 511-526.

Beyer, A. (1983), '*The Winning Horseplayer*', Houghton-Mifflin Co., Boston.

Bhandari, L.C. (1988) 'Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence', *Journal of Finance*, 43, pp. 507-528.

Bird, R. and McCrae, M. (1987), 'Tests of the Efficiency of Racetrack Betting Using Bookmaker Odds', *Management Science*, 33, pp. 1552-1562.

Bird, R. and McCrae, M. (1994) 'Efficiency of Racetrack Betting Markets: Australian Evidence', in *Efficiency of Racetrack Betting Markets* (Eds.) Donald B. Hausch, Victor S. Y. Lo and

William T. Ziemba, London: Academic Press, pp. 575-82.

Bird, R., McCrae, M. and Beggs, J. (1987) 'Are Gamblers Really Risk Takers?', *Australian Economic Papers*, December, pp. 237-253.

Bjerring, J.H., Lakonishok, J. and Vermaelen, T. (1983) 'Stock Market Prices and Financial Analysts' Recommendations', *Journal of Finance*, 38, pp. 187-204.

Black, F. (1972) 'Capital Market Equilibrium with Restricted Borrowing', *Journal of Business* ,45, pp. 444-455.

Black, F. (1973) 'Yes Virginia, There is Hope: Tests of the Value Line Ranking System', *Financial Analysts Journal*, 29, September-October, pp. 10-14.

Black, F. (1986) 'Presidential Address: Noise', *Journal of Finance*, 41(3), pp. 529-543.

Blackburn, Philip and Peirson, John (1995) 'Betting at British racecourses: An Analysis of Semi-strong Efficiency between Bookmaker and Tote odds', *Studies in Economics*, University of Kent, 95/4.

Blaug, M. (1968) '*Economic Theory in Retrospect*', London: Heinemann.

Bodie, Z. (1976) 'Common Stocks as a Hedge Against Inflation', *Journal of Finance*, 41, pp. 529-543.

Bolton, R.N. and Chapman, R.G. (1986) 'Searching for Positive Returns at the Track: A Multinomial Logit Model for Handicapping Horse Races', *Management Science*, 32, pp. 1040-1059.

Bonin, J.M. and Moses, E.A. (1974) 'Seasonal Variation in Prices of Individual Dow Jones Industrial Stocks', *Journal of Financial and Quantitative Analysis*, 9, December, pp. 963-991.

Borch, K. (1962) 'Equilibrium in a Reinsurance Market', *Econometrica*, 30, July, pp. 424-444.

Branch, B. (1977) 'A Tax Loss Trading Rule', *Journal of Business*, April, pp. 198-207.

Brauer, G.A. (1984) 'Open-ending Closed-End Funds', *Journal of Financial Economics*, 13, pp. 491-507.

Brealey, R.A. (1970) 'The Distribution and Independence of Successive Rates of Return from the British Equity Market', *Journal of Business Finance*, Summer, pp. 29-40.

Brickley, J.A. and Schallheim, J.S. (1985) 'Lifting the Lid on Closed-End Investment Companies: A Case of Abnormal Returns', *Journal of Finanncial and Quantitative Analysis*, 20, pp. 107-118.

Brock, W., Lakonishok, J. and Le Baron, B. (1992) 'Simple Technical Trading Rules and the Stochastic Properties of Stock Returns', *Journal of Finance*, 47(5), December, pp. 1731-1764.

Brown, G., Draper, P. and MacKenzie, E. (1993) 'Consistency of U.K. Pension Fund Investment

Performance', Unpublished Paper, Strathclyde University, 1993.

Brown, K.C. and Harlow, W.V. (1988) 'Assessing the Magnitude and Intensity of Stock Market Overreaction', *Journal of Portfolio Management*, pp. 6-13.

Brown, P. and Kennelly, J.W. (1972) 'The Informational Content of Quarterly Earnings : An Extension and Some Further Evidence', *Journal of Business*, 45, July, pp. 403-415.

Brown, S., Geotzmann, W,. Ibbotson, R. and Ross, S. (1992) 'Survivorship Bias in Performance Studies', *Review of Financial Studies*, 5, pp. 553-580.

Brown, S.J. and Warner, J.B.(1985) 'Using Daily Stock Returns: The Case of Event Studies', *Journal of Financial Economics*, 14, pp. 3-32.

Bruce, A.C. and Johnson, J.E.V. (1992) 'Toward an Explanation of Betting as a Leisure Pursuit', *Leisure Studies*, 11, pp. 201-218.

Busche, K. (1994) 'Efficient Market Results in an Asian Setting',
In: *Efficiency of Racetrack Betting Markets* (1994), ed. Hausch, D.B., Lo, S.Y. and Ziemba, W.T., pp. 615-616. London: Academic Press.

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

Cain, M., Law, D. and Peel, D.A. (1992) Greyhound Racing: Further Empirical Evidence on Market Efficiency in a Wagering Market. Aberystwyth Economic Research Papers, March.

Campbell, J.Y. (1987) 'Stock Returns and the Term Structure', *Journal of Financial Economics*, 18, pp. 373-399.

Campbell, J.Y. and Hamao, Y. (1989) 'Predictable Stock Returns in the United States and Japan: A Study of Long-Term Capital Market Integration', *LSE Financial Markets Group Discussion Paper*.

Campbell, J.Y. and Shiller, R. (1988) 'The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors', *Review of Financial Studies*, 1, pp. 195-228.

Canfield, B., Fauman, B.C. and Ziemba, W.T. (1987), 'Efficient Market Adjustment of Odds Prices to Reflect Track Biases,' *Management Science*, 33, pp. 1428-1439.

Chadha and Quandt (1996), 'Betting Bias and Market Equilibrium in Racetrack Betting', *Applied Financial Economics*, 6 (3), pp. 287-292.

Chan, K.C. (1986) 'Can Tax-Loss Selling Explain the January Seasonal Effect in Stock Returns?', *Journal of Finance*, December, pp. 1115-1128.

Chan, K.C. (1987) 'On the Return of the Contrarian Investment Strategy', Working Paper, Faculty of Finance, Ohio State University, January.

Chan, K.C. and Chen, N. (1991) 'Structural and Return Characteristics of Small and Large Firms', *Journal of Finance*, 46, pp. 1467-1484.

Chan, L.K.C., Hamao, Y. and Lakonishok, J. (1991) 'Fundamentals and Stock Returns in Japan', *Journal of Finance*, 46, pp. 1739-1789.

Chandy, P.R., Peavy, J.W. and Reichenstein, W. (1993) 'A Note on the Value Line Stock Highlight Effect', *Journal of Financial Research*, 16(2), pp. 171-179.

Chapman, R.G. (1994), 'Still Searching for positive returns at the Track: Empirical Results from 2,000 Hong Kong Races', in *Efficiency of Racetrack Betting Markets* (Eds.) Donald B. Hausch, Victor S. Y. Lo and William T. Ziemba, London: Academic Press, pp. 173-181.

Charest, G. (1978) 'Dividend Information, Stock Returns and Market Efficiency-II', *Journal of Financial Economics*, 6, June-September, pp. 297-330.

Chelley-Steeley, P.L. and Pentecost, E.J. (1994) 'Stock Market Efficiency, the Small Firm Effect and Cointegration', *Applied Financial Economics*, 4, pp. 405-411.

Chopra, N., Lakonishok, J. and Ritter, J. (1992) 'Measuring Abnormal Returns: Do Stocks Overreact?', *Journal of Financial Economics*, 31, pp. 235-268.

Choudhry, T. (1994), 'Stochastic Trends and Stock prices: An International Inquiry', *Applied Financial Economics*, 4(6), December.

Clare, A.D. and Thomas, S.H. (1992) 'International Evidence for the Predictability of Stock and Bond Returns', *Economics Letters*, 40, pp. 105-112.

Clotfelter, C. and Cook, P.J. (1989), '*Selling Hope: State Lotteries in America*', Cambridge: Harvard University Press.

Clotfelter, C. and Cook, P.J. (1991), 'Lotteries in the Real World', *Journal of Risk and Uncertainty*, 4, pp. 227-232.

Clotfelter, C. and Cook, P.J. (1993), 'The "Gambler's Fallacy" in Lottery Play', *Management Science*, 39 (12), December, pp. 1521-1525.

Cochrane, J.H. (1988) 'How Big is the Random Walk in GNP?' *Journal of Political Economy*, 96(5), pp. 893-920.

Cochrane, J.H. (1991) 'Production-Based Asset Pricing and the Link between Stock Returns and Economic Fluctuations', *Journal of Finance*, 46, pp. 209-238.

Cohen, K.J., Hawawini, G.A., Maier, S.F., Schwartz, R.A. and Whitcomb, D.K. (1980) 'Implications of Microstructure Theory for Empirical Research on Stock Price Behaviour', *Journal of Finance*, 35 (2), pp. 249-257.

Cohen, I.S. and Stephens, G.D. (1963), '*Scientific Handicapping: Tested Ways to Win at the Track*', Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Colling, P.L. and Irwin, S,H. (1990) 'The Reaction of Live Hog Future Prices to USDA Hogs and Pigs Reports', *American Journal of Agricultural Economics*, 72 (1), February, pp. 84-94.

Collins, D. (1975) 'SEC Product-line Reporting and Market Efficiency', *Journal of Financial Economics*, 2, June, pp. 125-164.

Connolly, R.A. (1989) 'An Examination of the Robustness of the Weekend Effect', *Journal of Financial and Quantitative Analysis*, 24, pp. 133-169.

Conrad, J. and Kaul, G. (1988) 'Time-Variation in Expected Returns', *Journal of Business*, 61, pp. 409-425.

Consumer Guide (1988). Top Performing Mutual Funds.
Publications International, Ltd. : Lincolnwood, Illinois.

Cooper, J.C.B. (1982), 'World Stockmarkets: Some Random Walk Tests', *Applied Economics*, October, pp. 515-531.

Cootner, P. (1962), 'Stock Prices : Random vs. Systematic Changes', *Industrial Management Review*, 3, Spring, pp. 24-45, in: Cootner, P. (ed.) (1964), *The Random Character of Stock*

Market Prices, Cambridge: MIT Press.

Copeland, T. and Mayers, D. (1982), 'The Value Line Enigma (1965-1978): A Case Study of Performance Evaluation Issues', *Journal of Financial Economics*, 10, pp. 289-321.

Coslow, S. and Schultz, H.D. (1966) '*A Treasury of Wall Street Wisdom*', Investors Press Inc., Pallisades Park, N.Y.

Cowles, A. (1933) 'Can Stock Market Forecasters Forecast?' *Econometrica*, 1(4), July, pp. 309-324.

Crafts, N.F.R. (1985), 'Some Evidence of Insider Knowledge in Horse Race Betting in Britain', *Economica*, 52, pp. 295-304.

Crafts, N.F.R. (1994), 'Winning Systems? Some Further Evidence on Insiders and Outsiders in British Horse Race Betting', in *Efficiency of Racetrack Betting Markets* (Eds.) Donald B. Hausch, Victor S. Y. Lo and William T. Ziemba, London: Academic Press, pp. 545-549.

Cross, F. (1973) 'The Behaviour of Stock Prices on Fridays and Mondays', *Financial Analysts Journal*, 29(6), November-December, pp. 2-69.

Cunningham, L.A. (1994) 'From Random Walks to Chaotic Crashes: The Linear Genealogy of the Efficient Capital Market Hypothesis', *George Washington Law Review*, 64 (4), pp. 546-573.

Cunningham, S.W. (1973) 'The Predictability of British Stock Market Prices', *Applied Statistics*, 22 (3), pp. 315-331.

Cutler, D.M., Poterba, J.M. and Summers, L.H. (1990) 'Speculative Dynamics and the Role of Feedback Traders', *American Economic Review*, 80, May, pp. 63-68.

Cutler, D.M., Poterba, J.M. and Summers, L.H. (1991) 'Speculative Dynamics', *Review of Economic Studies*, 58, pp. 529-546.

Damodaran, A. and Liu, C.H. (1993) 'Insider trading as a Sign of Private Information', *Review of Financial Studies*, 6 (1), pp. 79-119.

Dark, F.H. and Kato, K. (1986) 'Stock Market Over-reaction in the Japanese Stock Market', Working Paper, Iowa Stata University.

Davies, P.L., and Canes, M. (1978) 'Stock Prices and the Publication of Second-Hand Information', *Journal of Business*, 51, pp. 43-56.

De Bondt, W. and Thaler, R. (1985) 'Does the Stock Market Overreact?', *Journal of Finance*, 40(3), pp. 793-805.

DeBondt, W. and Thaler, R. (1987) 'Further Evidence on Investor Overreaction and Stock Market Seasonality', *Journal of Finance*, July, 42 (3), pp. 557-581

DeBondt, W. and Thaler, R. (1990), 'Do Security Analysts Overreact?', *American Economic Review*, 80 (Papers and Proceedings), pp. 52-57.

Debreu, G. (1959) '*Theory of Value*', New York : Wiley.

De Leeuw, F. and McKelvey, M.J. (1984), 'Price Expectations of Business Firms: Bias in the Short and Long Run,' *American Economic Review*, 74, March, pp. 99-110.

DeLong, J., Shleifer, A., Summers, L.H. and Waldman, R.J. (1990) 'Positive Feedback Investment Strategies and Destabilizing Rational Speculation', *Journal of Finance*, 45, pp. 379-396.

De Long, J.B., Shleifer, A., Summers, L.H., and Waldman, R.J. (1990) 'Noise Trader Risk in Financial Markets', *Journal of Political Economy*, 98, August, pp. 703-738.

De Long, B.J., Shleifer, A., Summers, L.H. and Waldmann, R.J. (1991) 'The Survival of Noise Traders', *Journal of Business*, 64, pp. 1-19.

Detta, S. and Iskandardatton, M.E. (1996) 'Does Insider Trading have Information Content for the Bond Market?', *Journal of Banking and Finance*, 20 (3), pp. 555-575.

Diamandis, P.F. and Kouretas, G.P. (1995) 'Cointegration and Market Efficiency - A Time-Series Analysis of the Greek Drachma', *Applied Economics Letters*, 1995, 2 (8), pp. 271-277

Dickey, D.A. and Fuller, W.A. (1979) 'Distribution of the Estimators for Autoregressive Time Series with a Unit Root', *Journal of the American Statistical Association*, 74, pp. 427-431.

Dickey, D.A. and Fuller, W.A. (1981) 'Likelihood Ratio Statistic for Autoregressive Time Series with a Unit Root', *Econometrica*, pp. 1057-1072

Dimson, E. and Fraletti, P. (1986) 'Brokers Recommendations: The Value of a Telephone Tip', *Economic Journal*, 96, March, pp. 139-159.

Dimson, E. and Marsh, P. (1984) 'An Analysis of Brokers' and Analysts' Unpublished Forecasts of U.K. Stock Returns', *Journal of Finance*, 39 (5), pp. 1257-1292.

Dorfman, J.R. (1993) 'Luck or Logic? Debate Rages on over Efficient Market Theory', in: *Wall Street Journal*, (J), p. C1, 11th. April.

Dowie, J. (1976) 'On the Efficiency and Equity of Betting Markets', *Economica*, 43 (170), May, pp. 139-150.

Dowie, J. (1992a), 'The Ethics of Parimutuel Systems', *Journal of Gambling Studies*, 8(4), Winter, pp. 371-381.

Dowie, J. (1992b), 'The Ethics of Parimutuel Systems', in Eadington, W.R. (ed.), '*Gambling and Commercial Gaming: Essays in Business, Economics, Philosophy and Science*', University of Nevada, Reno, pp. 345-355.

Dowie, J., Coton, M. and Miers, D.M. (1991) 'Consumer Protection in Betting', Paper presented at the Eighth International Conference on Risk and Gambling, London. In: Eadington, W.R. and Cornelius, J. (ed.) '*Gambling and Public Policy: International Perspectives*', pp. 433-444.

Dyl, E.A. (1977) 'Capital Gains Taxation and Year-End Stock Market Behaviour', *Journal of Finance*, March, pp. 165-175.

Dyl, E.A. and Maxfield, K. (1987) 'Does the Stock Market Over-react?', Working Paper, University of Arizona.

Economist (1989), 23rd. September, p. 135. Chartism: An M That Works.

Economist (1992a), 8th.August, pp. 67-72. Fidelity Changes Tack.

Economist (1992b), 5th. December. U.K. Special: Beating The Market: Yes, It Can Be Done, pp. 23-26.

Economist, (1993), 23rd. October. Frontiers of Finance: Survey, pp. 1-24.

Ederington, L.H. and Lee, J.H. (1993) 'How Markets Process Information: News Releases and Volatility', *Journal of Finance*, 48, pp. 1161-1191.

Eiser, J.R. and van der Pligt, J. (1988), '*Attitudes and Decisions*', Routledge, London.

Engel, C. and Morris, C.S. (1991) 'Challenges to Stock Market Efficiency: Evidence from Mean Reversion Studies', *Economic Review*, Sept./Oct. 1991, pp. 21-34.

Engle, R.F. and Granger, C.W.J. (1987) 'Cointegration and Error Correction; Representation, Estimation and Testing', *Econometrica*, 55, pp. 987-1007.

Euromoney: Forbes (1992a): U.S.A., 17th.August, p.135, Academics Say Past Performance is a Fair Guide to Future Investment Performance.

Euromoney: Forbes (1992b): U.S.A., 30th. March, p. 148: Value Approach to Investing Looks Better Than Ever.

Fabricand, B.P. (1965), '*Horse Sense*', New York: David McKay.

Fama, E.F. (1965) 'The Behaviour of Stock Market Prices', *Journal of Business*, 38, January, pp. 34-105.

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

Fama, E.F. (1976) 'Efficient Capital Markets: Reply', *Journal of Finance*, October, 3 (4), pp. 143-145.

Fama, E.F. (1981) 'Stock Returns, Real Activity, Inflation and Money', *American Economic*

Review, 71, pp. 545-565.

Fama, E. (1991) 'Efficient Capital Markets, II', *Journal of Finance*, 46 (5), December, pp. 1575-1617.

Fama, E.F. and Blume, M. (1966) 'Filter Rules and Stock Market Trading Profits', *Journal of Business*, 39 (1) (Special Supplement, January), pp. 226-241.

Fama, E., and French, K. (1988a) 'Permanent and Temporary Components of Stock Prices', *Journal of Political Economy*, 96 (2), April, pp. 246-273.

Fama, E.F. and French, K.R. (1988b) 'Dividend Yields and Expected Stock Returns', *Journal of Financial Economics*, 22 (1), pp. 3-25.

Fama, E.F. and French, K.R. (1989) 'Business Conditions and Expected Returns on Stocks and Bonds', *Journal of Financial Economics*, 25, pp. 23-49.

Fama, E.F. and French, K.R. (1992) 'The Cross-Section of Expected Stock Returns', *Journal of Finance*, 47 (2), pp. 427-465.

Fama, E.F. and MacBeth, J.D. (1973) 'Risk, Return, and Equilibrium: Empirical Tests', *Journal of Political Economy*, 81 (3), May/June, pp. 607-636.

Fama, E.F. and Miller, M.H. (1972) '*The Theory of Finance*', Hinsdale, Illinois : Dryden Press.

Fama, E.F. and Schwert, G.W. (1977) 'Asset Returns and Inflation', *Journal of Financial Economics*, 5, pp. 115-146.

Fama, E.F., Fisher, L., Jensen, M.C. and Roll, R. (1969) 'The Adjustment of Stock Prices to New Information', *International Economic Review*, 10 (2), February, pp. 1-21.

Figgis, E.L. (1951), *Focus on Gambling*, London:Barker.

Figgis, E.L. (1974a), *Sporting Life*, 11th.March. Quoted in *Royal Commission on Gambling* (1978), Final Report, vol.2, p.469. London: H.M.S.O.

Figgis, E.L. (1974b), 'Betting to Win', London: Playfair.

Figgis, E.L. (1976), 'Gamblers Handbook', London: Hamlyn.

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

Filby, M.P. and Harvey, L. (1988), 'Recreational Betting: Everyday Activity and Strategies', *Leisure Studies*, 7, pp. 159-172.

Financial Times (1989), 1st.September, p.10, Chartists with a Hot Line on Currency.

Financial Times (1993), 28th.October, p.22. Economic Viewpoint: Equity Delusions.

Fingleton, J. and Waldron, P. (1996), 'Optimal Determination of Bookmakers' Betting Odds: Theory and Tests', *Trinity Economic Papers*, Technical paper No.9, December.

Finnerty, J.E. (1976) 'Insiders and Market Efficiency', *Journal of Finance*, 16, September, pp. 1141-1148.

Firth, M.A. (1976) 'The Impact of Earnings Announcements on the Behaviour of Similar Type Firms', *Economic Journal*, 86, June, pp. 296-306

Firth, M.A. (1977) '*The Valuation of Shares and the Efficient-Markets Theory*', London and Basingstoke: Macmillan Press Ltd.

Fisher, L. (1966) 'Some New Stock Market Indexes', *Journal of Business*, 39, pp. 191-225.

Forbes, W.P. (1996) 'Picking Winners? A Survey of the Mean Reversion and Overeaction of Stock Prices Literature', *Journal of Economic Surveys*, 10 (2), pp. 123-158.

Fortune,P. (1991) 'Stock Market Efficiency: An Autopsy?', *New England Economic Review*, 0 (0), March-April, pp. 17-40.

Foster, G. (1973) 'Stock Market Reaction to Estimates of Earnings per Share by Company Officials', *Journal of Accounting Research*, 11, pp. 25-37.

Franks, J., Harris, R.S. and Titman, S. (1991) 'The Postmerger Share Price Performance of

Acquiring Firms', *Journal of Financial Economics*, 29, pp. 81-96.

French, K.R. (1980) 'Stock Returns and the Weekend Effect', *Journal of Financial Economics*, 8 (1), March, pp. 55-70.

French, K. and Roll, R.W. (1986) 'Stock Return Variances: The Arrival of Information and the Reaction of Traders', *Journal of Financial Economics*, 17 (1), September, pp. 5-26.

Frennberg, P. and Hansson, B. (1993) 'Testing the Random Walk Hypothesis on Swedish Stock Prices: 1919-1990',

Journal of Banking and Finance, 17, pp. 175-191.

Friedman, M. and Savage, L.J. (1948), 'The Utility Analysis of Choices Involving Risk', *Journal of Political Economy*, August, pp. 279-304.

Friend, I.,Blume, M. and Crockett, J. (1970) '*Mutual funds and other Institutional Investors, a New Perspective*', McGraw-Hill.

Froot, K. and Perold, A. (1990) 'New Trading Practices and Short-Run Market Efficiency', National Bureau of Economic Research, Working Paper no. 3498.

Fuller, W.A. (1976) '*Introduction to Statistical Time Series*.

John Wiley: New York.

Gabriel, P. E. and Marsden, J. R. (1990) 'An Examination of Market Efficiency in British Racetrack Betting', *Journal of Political Economy*, 98, pp. 874-885.

Gabriel, P. E. and Marsden, J. R. (1991) 'An Examination of Market Efficiency in British Racetrack Betting: Errata and Corrections *Journal of Political Economy*, 99, pp. 657-659.

Gartley, H.M. (1930) '*Profits in the Stock Market*',
H.M. Gartley Inc., New York, NY.

Gencay, R. (1996) 'Nonlinear Prediction of Security Returns with Moving Average Rules', *Journal of Forecasting*, 15 (3), pp. 165-174.

Gibbons, M.R. and Hess, P.J. (1981) 'Day of the Week Effects and Asset Returns', *Journal of Business*, 54 (4), October, pp. 579-596.

Giliberto, S. and Varaiya, N.P. (1989) 'The Winner's Curse and Bidder Competition in Acquisitions: Evidence from Failed Bank Auctions', *Journal of Finance*, 44 (1), March, pp. 59-76.

Goetzmann, W. and Ibbotson, R. (1994) 'Do Winners Repeat?', *Journal of Portfolio Management*, 20, pp. 9-18.

Goldstein, R. (1992) Cook-Weisberg test of heteroscedasticity *Stata Technical Bulletin*, **10**, 27-28.

Golec, J. and Tamarkin, M. (1995) 'Do Bettors Prefer Long Shots Because They Are Risk-Lovers, or Are They Just Overconfident?', *Journal of Risk and Uncertainty*, 11, pp. 51-64.

Grammatikos, T. and Saunders, A. (1990) 'Additions to Loan-Loss Reserves', *Journal of Monetary Economics*, 22, pp. 289-304.

Granger, C.W.J. (1986) 'Developments in the Study of Cointegrated Variables', *Oxford Bulletin of Economics and Statistics*, 48, pp. 213-228.

Granger, C. (1992) 'Forecasting Stock Market Prices: Lessons for Forecasters', *International Journal of Forecasting*, 8 (1), June, pp. 3-13.

Granger, C.W.J. and Morgenstern, O. (1963) '*Spectral Analysis of New York Stock Market Prices*', in: Cootner,P.H. (ed.) (1964)

The Random Character of Stock Market Prices, pp. 162-188.

Cambridge: MIT Press.

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

Griffith, R.M. (1961) 'A Footnote on Horse Race Betting', *Transactions Kentucky Academy of Science*, 22, pp. 78-81.

In: *Efficiency of Racetrack Betting Markets* (1994), ed. Hausch, D.B., Lo, S.Y. and Ziemba, W.T., pp. 27-30. London: Academic Press.

Griffiths, M.D. and White, R.W. (1993) 'Tax-Induced Trading and the Turn-of-the-Year Anomaly: An Intraday Study', *Journal of Finance*, 48(2), June, pp. 575-598.

Grinblatt, M. and Titman, S. (1992) 'The Persistence of Mutual Fund Performance', *Journal of Finance*, 47, pp. 1977-1984.

Grossman, S.J. and Stiglitz, J. (1976) 'Information and Competitive Price Systems', *American Economic Review*, 66 (2), May, pp. 246-53.

Grossman, S. and Stiglitz, J.E. (1980) 'On the Impossibility of Informationally Efficient Markets', *American Economic Review*, 70 (3), June, pp. 393-408.

Guletkin, M.N. and Guletkin, N.B. (1987) 'Stock Return Anomalies and the Tests of the APT', *Journal of Finance*, December, 42 (5), pp. 1213-1224.

Hagerman, R.L. and Richmond, R.D. (1973), 'Random Walks, Martingales and the OTC', *Journal of Finance*, 28, pp. 897-909.

Harris, L. (1986) 'A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns', *Journal of Financial Economics*, May, 16 (1), pp. 99-117.

Harvey, C.R. (1991) 'The World Price of Covariance Risk', *Journal of Finance*, 46, pp. 111-157.

Harvey, C.R. and Huang, R.D. (1991) 'Volatility in the Foreign Currency Futures Market', *Review* *of Financial Studies*, 4, pp. 543-569.

Hausch, D.B. and Ziemba, W.T. (1985), 'Transactions Costs, Extent of Inefficiencies, Entries and Multiple Wagers in a Racetrack Betting Model', *Management Science*, 31, pp. 381-394.

Hausch, D.B. and Ziemba, W.T., (1990) 'Arbitrage Strategies for Cross-track Betting on Major Horse Races', *Journal of Business*, 33, pp. 61-78.

Hausch, D.B., Ziemba, W.T. and Rubinstein, M. (1981) 'Efficiency of the Market for Racetrack Betting', *Management Science*, 27, pp. 1435-1452.

Hawawini, G. (ed.) (1984) 'European Equity Markets: Price Behaviour and Efficiency', Salomon Brothers Center/NYU Monograph 4/5.

Hendricks, D., Patel, J. and Zeckhauser, R. (1993), 'Hot Hands in Mutual Funds', *Journal of Finance*, 48, March, pp. 3-130.

Hendricks, K. and Porter, R.H. (1988) 'Empirical Study of an Auction with Asymmetric Information', *American Economic Review*, 78 (5), December, pp. 865-883.

Henery, R.J. (1985) 'On the Average Probability of Losing Bets on Horses with Given Starting Price Odds', *Journal of the Royal Statistical Society*, 148 (4) pp. 342-349.

Hoerl, A.E. and Fallin, H.K. (1974), 'Reliability of Subjective Evaluations in a High Incentive Situation', *Journal of the Royal Statistical Society A*, 137, pp. 227-230.

Holloway, C. (1981) 'A Note on Testing on Aggressive Investment Strategy Using Value Line Ranks', *Journal of Finance*, 36, June, pp. 711-719.

Hsieh, D.A. (1991) 'Chaos and Non-Linear Dynamics: Applications to Financial Markets', *Journal of Finance*, 46, pp. 1839-1877.

Huang, B. (1995) 'Do Asian Stock Market Prices Follow Random Walks? Evidence from the Variance Ratio Test', *Applied Financial Economics*, 1995, 5, pp. 251-256

Huang, B. and Yang, C.W. (1995) 'The Fractal Structure in Multinational Stock Returns', *Applied Economics Letters*, 2, pp. 67-71.

Huberman, G. and Kandel, S. (1987), 'Value Line Rank and Firm Size', *Journal of Business*, 60, pp. 577-589.

Huberman, G. and Kandel, S. (1990), 'Market Efficiency and Value Line's Record', *Journal of Business*, 63, pp. 187-216.

Hulbert, M. (1990) 'Proof of Pudding', *Forbes*, December, 10, p.316.

Hurley, W. and McDonough, L. (1995), 'A Note on the Hayek Hypothesis and the Favourite-

Longshot Bias in Parimutuel Betting', *American Economic Review*, 85, pp. 949-955.

Hurst, H. (1951) 'Long Term Storage Capacity of Reservoirs', *Transactions of the American Society of Civil Engineers*, 116, pp. 770-799.

Ibbotson Associates (1990) '*Stocks, Bonds, Bills and Inflation: 1990 Yearbook*, Chicago, Illinois: Ibbotson Associates.

Ikenberry, D. and Lakonishok, J. (1989) 'Seasonal Anomalies in Financial Markets: A Survey', in: Guimaraes, R.M.C., Kingsman, B.G. and Taylor, S.J. (1989), *A Reappraisal of the Efficiency of Financial Markets*, Springer-Verlag : Berlin.

Insurance Age (1988), 10, pp. 10-11.

Investors Chronicle (1991), 25th. November, pp. 16-17. Doing the Random Walk - Selecting a Share Portfolio.

Jaffe, J.F. (1974) 'Special Information and Insider Trading', *Journal of Business*', 47, July, pp. 410-429.

Jaffe, J.F. and Mandelkar, G. (1976) 'The "Fisher Effect" for Risky Assets: An Empirical Investigation', *Journal of Finance*, 31, pp. 447-458.

Jaffe, J. and Westerfield, R. (1985) 'The Week-End Effect in Common Stock Returns: The International Evidence', *Journal of Finance*, 40 (2), June, pp. 433-454.

Janis, I.L. and Mann, L. (1977), '*Decision Making: A Psychological Analysis of Conflict, Choice and Commitment*', Free Press, New York.

Jarrell, G.A. and Poulsen, A.B. (1989) 'Stock Trading Before the Announcement of Tender Offers: Insider Trading or Market Anticipation?', *Journal of Law, Economics, and Organization*, 5 (2), pp. 225-248.

Jegadeesh, N. (1990) 'Evidence of Predictable Behavior of Security Returns', *Journal of Finance*, 45, pp. 881-898.

Jegadeesh, N. and Titman, S. (1991)
Short Horizon Return Reversals and the Bid-Ask Spread.
Working Paper, University of California at Los Angeles.

Jegadeesh, N. and Titman, S. (1993) 'Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency', *Journal of Finance*, 48 (1), March, pp.65-91.

Jensen, M.C. (1978) 'Some Anomalous Evidence Regarding Market Efficiency', *Journal of Financial Economics*, 6, pp. 95-101.

Jensen, M.C. and Bennington, G.A. (1970) 'Random Walks and Technical Theories : Some

Additional Evidence', *Journal of Finance*, 25, May, pp. 469-482.

Johansen, S. (1988) 'Statistical Analysis of Cointegration Vectors', *Journal of Economic Dynamica and Control*, 12, pp. 231-254.

Johnson, Johnnie E.V. and Bruce, Alistair C. (1993) 'Gluck's Second Law: An Empirical Investigation of Horserace Betting i Early and Late Races', *Psychological Reports*, 72, pp. 1251-1258.

Kahneman, D. and Tversky, A. (1973) 'On the Psychology of Prediction', *Psychological Review*, 80, pp. 237-251.

Kahneman, D. and Tversky, A. (1979), Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, 47, pp. 263-291.

Kahneman, D. and Tversky, A. (1982) 'Intuitive Prediction: Biases and Corrective Procedures', in: Kahneman, D., Slovic, P. and Tversky, A. (eds.)(1982) '*Judgment under Uncertainty: Heuristics and Biases*' Cambridge : Cambridge University Press.
pp.414-421.

Kahneman, D. and Tversky, A. (1984), 'Choices, Values and Frames', *American Psychologist*, 39, pp. 341-350.

Kaplan, R. and Roll, R. (1972) 'Investor Evaluation of Accounting Information : Some Empirical

Evidence', *Journal of Business*, 45, April, pp. 225-257.

Keane, S.M. (1987) '*Efficient Markets and Financial Reporting*' Edinburgh : Institute of Chartered Accountants of Scotland.

Keane, S.M. (1991) 'Paradox in the Current Crisis in Efficient Market Theory', *Journal of Portfolio Management*, Winter.

Keane, S.M. (1993) 'Emerging Markets- The Relevance of Efficient Market Theory', Occasional Research Paper no.15, Glasgow University.

Keim, D.B. (1983) 'Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence', *Journal of Financial Economics*, 12 (1), June, pp. 13-32.

Keim, D. (1989) 'Earnings Yield and Size Effects: Unconditional and Conditional Estimates', in: Guimaraes, R.M.C., Kingsman, B.G. and Taylor, S.J. (1989), ' A Reappraisal of the Efficiency of Financial Markets', Springer-Verlag: Berlin.

Keim, D.B. and Stambaugh, R.F. (1984) 'A Further Investigation of the Weekend Effect in Stock Returns', *Journal of Finance*, 39 (3), July, pp. 819-837.

Keim, D.B. and Stambaugh, R.F. (1986) 'Predicting Returns in the Stock and Bond Markets', *Journal of Financial Economics*, 17, pp. 357-390.

Kendall, M.G. (1953) 'The Analysis of Economic Time Series, part 1 : Prices', *Journal of the Royal Statistical Society*, 96 (Part 1) pp. 11-25, in: Cootner, P.H. (ed.) (1964), *The Random Character of Stock Market Prices*, Cambridge : MIT Press.

Kim, M.J., Nelson, C.R. and Startz, R. (1991) 'Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence', *Review of Economic Studies*, 58, May, pp. 515-528.

Kopelman, R.E. and Minkin, B.L. (1991), 'Toward a Psychology of Parimutuel behaviour: Test of Gluck's Laws', *Psychological Reports*, 68, pp. 701-702.

Kraus, A. and Stoll, H. (1972) 'Price Impacts of Block Trading on the New York Stock Exchange', *Journal of Finance*, 27, June, pp. 569-588

Lakonishok, J. and Levi, M. (1982) 'Weekend Effects on Stock Returns: A Note', *Journal of Finance*, 37 (3), June, pp. 883-889.

Lakonishok, J., Shleifer, A. and Vishny, R. (1992) 'The Structure and Performance of the Monetary Management Industry', *Brookings Papers on Economic Activity*, 1992.

Lakonishok, J. and Smidt, S. (1988) 'Are Seasonal Anomalies Real? A Ninety-Year Perspective', *Review of Financial Studies*, 1 (4), Winter, pp. 403-425.

Langer, E.J. (1975), 'The Illusion of Control', *Journal of Personality and Social Psychology*, 32, pp. 311-328.

Langer, E.J. and Roth, J. (1975), 'The Effect of Sequence Outcome in a Chance Task on the Illusion of Control', Journal of Personality and Social Psychology, 32, pp. 951-955.

Langer, E.J. (1983), *'The Psychology of Control'*, Sage: London.

Le Roy, S.F. (1976) 'Efficient Capital Markets: Comment', *Journal of Finance*, 3, pp. 139-141

Leamer, E. (1978) '*Specification Searches*', Wiley: New York

Lee, C.F. and Park, H.Y. (1987) 'Value Line Investment Survey Rank Changes and Beta Coefficients', *Financial Analysts Journal*, September/October, pp. 70-72.

Lee, C.M.C., Shleifer, A. and Thaler, R.H. (1991)
Explaining Closed-End Fund Discounts.
Unpublished Manuscript, University of Michigan, Harvard University and Cornell University.

Lee, C., Shleifer, A. and Thaler, R. (1990) 'Closed-End Mutual Funds', *Journal of Economic Perspectives*, Fall, pp. 153-164.

Lee, C., Shleifer, A. and Thaler, R. (1990) 'Investor Sentiment and the Closed-End Fund Puzzle', *Journal of Finance*, 46, pp. 75-109.

Lehmann, B. (1990) 'Fads, Martingales, and Market Eficiency', *Quarterly Journal of Economics*, 105 (1), pp. 1-28.

Leonard, D.C. and Shull, D.M. (1996) 'Investor Sentiment and the Closed-End Fund Evidence: Impact of the January Effect', *Quarterly Review of Economics and Finance*, 36 (1), Spring, pp. 117-126.

LeRoy, S.F. (1989) 'Efficient Capital Markets and Martingales', *Journal of Economic Literature*, 27, December, pp. 1583-1621.

Letarte, A., Ladouceur, R. and Mayrand, M. (1986), '*Primary and Secondary Illusory Control and Risk-Taking in Gambling*', Psychological Reports, 58, pp. 299-302.

Levy, R. (1967) 'Relative Strength as a Criterion for Investment Selection', *Journal of Finance*, 22, December, pp. 595-610.

Liano, K. (1992) Macroeconomic Events and Seasonality of Risk and Return', *Applied Financial Economics*, 2, pp. 205-209.

Liano, K. and Gup, B.E. (1989) 'The Day-of-the-Week Effect in Stock Returns over Business Cycles', *Financial Analysts Journal*, 45, pp. 74-77.

Liano, K. and White, L.R. (1994) 'Business Cycles and the Pre-Holiday Effect in Stock Returns', *Applied Financial Economics*, 4, pp. 171-174.

Liano, K, Huang, G.C. and Gup, B.E. (1993) 'A Twist on the Monday Effect in Stock Returns: A Note', *Journal of Economics and Business*, 45, pp. 61-67.

Liano, K., Manakyan, H. and Marchand, P.H. (1992) 'Economic Cycles and the Monthly Effect in the OTC Market' *Quarterly Journal of Business and Economics*, 31, pp. 41-50.

Liano, K., Marchand, P.H. and Huang, G.C. (1992) 'The Hot Hand Effect in Stock Returns: Evidence from the OTC Market', *Review of Financial Economics*, 2, pp. 45-54.

Lin, J-C. and Howe, J.S. (1990) 'Insider Trading in the OTC Market', *Journal of Finance*, 45, pp. 1273-1284.

Lintner, J. (1965) 'The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets' *Review of Economics and Statistics*, 47 (1), February, pp. 13-37.

Livingston, J. (1974), '*Compulsive Gamblers*', New York: Harper and Row.

Lo, A.W. (1991) 'Long-Term Memory in Stock Market Prices', *Econometrica*, 59, pp. 1279-1313.

Lo, A.W. and MacKinlay, A.C. (1988) 'Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test', *Review of Financial Studies*, 1(1), Spring, pp. 41-66.

Lo, A.W. and Mackinlay, A.C. (1989) 'The Size and Power of the Variance Ratio Test in Finite samples: a Monte carlo Investigation', *Journal of Econometrics*, 22(7), pp. 27-59.

Lo, A.W. and MacKinlay, A.C. (1990a) 'An Econometric Analysis of Nonsynchronous Trading', *Journal of Econometrics*, 45 (1-2), July-August, pp. 181-211

Lo, A.W. and MacKinlay, A.C. (1990b) 'Data-Snooping Biases in Tests of Financial Asset Pricing Models', *Review of Financial Studies*, 3, pp. 431-468.

Lo, A.W. and Mackinlay, C. (1990c) 'When are Contrarian Profits Due to Stock Market Overreaction?', *Review of Financial Studies*, 3, pp. 175-205.

Lo, V.S.Y. and Busche, K. (1994), 'How Accurately Do Bettors Bet in Doubles?', in *Efficiency of Racetrack Betting Markets* (Eds.) Donald B. Hausch, Victor S. Y. Lo and William T. Ziemba, Academic Press, London, pp. 465-468.

Lorie, J.H. and Niederhoffer, V. (1968) 'Predictive and Statistical Properties of Insider Trading', *Journal of Law and Economics*, 11 ,April, pp. 35-53.

Losey, R.L. and Talbott, J.C. Jr. (1980), 'Back on the Track with the Efficient Markets Hypothesis', *Journal of Finance*, 35, pp. 1039-1043.

Lowenstein, R. (1996) '*Buffet: The Making of An Americal Capitalist*, Weidenfeld and Nicholson.

Lukac, L.P. and Brorsen, B.W. (1990) 'A Comprehensive Test of Futures Market Disequilibrium', *Financial Review*, 25 (4), November, pp. 593-622.

Lustgarten, S. and Mande, V. (1995) 'Financial Analysts' Earnings Forecasts and Insider Trading', *Journal of Accounting and Public Policy*, 14(3), pp. 233-261.

MacDonald, R. and Power, D. (1991) 'Stock Prices, Efficiency and Cointegration', Working Paper,Department of Economics, Dundee University, August.

Madura, J. and Wiant, K.J. (1995) 'Information Content of Bank Insider Trading', *Applied Financial Economics*, 5, pp. 219-227.

Malkiel, B. (1977) 'The Valuation of Closed-End Investment Company Shares', *Journal of Finance*, 32, June, pp. 847-859.

Malkiel, B.G. (1995) 'Returns from Investing in Equity Mutual Funds, 1971-1991', *Journal of Finance*, 50, pp. 549-572.

Mandelbrot, B. (1963) 'The Variation of Certain Speculative Prices.' *Journal of Business*, 36, October, pp. 394-419

Mandelbrot, B. (1972) 'Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis,' *Annals of Economic and Social Measurement*, 1, pp. 259-290.

Mandelbrot, B. (1975) 'Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes',

Z. Wahrscheinlichkeitstheorie verw. Gebiete, 31, pp. 271-285.

Mandelbrot, B. and Takku, M. (1979) 'Robust R/S Analysis of Long-Run Serial Correlation', *Bulletin of the International Statistical Institute*, 48, Book 2, pp. 59-104.

Markowitz, H. (1952), 'The Utility of Wealth', *Journal of Political Economy*, 60, April, pp. 151-158.

McGlothlin, W.H. (1956), 'Stability of Choices Among Uncertain Alternatives', *American Journal of Psychology*, 69, pp. 604-619.

McQuaid, C. (1971), ed., *Gambler's Digest*, Chicago: Follett Publishing Co.

McQueen, G.R. (1992) 'Long-Horizon Mean-Reverting Stock prices Revisited', *Journal of Financial and Quantitative Analysis*, 27 (1), March, pp. 1-18.

Mead, W.J., Moseidjord, A. and Sorensen, P.E. (1984) 'Competitive Bidding Under Asymmetrical Information: Behavior and Performance in Gulf of Mexico Drainage Lease Sales, 1959-1969', *Review of Economics and Statistics*, 66 (3), August, pp. 505-508.

Merton, R.C. (1987) 'On the Current State of the Stock Market Rationality Hypothesis', in: Dornbusch, R., Fischer, S. and Bossons, J. (ed.) (1987), '*Macroeconomics and Finance: Essays in Honor of Franco Modigliani*.' MIT Press, Cambridge, MA.

Metcalf, G.E. and Malkiel, B.G. (1994) The Wall Street Journal Contests: The Experts, the Darts, and the Efficient Market Hypothesis', *Applied Financial Economics*, 4, pp. 371-374.

Metzger, M.A. (1985), 'Biases in Betting: An Application to Laboratory Findings', *Psychological Reports*, 56, pp. 883-888.

Mitchell, M.L. and Lehn, K. (1990) 'Do Bad Bidders Become Good Targets?', *Journal of Political Economy*, 98, pp. 372-398.

Mitchell, M.L. and Mulherin, J.H. (1994) 'The Impact of Public Information on the Stock Market', *Journal of Finance*, pp. 923-950.

Moore, A.B. (1962) 'A Statistical Analysis of Common Stock Prices', Unpublished Ph.D. thesis, Graduate School of Business, University of Chicago, 1962.

Moore, A.B. (1964) 'Some Characteristics of Changes in Common Stock Prices', in: Cootner,P.(ed.) (1964) : *The Random Character of Stock Market Prices*, pp.262-296.

Mossin,J. (1966) 'Equilibrium in a Capital Asset Market', *Econometrica*, 34 (4), October, pp. 768-783.

Neill, H.B. (1931) '*Tape Reading and Market Tactics*.'
Forbes Publishing, New York, NY.

Nelson, C.R. (1976) 'Inflation and Rates of Return on Common Stocks', *Journal of Finance*, 31, pp. 471-483.

Nelson, C.R. and Kim, M.J. (1990) 'Predictable Stock Returns: Reality or Statistical Illusion?', Working Paper, Economics Department, University of Washington, Seattle.

Niederhoffer, V. and Osborne, M.F.M. (1966) 'Market Making and Reversal on the Stock Exchange', *Journal of the American Statistical Association*, 61, December, pp.97-916.

Panas, E. (1991) 'A Weak Form Evaluation of the Efficiency of the Rotterdam and Italian Oil Spot Markets', *Energy Economics*, 13 (1), January, pp. 26-32

Patel, J., Zeckhauser, R. and Hendricks, D. (1992), 'Investment Flows and Performace: Evidence from Mutual Funds, Cross-border Investments, and New Issues', in: Sato, R., Levich, R. and Ramachandran, R., 'Japan and International Financial Markets: Analytical and Empirical Perspectives. Cambridge University Press

Patell, J.M. and Wolfson, M.A. (1984) 'The Intraday Speed of Adjustment of Stock Prices to Earnings and Dividend Announcements', *Journal of Financial Economics*, 13, pp. 223-252.

Pesaran, M.H. and Timmermann, A. (1995) 'Predictability of Stock Returns: Robustness and Economic Significance', *Journal of Finance*, L(4), pp. 1201-1228.

Peters, E.E. (1989) 'Fractal Structure in the Capital Markets', *Financial Analysts Journal*, July-August, pp. 32-37.

Peterson, D.R. (1995) 'The Informative Role of the Value Line Investment Survey - Evidence From Stock Highlights', *Journal of Financial and Quantitative Analysis*, 30(4), pp. 607-618.

Pettit, R.R. (1972) 'Dividend Announcements, Security Performance and Capital Market

Efficiency', *Journal of Finance*, 27, December, pp. 993-1007.

Pettit, R.R. and Venkatesh, P.C. (1995) 'Insider Trading and Long-Run Return Performance', *Financial Management*, 24 (2), pp. 88-103.

Pontiff, J. (1995) 'Closed-end Fund Premia and Returns - Implications for Financial Market Equilibrium', *Journal of Financial Economics*, 37 (3), pp. 341-370.

Pratt, S.P. and De Vere, C.W. (1968)
Relationship between Insider Trading and Rates of Return for N.Y.S.E. Common Stocks, 1960-1966. Paper presented for the Seminar on the Analysis of Security Prices, University of Chicago,May. Reprinted in: Lorie, J.H., and Brealey, R.A. (ed.), '*Modern Developments in Investment Management*', 2nd.ed., Hinsdale, Illinois : Dryden Press (1978)

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

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

Quirin, W.L. (1979), '*Winning at the Races: Computer Discoveries in Thoroughbred Handicapping*', William Morrow: New York.

Raj, M. and Thurston, D. (1996) 'Effectiveness of Simple Technical Trading Rules in the Hong

Kong Futures Market', *Applied Economics Letters*, 3 (1), pp. 33-36.

Reinganum, M.R. (1982) 'A Direct Test of Roll's Conjecture on the Firm Size Effect', *Journal of Finance*, 37(1), March, pp. 27-35.

Reinganum, M.R. (1983) 'The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax-Loss Selling Effects', *Journal of Financial Economics*, 12 (1), June, pp. 89-104.

Reynolds, R. (1971), 'Everything You Should Know About Making Money At the Races,' Toronto, Canada: Pagurian Press Ltd.

Roberts, H.V. (1959) 'Stock Market "Patterns" and Financial Analysis : Methodological Suggestions', *Journal of Finance*, 14(1), pp. 1-10.

Roberts, H.V. (1959) 'Stock Market "Patterns" and Financial Analysis : Methodological Suggestions', *Journal of Finance*, 14(1), March, pp. 1-10.
In:
Cootner, P.H. (ed.) (1964)
The Random Character of Stock Market Prices.
Cambridge: MIT.

Roberts, H.V. (1967) '*Statistical versus Clinical Prediction of the Stock Market*', Unpublished paper presented to the Seminar on the Analysis of Security Prices, Univesity of Chicago, May.

Roberts, P.M. and Newton, B.A. (1987), '*The Intelligent Punter's Survey*', Weymouth: T.I.P.S. Publishing.

Rogalski, R. (1984) 'New Findings Regarding Day-of-the-Week-Returns over Trading and Non-Trading Periods: A Note', *Journal of Finance*, 39 (5), December, pp. 1603-1614.

Rogalski, R.J. and Tinic, S.M. (1986), 'The January Size Effect: Anomaly or Risk Measurement?', *Financial Analysts Journal*, November/December, pp. 63-70.

Roll, R. (1983) 'Vas ist das? The Turn-of-the-Year Effect and the Return Premia of Small Firms', *Journal of Portfolio Management*, 9 (2), Winter, pp. 18-28.

Roll, R. (1986) 'The Hubris Hypothesis of Corporate Takeovers', *Journal of Business*, 59, pp. 197-216.

Rosenberg, B., Reid, K. and Lanstein, R. (1985) 'Persuasive Evidence of Market Inefficiency', *Journal of Portfolio Management*, 11, pp. 9-16.

Rosett, R.N. (1965) 'Gambling and Rationality', *Journal of Political Economy*, 6, pp. 595-607.

Royal Commission on Gambling (1978), Final Report, vol.2, London:HMSO.

Rozeff, M.S. (1974) 'Money and Stock Prices; Market Efficiency and the Lag in Effect of

Monetary Policy', *Journal of Financial Economics*, 1, September, pp. 245-302.

Rozeff, M.S. (1984) 'Dividend Yields are Equity Risk Premiums', *Journal of Portfolio Management*, 11, pp. 68-75.

Rozeff, M.S. (1985) 'The December Effect in Stock Returns and the Tax-Loss Selling Hypothesis', Working Paper, No. 85-18, College of Business Administration, University of Iowa, May.

Rozeff, M.S. and Zaman, M.A. (1988) 'Market Efficiency and Insider Trading : New Evidence', *Journal of Business*, 61, pp. 25-44.

Rugg, D. (1986) '*Dow Jones-Irwin Guide to Mutual Funds*, 3rd.edn.
Dow Jones-Irwin: Homewood, Illinois.

Samuelson, P.A. (1965) 'Proof That Properly Anticipated Prices Fluctuate Randomly', *Industrial Management Review*, 6, pp.1-49.

Sant, R. and Zaman, M.A. (1996) 'Market Reaction to Business Week Inside Wall Street Column - A Self-Fulfilling Prophecy', *Journal of Banking and Finance*, 20 (4), pp. 617-643.

Saunders, E.M., Jr. (1993) 'Stock Prices and Wall Street Weather', *American Economic Review*, December, pp. 1337-1345.

Savit, R.A. (1988) 'When Random is not Random: An Introduction to Chaos in Market Prices', *The Journal of Futures Markets*, 8 (3), pp. 271-289.

Savit, R.A. (1992) '*Chaos on the Trading Floor*', in: '*The New Scientist Guide to Chaos*' ed. N.Hall, London : Penguin.

Schabacker, R. (1930) '*Stock Market Theory and Practice*', Forbes Publishing : New York, NY.

Schnytzer, A. and Shilony, Y. (1995) Inside Information in a Betting Market, *Economic Journal*, 105, pp. 963-971.

Scholes, M.S. (1972) 'The Market for Securities: Substitution versus Price Pressue and the Effects of Information on Share Prices', *Journal of Business*, 45, April, pp. 179-211.

Seligman, D. (1975), 'A Thinking Man's Guide to Losing at the Track', *Fortune* 92, September, pp. 81-87.

Seyhun, H.N. (1986) 'Insiders' Profits, Costs of Trading, and Market Efficiency', *Journal of Financial Economics*, 16, pp. 189-212.

Shah, M. and Wadhwani, S.B. (1990) 'The Effect of the Term Spread, Dividend Yield and Real Activity on Stock Returns: Evidence From 15 Countries', *LSE Financial Markets Group Discussion Paper*, 98.

Sharpe, W.F. (1964) 'Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk', *Journal of Finance*, 19 (4) September, pp. 425-442.

Sharpe, W.F. (1966) 'Mutual Fund Performance', *Journal of Business*, 39, January, pp. 119-138.

Shiller, R.J. (1979) 'The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure', *Journal of Political Economy*, 87(6), December, pp. 1190-1219.

Shiller, R.J. (1981) 'Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?', *American Economic Review*, 71(3), June, pp. 421-436.

Shiller, R.J. (1984) 'Stock Prices and Social Dynamics', *Brookings Papers on Economic Activity*, 2, pp. 457-510.

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

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

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

Shleifer, A. and Summers, L.H. (1990) 'The Noise Trader Approach to Finance', *Journal of*

Economic Perspectives, Spring, pp. 19-33.

Sirri, E. and Tufano, P. (1993) 'Competition and Change in the Mutual Fund Industry', in: Hayes, S. (ed.), *Financial Services: Perspectives and Challenges*. Harvard Business School Press, 1993.

Sivakumar, K. and Wagmire, G. (1994) 'Insider Trading Following Material News Events - Evidence from Earnings', *Financial Management*, 23(1), pp. 23-32.

Slovic, P. (1972), 'Psychological Study of Human Judgment: Implications for Invcestment Decision Making', *Journal of Finance*, 27, pp. 779-799.

Snyder, W.W. (1978a) 'Horse Racing : Testing the Efficient Markets Model', *Journal of Finance*, 33 (4), pp. 1109-1118.

Snyder, W.W. (1978b) 'Decision-making with Risk and Uncertainty: The Case of Horse Racing', *American Journal of Psychology*, 91(2), June, pp. 201-209.

Solnik, B.H. (1973) 'Note on the Validity of the Random Walk for European Stock Prices', *Journal of Finance*, 28, pp. 1151-1159.

Solnik, B. and Bousquet, L. (1990) 'Day-of-the-Week Effect on the Paris Bourse', *Journal of Banking and Finance*, 14, pp. 461-468.

Stael von Holstein, C.A.S. (1972) 'Probabilistic Forecasting: An Experiment Related to the Stock

Market', *Organizational Behavior and Human Performance*, 30, pp. 132-156.

Stattman, D. (1980) 'Book Values and Stock Returns', *The Chicago MBA: A Journal of Selected Papers*, 4, pp. 25-45.

Stickel, S.E. (1985) 'The Effect of Value Line Investment Survey Rank Changes on Common Stock Prices', *Journal of Financial Economics*, 14, pp. 121-144.

Stock, D. (1990) 'Winner and Loser Anomalies in the German Stock Market', *Journal of Institutional and Theoretical Economics*, 146(3), pp. 518-529.

Stottner, R. (1990)
P & F Filteranalyse, Averaging Strategie und Buy & Hold-Anlageregel--Ein Beitrag zur Efficient Market Hypothese. (Point & Figure Analysis, Averaging and Buy & Hold Strategies--Challenging the Efficient Market Hypothesis. With English summary.) *Jahrbucher fur Nationalokonomie und Statistik*, 207 (4), July, pp. 374-390.

Sundali, J.A. and Atkins, A.B. (1994) 'Expertise in Investment Analysis- Fact or Fiction', *Organizational Behaviour and Human Decision Processes*, 59 (2), pp. 223-241.

Sunder, S. (1973) 'Relationships Between Accounting Changes and Stock Prices: Problems of Measurement and Some Empirical Evidence', *Journal of Accounting Research*, 11, Supplement, pp. 1-45.

Swidler, Steve and Shaw, Ron (1995) 'Racetrack Wagering and the Uninformed Bettor: A Study of Market Efficiency', *Quarterly Review of Economics and Finance*, 35 (3) pp. 305-14.

Teger, A.I. and Pruitt, D.B. (1967), 'Components of Group Risk Taking', *Journal of Experimental Social Psychology*, 3, pp. 189-205.

Terrell, Dek (1994) 'A Test of the Gambler's Fallacy: Evidence From Pari-Mutuel Games', *Journal of Risk and Uncertainty*, 8, pp. 309-317.

Terrell, Dek and Farmer, Amy (1996), 'Optimal Betting and Efficiency in Parimutuel Betting Markets With Information Costs', *Economic Journal*, 106, pp. 846-868.

Thaler, R. (1985) 'Mental Accounting and Consumer Choice', *Marketing Science*, 4, pp. 199-214.

Thaler (1992) 'The Winner's Curse: *Paradoxes and Anomalies of Economic Life*. New York: Free Press.

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

Thiel, S.E. (1988) 'Some Evidence on the Winner's Curse', *American Economic Review*, 78 (5), December, pp. 884-895.

Tinic, S.M. and West, R.R. (1984) 'Risk and Return: January vs. the Rest of the Year', *Journal* *of Financial Economics*, 13 (4), December, pp. 561-574.

Tinic, S.M. and West, R.R. (1986) 'Risk, Return and Equlilibrium: A Revisit', *Journal of Political Economy*, 94 (1), February, pp. 126-147.

Tuckwell, R. (1981) 'Anomalies in the Gambling Market', *Australian Journal of Statistics*, December, pp. 287-295.

Tuckwell, R. (1983) 'The Thoroughbred Gambling Market: Efficiency, Equity and Related Issues', *Australian Economic Papers*, 22, pp. 106-108.

Tversky, A. and Kahneman, D. (1974) 'Judgment Under Uncertainty: Heuristics and Biases', *Science*, 185, pp. 1124-1131.

Tversky, A. and Kahneman, D. (1981) 'The Framing of Decisions and the Psychology of Choice', *Science*, 211, pp. 453-458.

Tversky, A., Slovic, P. and Kahneman, D. (1990) 'The Causes of Preference Reversal', *American Economic Review*, 80, pp. 205-217.

Ukpolo, V. (1995) 'Exchange Rate Market Efficiency - Further Evidence from Cointegration Tests', *Applied Economics Letters*, 2 (6), pp. 196-198.

Van Horne, J.C. and Parker, G.G.C. (1967) 'The Random Walk Theory: An Empirical Test', *Financial Analysts Journal*, 23, pp. 87-94.

Van Zijl, A. (1984) 'Returns and Weak Form Efficiency: Betting Markets', Victoria University of Wellington Working Paper.

Vaughan Williams, Leighton and Paton, David (1996), 'Risk, Return and Adverse Selection: A Study of Optimal Behaviour Under Asymmetric Information', *Rivista di Politica Economica*, XI-XII, November-December, 11-12, pp. 63-81.

Vaughan Williams, Leighton and Paton, David (1997a), 'Why is There a Favourite-Longshot Bias in British Racetrack Betting Markets?', *Economic Journal*, 107 (1), pp. 150-158.

Vaughan Williams, Leighton and Paton, David (1997b), 'Does Information Efficiency Require a Perception of Information Inefficiency?', *Applied Economics Letters* (forthcoming, October).

Vergin, R.C. (1977), 'An Investigation of Decision Rules for Thoroughbred Racehorse Wagering,' *Interfaces*, 8(1), November, pp. 34-45.

Vermaelen, T. and Verstringe, M. (1986) 'Do Belgians Overreact?', Working Paper, Catholic University of Louvain, Belgium, November.

Wall Street Journal (1989), 3rd.October. 'Pros Outperform Investment Dartboard in Stock Picking'.

Wall Street Journal (1992), March. Quoted in Reuters Texline service, as "Wall Street Journal" C/C 3/92.

Wall Street Journal (1993) (J), p.C1, 11th.April.
'Luck or Logic? Debate Rages on over Efficient Market Theory'.

Waud, R. (1970) 'Public Interpretation of Federal Reserve Discount Rate Changes: Evidence on the Announcement Effect',
Econometrica, 38, March, pp. 231-250.

Weitzman, M. (1965), 'Utility Analysis and Group Behaviour: An Empirical Study', *Journal of Political Economy*, 73 (1), pp. 18-26.

Wilson, J.W. and Jones, C.P. (1993) 'Comparison of Seasonal Anomalies Across Major Equity Markets: A Note.', *Financial Review*, 28, pp.107-115.

Womack, K.L. (1996) 'Do Brokerage Analysts' Recommendations Have Investment Value?', *Journal of Finance*, 51(1), March, pp. 137-167.

Woodland, Linda M. and Woodland, Bill M. (1994), 'Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market', *Journal of Finance*, 49 (1), pp. 269-279.

Working, H. (1934) 'A Random Difference Series for Use in the Analysis of Time Series', *Journal of the American Statistical Association*, 29, March, pp. 11-24.

Working, H. (1960) 'Note on the Correlation of First Differences of Averages in a Random Chain', *Econometrica*, 28, October, pp. 916-918. Reprinted in: Cootner, P. (ed.) (1964), *The Random Character of Stock Market Prices*, pp.129-131.
Cambridge: MIT Press.

Wyckoff, R. (1910) '*Studies in Tape Reading*', Fraser Publishing Company : Burlington, Vermont.

Yaari, M.E. (1965), 'Convexity in the Theory of Choice Under Risk', *Quarterly Journal of Economics*, 79, pp. 278-290.

Yates, J.F., McDaniel, L.S. and Brown, E.S. (1991) 'Probabilistic Forecasts of Stock Prices and Earnings: The Hazards of Nascent Expertise', *Organizational Behavior and Human Decision Processes*, 49, pp.60-79.

Zarowin, P. (1989) 'Does the Stock Market Overreact to Corporate Earnings Information?' *Journal of Finance*, 44, pp. 1385-1399.

Ziemba, W.T. (1994) 'Investing in the Turn-of-the-Year Effect in the United States Futures Markets', *Interfaces*, 24 (3), pp. 46-61.

Ziemba, W.T. and Hausch, D.B. (1984), '*Beat the Racetrack*', Harcourt Brace Jovanovich, San Diego.

Ziemba, W.T. and Hausch, D.B. (1986), '*Betting at the Racetrack*', Vancouver and Los Angeles: Dr.Z Investments, Inc.

Ziemba, W.T. and Hausch, D.B. (1987), '*Dr.Z's Beat the Racetrack*', Revised Edition, William Morrow & Co.: New York.

Zivney, T.L., Bertin, W.J. and Torabzadeh, K.M. (1996) 'Overreaction to Takeover Speculation', *Quarterly Review of Economics and Finance*, 36(1), Spring, pp. 89-115.

Zuber, R.A., Gandar, J.M. and Bowers, B.D. (1985), 'Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games', *Journal of Political Economy*, 93, August, pp. 800-806.