UNIVERSITY OF SOUTHAMPTON

FACULTY OF BUSINESS AND LAW

SCHOOL OF MANAGEMENT

Biased decision making in a naturalistic environment: Implications for forecasts of competitive events

David C. J. McDonald

Thesis for the degree of Doctor of Philosophy (Management)

June 2012

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF BUSINESS AND LAW

Management

Doctor of Philosophy

BIASED DECISION MAKING IN A NATURALISTIC ENVIRONMENT:

IMPLICATIONS FOR FORECASTS OF COMPETITIVE EVENTS

by David C. J. McDonald

This thesis, which is divided into five papers, explores biased decision making in naturalistic environments and its implications for the efficiency of financial markets and forecasts of competitive event outcomes. Betting markets offer a valuable real world decision making context, allowing analysis that is not possible using regular financial market data. The first paper surveys studies that have employed betting markets to investigate biased decision making and discusses why the extent of these biases is significantly less than in the laboratory.

The second paper addresses unresolved issues relating to noise trading and herding in financial markets, by showing that noise trading is associated with increased market efficiency, that the extent of herding differs depending on the direction and timing of changes in market prices, and that this results in an economically significant inefficiency. The findings of this paper have important policy implications for wider financial markets: regulatory measures to protect investors from the destabilizing effects of noise appear to be self-defeating and herding is particularly prevalent when uninformed traders perceive that informed traders are participating in the market.

The third and fourth papers address the favourite-longshot bias (FLB), where market prices under-/over-estimate high/low probability outcomes. These papers demonstrate that previous explanations of the bias are inconsistent with evidence of trading in UK betting markets by developing and testing the predictions of models that explain the bias in terms of competition between market makers and the demand preferences of bettors. Moreover, it is definitively shown that, when no market maker is involved, the bias is due to cognitive errors of traders rather than their preference for risk, because only prospect theory, and not risk-love, can explain a reduced FLB in events with strong favourites.

The final chapter explores methodological concerns relating to estimates of forecast accuracy in models of discrete choice, and arrives at a much more rigorous understanding of the value of these estimates.

Table of contents

List of tables and figures

Table

Figure

Declaration of authorship

I, **David C J McDonald**, declare that this thesis, **Biased decision making in a naturalistic environment: Implications for forecasts of competitive events**, and the work presented in it are my own and have been generated by me as the result of my own original research. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission, with the exception of a version of Chapter 1, which is to be published as a chapter entitled "Evidence of biased behaviour in betting markets" in the forthcoming title *The Economics of Gambling* (Oxford: Oxford University Press).

This thesis is based on work done by myself jointly with my supervisors, **Dr Ming-Chien Sung** and **Professor Johnnie E. V. Johnson** (School of Management, University of Southampton), who each provided advice and suggestions based on earlier drafts of the work.

Signed: ………………………………………………………………………..

Date: ………………………………………………………………………

Acknowledgements

A thesis is a very personal thing, but I surely would not have been able to complete it without the guidance and support of many individuals and organisations to whom I am extremely grateful. First, I would like to thank the EPSRC and World Bet Exchange Ltd. for sponsoring my research. I also thank the many students and staff in the School of Management and elsewhere in the University who have supported me in one way or another. In particular, my supervisors Vanessa and Johnnie for their unhesitating commitment to guiding me through this process and to my development as a person. I have learned so much from them, especially how to have a positive attitude towards everything. Thanks also to Professor William Benter for generously granting me his time to discuss and give insights on my research findings. I would also thank Pete and Tiejun for sharing their thoughts and opinions and being great fun to be around, and helping me to learn new skills and improve on some old ones (like my badminton drop shots). Also thanks to my friends Jeff, Ezra, Charlotte, Sandra, Claire, Kat, James, and Aaron, and others, for being there to take my mind off things once in a while, for my parents, for keeping my mind *on* things, and finally my brothers Howard and Ross, and extended family members Carla and Jenny.

Introduction

In contrast to normative economic theory, decision makers have been shown to be susceptible to judgemental biases when making choices (Simon, 1955). However, the vast majority of evidence for this conclusion has been obtained from laboratory-based studies. The generalizability of these results to the population as a whole has been questioned because it is well understood that controlled laboratory conditions cannot replicate the richness and complexity of real life settings, with subjects often lacking the relevant experience and meaningful incentives (Bruce and Johnson, 2003; Levitt and List, 2007). Hence, an attractive alternative for researchers is to seek out evidence of biased judgements in the natural environment of the decision maker. The decisions of traders operating in financial markets, for instance, can be analyzed by seeing that the market price of a speculative asset results from the combined decisions of all traders, and so the average subjective probability estimates of the asset's payoffs can be determined; this can be compared with objective outcomes in order to determine whether participants in the market are biased in their decision making (Griffith, 1949). However, it is difficult to assess the extent of biases in regular financial market prices because uncertainty is always present, with prices reflecting the current subjective expectation of future prices rather than objective fundamental information (Shleifer and Summers, 1990). Hence, one can never be certain whether anomalies in financial market prices truly represent biased decision making or simply reflect the expectation of possible future events that may or may not occur in practice.

Betting markets are an ideal real world decision making environment in which to explore biased decision making, sharing many features with other financial markets and offering many advantages over laboratory experiments. In particular, they involve a large number of regular traders who have access to widely available information, a smaller number of traders who are particularly adept at combining information in such a way as to make excess returns, and a minority of traders who have access to privileged information (Vaughan Williams and Paton, 1997). In contrast to laboratory conditions, betting markets are associated with rich, dynamic information sets, offer strong incentives for success, require the commitment of the individual's own resources, and involve repeated trials. Most importantly, they have a defined end point at which all uncertainty in the relation between prices and fundamental information is resolved, thus overcoming the problem of expectation in regular financial markets. In addition, "in its simplest formulation, the market for bets in an *n*-horse race corresponds to a market for contingent claims with *n* states in which the *i*th state corresponds to the outcome in which the *i*th horse wins the race" (Shin, 1993, p.1142). Betting markets also offer the opportunity of quantifying the proportion of market activity attributable to informed trading (using a model developed by Shin, 1993). Consequently, betting markets appear to offer considerable advantages for the study of biased decision making in a naturalistic environment.

This thesis is divided up into five separate but inter-related papers. The theme running through all the papers is that they examine to what extent and why decision makers make biased choices in a naturalistic environment and the implications of such biased behaviour for the efficiency of financial markets and forecasts of competitive event outcomes. Throughout, these papers develop new insights relating to behavioural and economic biases in general by developing innovative models and carrying out empirical tests on recent data drawn from UK betting markets. Overall, this thesis makes a significant contribution towards understanding the extent and nature of biased decision making in naturalistic environments.

Chapter 1, entitled "**Evidence of biased decision making in a naturalistic environment",** a version of which is to be published in the forthcoming title *The Economics of Gambling* (Oxford: Oxford University Press), is a discussion of previous studies that have employed betting markets to study biased decision making, with particular reference to systematic biases that were first discovered in the laboratory. Hence, it serves as an overall literature review for the remainder of the thesis, although, as each chapter can be considered a stand-alone paper, relevant literature is covered at the beginning of each subsequent chapter. This paper begins by noting (as have a number of other researchers, e.g., Ebbeson and Konecni, 1980; Funder, 1987) that, while there is a substantial literature consisting of evidence of biased decision making, such as the work of Simon (1955) on bounded rationality and Kahneman and Tversky on heuristics and biases (Kahneman et al., 1982), most of this research has been carried out in controlled laboratory conditions with subjects who are often students lacking the relevant experience for the task at hand. It outlines a number of reasons why there is a problem with generalizing the results of these studies to the wider population and shows how studies of betting markets overcome this problem. Then it summarizes and discuss the results of betting market studies of a number of important decision making biases: the favourite-longshot bias, anchoring and adjustment, herding, the gambler's fallacy, and the hot hand.

Chapter 2 consists of a paper that will be submitted to *The Journal of Financial Markets*: "**Noise, herding, and the efficiency of market prices: insights from markets for state contingent claims**". This paper shows how data from betting markets can be employed to provide new insights into a number of unresolved issues relating to two types of biased decision making that are prominent in the financial markets literature. First, it explores the role played in the efficiency of financial markets by noise trading, which is trading that is based on anything except information, and so appears to be a universally loss-making strategy (Black, 1986). An unanswered question has been whether noise trading results in excessively volatile, inefficient markets, in which added risks limit the possibility of arbitrage by informed investors (De Long et al., 1990), or if noise is essential in providing liquidity to informed investors in order that markets are efficient (Bloomfield et al., 2009). It has been predicted and, to an extent, verified empirically (Campbell and Kyle, 1993) that a consequence of noise trading is increased volatility in market prices. So, on the one hand, it has been argued that noise results in volatility in excess of the variations justified by underlying fundamental information (Shiller, 1990), in which case, noise is detrimental to market efficiency because of its destabilizing effect on long-run equilibrium values, with risk-averse informed investors limiting their arbitrage to avoid liquidity risks (Shleifer and Summers, 1990). On the other hand, it has been suggested that noise may be essential for generating liquid, and thus efficient, markets (Black, 1986; Grossman and Stiglitz, 1980; Kyle, 1985). Noise trading might harm market efficiency, but only when prices are extreme; when prices are not extreme, noise traders help to make prices *more* efficient by providing liquidity to informed traders (Bloomfield et al., 2009). The first aim of this paper is to answer these questions: are markets associated with a greater degree of noise trading also more volatile, and are these markets more or less efficient?

Second, this paper investigates the related decision making bias of herding, which occurs when market participants neglect their own private information and adjust their actions to be more representative of those of other traders. They do this in the belief (perhaps mistaken) that other traders are more informed than themselves. The combined activity of many herding traders can result in extraordinary changes in asset values over a short period, possibly leading to bubbles, crashes and bank runs (Devenow and Welch, 1996). However, empirical evidence for the phenomenon is inconclusive in both financial markets (Lakonishok et al., 1992; Wermers, 1999; Sias, 2004) and laboratory investigations (Cipriani and Guarino, 2005; Spiwoks et al., 2008). One reason for this mixed evidence could be that previous studies failed to take account of differing levels of actual and perceived trading by investors with priviliged information at different times of the market as well as depending on the direction of price movements. The second aim of this paper, therefore, is to explore the extent and reasons for herd behaviour while accounting for these concerns.

The third and fourth papers address the favourite-longshot bias (FLB), which is the widely-reported phenomenon wherby prices in markets for state contingent claims systematically under-/over-estimate high/low probability outcomes (e.g., Dowie, 1976; Ali, 1977; Snowberg and Wolfers, 2010). It has been attributed to a wide range of causes, such as the risk-loving nature of traders (Weitzman, 1965), errors in the estimation of probabilities (Henery, 1985), the pricing policies of bookmakers (Shin, 1993), and limited information of traders (Sobel and Raines, 2003). Each of these papers demonstrates that previous explanations of the bias are inconsistent with theoretical and empirical considerations relating to recent trading in UK betting markets. The first of these papers addresses bookmaker and exchange markets and how competition between these markets, along with the demand preferences of bettors, can provide a more satisfactory explanation for the presence (or absence) of the bias in these markets. The second of these papers addresses pari-mutuel markets and investigates using a novel methodology whether, when no market maker is involved, the bias is due to cognitive errors of traders rather than their preference for risk.

Chapter 3 is titled "**The favourite-longshot bias in competing betting markets**" and consists of a paper that is under review by *The Economic Journal*. Many studies have sought to explain the enduring presence of the FLB and its absence in some markets, but little consensus has been reached (Jullien and Salanié, 2008), particularly with respect to the presence (and absence) of the FLB in the two major competing types of horserace betting market in the UK and Ireland (and in other jurisdictions, such as Australia): bookmakers and betting exchanges. In order to explain observed patterns of the extent of the FLB at various times in these markets, this paper explores two aspects of these parallel markets: competition and informed trading. The markets for horserace betting in the UK are increasingly competitive, with the rapid rise of exchange betting, along with many existing operators and a wealth of information regarding prices available to bettors via internet price comparison services. This paper develops a theoretical model to investigate the optimal pricing decisions of bookmakers when the betting public are able to rapidly compare prices, and also it is argued that informed trading has a significant effect on reducing the degree of the FLB in either type of market, but only when transaction costs are low. These considerations lead to hypotheses, which are tested empirically using a novel and unique dataset, consisting of over 5.5 million market prices: specifically, this paper analyzes how the bias develops over the course of the markets for 6058 races run between August 2009 and August 2010.

Chapter 4 ("**New evidence for a prospect theory explanation of systematic decision making bias in a market for state contingent claims**") will be submitted to *Economics Letters*, and hence is a concise but definitive account of whether cognitive errors of traders or risk preferences better explain the FLB in pari-mutuel markets (i.e., markets that are independent of a market maker). While there have been a range of explanations proposed for the

FLB in pari-mutuel markets, a simplified but useful categorization is that bettors either have unbiased expectations, but are risk-loving, or have biased expectations, but are risk-neutral or risk-averse (Snowberg and Wolfers, 2010). However, it is difficult to empirically discriminate between these competing explanations because each is observationally equivalent. Those that have attempted this task have relied on parametric assumptions or assumptions relating to the choice set of the decision maker. This paper develops a new methodology for choosing between the hypotheses that does not rely on these assumptions, consisting of a representative agent model that predicts a 'strong favourite' effect on the level of FLB. Specifically, it is demonstrated that the level of bias in an individual event varies in a predictable manner depending on the traders' risk preferences. If the representative agent is risk-seeking, the model predicts an increased FLB in events where the variance of the odds on competitors other than the favourite is relatively low, *ceteris paribus*. Convesely, if the representative agent is risk-averse, the level of FLB is reduced or a reverse bias is predicted when the same variable is relatively low. This prediction is independent of whether probabilities enter the decision process linearly (as in expected utility theory) or nonlinearly (as in prospect theory). Hence, empirical tests can be conducted that distinguish between hypotheses that do and do not require the representative agent to be risk-loving. The purpose of this paper therefore, is to test empirically the predictions of the model and definitively show whether expected utility theory or prospect theory better explains the FLB.

Chapter 5 consists of a paper with the title, "**Properties of pseudo-R^2 as an estimate of forecast accuracy for discrete choice models**", and is primarily concerned with methodological issues relating to statistical methods employed in studies that employ models of discrete choice, of which the conditional logit model, employed throughout this thesis, is an example. Hence, this paper will be submitted to *The Journal of the Royal Statistical Society, Series B (Methodological)*. The motivation for this paper is that in many studies that have used discrete choice models, particularly many of the studies cited in this thesis, pseudo-R^2s are reported as a measure of forecast accuracy. However, it is shown in this paper that there are significant concerns related to the evaluation of discrete choice models using pseudo-R^2s. A key property of any means of evaluation of a forecast is its comparability across empirical models, both in the sense that its interpretation should be the same, and that standard errors should be reported, in order to carry out significance tests. However, little attention has previously been given to whether pseudo-R^2s have this desirable property, and the consequences of this for the evaluation of discrete choice models using pseudo-R^2s. Consequently, there are three broad research questions that are asked in this paper. First, are pseudo-R^2s directly comparable across models estimated on datasets with different characteristics? Second, if they are, what is the most appropriate method for statistical comparisons? Finally, how useful are pseudo-R^2s in explaining the predictive power of model probabilities?

1. Evidence of biased decision making in a naturalistic environment

Abstract

The generalizability of laboratory-based research into behavioural biases has been questioned because it is well understood that laboratory experiments cannot replicate the richness and complexity of real life settings. Naturalistic environments and betting markets in particular therefore offer an attractive alternative for examining decision making behaviour. This paper discusses the results of studies that have employed betting markets of various kinds to investigate decision making, with particular reference to systematic biases that were first identified in the laboratory. The primary conclusion of this paper is that, while systematic biases reported in the laboratory have been found in naturalistic betting markets, the extent and generality of these biases in these real world environments is often significantly less. We attribute this to the context of the decision task, the incentives offered, the lack of scrutiny involved, the experience of the decision makers, and the effect of aggregation.

1.1. Introduction

Psychologists have long been aware of the limitations of normative models of judgement and decision making. Herbert Simon's (1955) work on bounded rationality criticized rational models of decision making for disregarding factors such as the limited cognitive capacity of individuals. Later, psychologists confirmed experimentally that decisions are systematically biased in many ways, with decision makers adopting rules of thumb or 'heuristics' in order to more rapidly solve complex problems (Kahneman et al., 1982). However, over the decades, it has become apparent that many of these conclusions have relied on experiments carried out under controlled laboratory conditions. This has led researchers to question the generalizability of the results (e.g., Bruce and Johnson, 2003; Levitt and List, 2007). In particular, it is well understood that laboratory experiments cannot replicate the richness and complexity of real life settings. As Festinger (1953, p.141) notes: "In the most excellently done laboratory experiment, the strength to which the various variables can be produced is extremely weak compared to the strength with which these variables exist and operate in real life situations". Naturalistic environments offer an attractive alternative for examining decision making behaviour, featuring subjects who, unlike in many laboratory experiments, are experienced in the task at hand and are not directly aware that their actions are under scrutiny.

Betting markets are naturalistic decision making environments that offer considerable potential for understanding decision making behaviour, sharing many features with other real world decision environments. In particular, they are associated with rich, dynamic information sets, offer strong incentives for success, require the commitment of the individual's own resources, and involve repeated trials. This paper provides a survey of previous studies that have employed betting markets of various kinds to investigate decision making, with particular reference to systematic biases that were first identified in the laboratory.

The remainder of this paper is structured in three main parts. First, we summarise the debate over the generalizability of laboratory findings and identify ways in which naturalistic environments offer an alternative for studying biased decision making. In particular, we outline the usefulness of betting markets and review a range of studies that have demonstrated that bettors are, in many ways, rational and well calibrated decision makers. Second, we discuss the psychology behind the widely documented favourite-longshot bias. Third, we address two decision biases, anchoring and herding, each of which involve judgements of some unknown quantity being unduly influenced by external stimuli. Finally, we survey studies that have investigated biases that result from a failure of individuals to recognise randomness: the gambler's fallacy and the hot hand fallacy.

1.2. Betting markets: decision making in a naturalistic environment

1.2.1. The generalizability of laboratory findings

At the heart of this discussion is the distinction between experiments carried out under controlled conditions in artificial laboratory settings and analysis of data obtained from naturalistic environments, such as casinos, lotteries, and markets for betting on horseraces or other sports. While some experiments can claim to have been carried in 'real world' environments, we define a naturalistic environment as one that "has not been artificially manipulated (i.e., a nonexperimental setting)" (Johnson and Bruce, 2001, p.266). This distinction is crucial, and there is a long running debate concerning the relative merits of the two alternative methodologies when employed in experimental psychology (e.g., Ebbeson and Konecni, 1980; Hogarth, 1981; Funder, 1987; Bruce and Johnson, 2003) and economics (Harrison and List, 2004; Levitt and List, 2007). A critical assumption in experimentation is that results generalize to the broader population, but this generalizability, or 'external validity', has been questioned because of significant variations in observed behaviour between laboratory and naturalistic environments (e.g., Ebbeson and Konecni, 1980; Koehler, 1996). The factors that have been identified as limiting the generalizability of laboratory experiments include the following:

1. *Context*: Laboratory environments often present simplified versions of tasks that are more complex in real world environments, and so may unintentionally omit variables that are influential in the natural setting. Significant differences in behaviour may depend only on small changes to the experimental conditions (Ayton and Wright, 1994), and Harrison and List (2004, p.1010) note that

> although it is tempting to view field experiments as simply less controlled variants of laboratory experiments, we argue that to do so would be to seriously mischaracterize them. What passes for "control" in laboratory

experiments might in fact be precisely the opposite if it is artificial to the subject or context of the task.

In addition, there are variables that the experimenter cannot control, such as past experiences or social norms, that can affect the results (Levitt and List, 2007). Finally, biases recorded in the laboratory may simply be a response to the specific laboratory stimulus, with those same biases not occurring under ordinary circumstances (even while resulting from the same cognitive processes). For example, when mistakes are made in visual perception tasks in the laboratory, it is usually assumed that the mechanisms that result in the error generally produce correct judgements in real life (Funder, 1987).[1]

2. *Experience*: Laboratory studies typically use university students, who may be inexperienced in the task at hand. Hogarth (1981) highlights the importance of feedback in making correct decisions over the continuous time period often associated with real world decision making. Feedback is simply not available in 'one-shot game' laboratory studies, so there is limited potential for participants to learn from their mistakes. Even worse, they frequently carry 'baggage': behaviour learned in the outside world entirely unsuited to the problem at hand (e.g., Burns, 1985). Furthermore, a number of studies demonstrate large differences between the decision strategies of experts and novices in terms of the way they think, the nature of the decision models they employ, and the speed and accuracy of their problem solving (e.g., Larkin et al., 1980).

3. *Scrutiny*: Laboratory subjects, who are usually aware that they are being investigated, may be keen to project a particular image (even if they have no

[1] As a further example of the importance of context in decision making, consider the following problem. There are four cards on the table, each with a letter on one side and a number on the other. The rule is, "If there is a vowel on one side of a card, then there is an even number on the other side." The cards show A, D, 4, and 7. Which cards must be turned over in order to determine whether the rule is true or false? This is known as Wason's four-card selection task (Wason, 1968), and usually less than 10% of people respond with the correct answer of A and 7 (most neglect to choose 7, or unnecessarily include 4). However, when this problem is reframed in terms of certain social contexts, such as asking subjects to test the rule, "If a person is over 18, they can drink alcohol", and replacing the cards with '16 years old', '22 years old', 'Coke', and 'Beer', the correct answer ('16 years old' and 'Beer') is given by most respondents, even though the problem is logically identical to the first, more abstract, task (e.g., Cox and Griggs, 1982).

idea of the purpose of the experiment). The student volunteers studied in most investigations are more likely to be 'scientific do-gooders' (e.g., interested in the research, or seeking approval from the experimenter) with unusually high awareness of the moral implications of their decisions. Scrutiny may therefore exaggerate the importance of pro-social behaviours such as altruism and fairness (Levitt and List, 2007). Conversely, the anonymity that is often present in real settings may allow decision makers to feel that they are able to avoid being judged morally.

4. *Incentives*: Laboratory experiments are usually conducted with relatively trivial rewards for success. However, in the real world, decision makers are often involved in high-stakes environments, where they must commit their own or others' resources and where the results of their decisions can have significant personal consequences. These high-stakes environments can, therefore, involve a meaningful degree of risk. This can lead to a marked difference in risk taking behaviour between laboratory and real world environments (Yates, 1992). For example, the lack of excitement and low arousal levels in laboratory studies may lead to behaviours that would not be present in real settings (Anderson and Brown, 1984).

The issues discussed above all limit the generalizability of biased behaviour often found amongst laboratory participants to the wider population. However, to discard laboratory findings outright would be naive (Hogarth, 1981). Rather, data gathered in the laboratory and under naturalistic conditions have their own strengths and weaknesses, so should be considered complementary (Keren and Wagenaar, 1985). For instance, naturalistic work suffers from the inability to use control groups and difficulties associated with the replication of results. In addition, laboratory-based investigations are usually more cost-effective and afford the possibility of isolating specific variables.

1.2.2. Betting markets as valuable naturalistic environments

Betting markets offer an ideal naturalistic environment in which to explore biased decision making. A key practical advantage is the availability of extensive and detailed quantitative data relating to bettors' decisions. Since markets are finite in nature, there is a continually expanding set of 'completed' markets, i.e., self-contained time periods of betting with pre-defined endpoints, when all bets are settled in an unambiguous manner[2]. Furthermore, there is potential for comparative analysis across different types of event or bet, according to criteria such as quality (e.g., Smith et al., 2006), time of day (e.g., McGlothlin, 1956), or complexity (e.g., Johnson and Bruce, 1998). Thus, it is possible to control for some aspects of the decision setting. Most importantly, betting markets include many of the factors regarded as distinctive to naturalistic decision making (Orasanu and Connolly, 1993): uncertain dynamic environments, poorly-structured problems, high stakes, time stress, action/feedback loops, and multiple players. Each element of the decision making event (i.e., the bet) is unique: no two horseraces or football matches are the same. Thus, the outcome is uncertain, and the information relating to that outcome is often (as it is in many real world decision environments) ambiguous, vague, or redundant. For example, it is not obvious how to combine the various factors that might enable one to predict participants' performance. The dynamic nature of betting markets is evidenced by the constantly-updating prices as bettors with diverging opinions participate in the market.

Bettors, like many decision makers in real world environments, often risk meaningful amounts of money while under stress from time pressures (the window of opportunity in a betting market may only last minutes, or even seconds). A further important feature of these markets is the repetitive nature of betting. Since events take place regularly and often, there is potential for gaining familiarity with and expertise in the task. Betting markets involve action-feedback loops; once bets have been placed and a market is closed and

[2] This is a particular advantage of betting markets over other types of financial market for naturalistic research. The payoffs in betting markets are entirely unambiguous, so there is a time when all uncertainty is resolved. This is not the case in regular financial markets, where prices continuously represent the current expectation of future prices.

decided, bettors receive relatively unambiguous feedback on the success of their decisions, and this can be incorporated into future decisions (Goodman, 1998). Also, betting markets involve multiple players, and it has been shown that the interaction between individuals in markets can significantly reduce errors (Wallsten et al., 1997). This results from a variety of causes, not least the fact that different individuals use different decision making procedures and have diverse information gathering skills. As a result, their reaction to the same information may vary. Consequently, the final prices that emerge in these markets take into account a wide range of information and the forecasts of many individuals, and studies show that combining diverse forecasts generally leads to significantly more accurate predictions (e.g. Vlastakis et al., 2009; Grant and Johnstone, 2010). In addition, betting markets are not subject to the limitations of laboratory investigations listed above. For example, bettors are unaware that their decisions may be scrutinized, as they are not directly volunteering to take part in an experiment; instead, betting patterns are analysed in such a way as to observe their decisions unobtrusively.

1.2.3. Betting market data and decision making

In any betting market, individuals are able to place bets on one or more of a set of outcomes of some future event. For instance, in the simplest of markets for betting on a horserace with n runners, n different bets are available, one for each horse to win the race. After the market has closed and the race has taken place, each bet pays, for each £1 staked, a return £R_i if horse i wins and nothing otherwise. While the returns R_i (usually referred to as the 'odds' against outcome i) are determined differently according to the type of market and event, they depend on the relative amounts bet on each outcome by all the market participants. Consequently, bettors have an incentive to continue to place money on each outcome until the returns reflect the market's best estimate of that outcome's probability of occurring (Figlewski, 1979). Therefore, a typical approach to assessing decisions in betting markets is summarized (with reference to horserace betting) by Griffith (1949, p.290):

> the odds on the various horses in any race are a functioning of the proportion of the total money that is bet on each and hence are socially determined. On the other hand, the objective probability for winners from any group of horses is given a posteriori by the percentage of winners. Thus the odds express (reciprocally) a psychological probability while the percentage of winners at any odds group measures the true probability; any consistent discrepancy between the two may cast light not only on the specific topics of horse-race betting and gambling but on the more general field of the psychology of probabilities.

Hence, the 'socially determined' prices in betting markets reflect the 'subjective probabilities' assigned to each possible outcome by the bettors, in aggregate. The results of the event then determine the 'objective probabilities'. Thus, biases can be detected by researchers by comparing subjective and objective probabilities.

A drawback of most betting market research is that, for ethical and/or practical reasons, it is usually not possible to obtain information relating to the decisions of individual bettors. Instead, subjective probabilities are an aggregation of opinions of all bettors. Hence, it is possible that ''certain biases present in an individual bettor's decisions are being counterbalanced by opposite biases in other bettors' decisions'' (Johnson and Bruce, 2001; p. 280). Camerer (1987, p.982) notes that a common argument for the rationality of market participants is that "random mistakes of individuals will cancel out", but also offers the counter-argument: "biases found by psychologists are generally *systematic* - most people err in the same direction". Thus, the best we can hope for in betting market research is evidence of systematic bias.

A further weakness of employing betting market data to examine decision making behaviour is that psychologically significant biases also hold an economic significance. Consequently, if some bettors (even a small group) become aware of an overall disparity between subjective and objective probabilities, they can potentially profit by betting against the bias. This could reduce the extent to which any systematic bias that exists amongst bettors is detectable from aggregate betting market data. Fortunately for researchers, transaction costs ensure that it is rarely possible to entirely arbitrage away biases.

1.2.4. Calibration of bettors' judgements

Given the above discussion, it might be expected that bettors display significantly less biased judgements in their natural domain than is demonstrated amongst naive participants in laboratory experiments. Indeed, Rosett (1965) found that horserace bettors are generally sophisticated and rational agents, who will not forego combinations or sequences of bets that offer higher probability of winning for the same return or higher return for the same probability of winning. Furthermore, results reveal a high correlation between subjective and objective probabilities, suggesting that bettors are familiar with their decision making environment and are able to accurately forecast risky outcomes[3]. Rosett (p.596) notes that

> if these gamblers behave as though they know statistical prediction methods and the probability calculus, it seems reasonable to suppose that, in a variety of other circumstances, human beings can be expected to respond appropriately to risky situations merely after having had sufficient experience with them.

Johnson and Bruce (2001) also investigated the calibration of horserace bettors' subjective probability judgements. They found that bettors' subjective probabilities are not significantly different from the observed objective probabilities. They noted that, while there is substantial evidence of poor calibration in decision makers, this may reflect on the specific laboratory experiments involved. For example, Shanteau (1992) suggests that task characteristics may account for differences observed in the quality of experts' judgments; specifically, more competent performance is likely if the decisions involve stimuli that are relatively constant, the tasks undertaken are repetitive,

[3] An exception holds for objective probabilities of less than 0.05, which is the favourite-longshot bias detailed in the second part of this chapter.

and where decision aids are widely available. Furthermore, it has been empirically observed that violations of rationality are reduced under the multiple play conditions that exist at the racetrack (e.g., Keren and Wagenaar, 1987). Johnson and Bruce's (2001) study therefore suggests that bettors are skilled in a similar way to weather forecasters, who are also required to make frequent risky forecasts (Murphy and Brown, 1984). Hoerl and Fallin (1974) also found no significant difference between subjective and objective probabilities in horseraces, and suggested that this is due to the high incentives available for successful gambling.

Not only are bettors well calibrated in general, they are able to adapt remarkably well to uncertain and dynamic information. Johnson, O'Brien, and Sung (2010) set out to test Gigerenzer's (2000) assertion that evolution has equipped humans to process probabilistic information from frequencies observed in a natural environment. They investigated the extent to which horserace bettors accounted for post position bias (an advantage/disadvantage afforded to the horses depending on their position in the starting stalls), a factor shown to be a particularly important determinate of race outcomes at the racetrack examined. Despite the fact that track managers employed a variety of procedures to alter the bias (even between two consecutive races, and often unannounced) bettors were able to account for most of the changes through regular outcome feedback over 6 years. This finding may be accounted for by the fact that (i) bettors have a strong motivation to make accurate probability judgements as their own financial resources and often their peer group esteem depend on the outcome of their decisions (Saunders and Turner 1987), and (ii) those who frequently make probability judgments are often better calibrated (Ferrell 1994). It has also been shown that bettors' calibration is generally improving over time (Smith and Vaughan Williams, 2010) and that expert bettors employ complex mental models encompassing a wide range of variables and interactions between these variables (Ceci and Liker, 1986).

In summary, naturalistic environments such as betting markets offer rich, complex settings in which to examine decision making biases that have been observed in the laboratory. Due to a number of factors, bettors appear to be more rational, well calibrated, and able to adapt to dynamic information than participants in laboratory studies. However, there are a number of ways in which bettors are biased, and the first, and most widely documented, of these is the favourite-longshot bias, which is the focus of the next section.

1.3. The favourite-longshot bias

By far the most reported departure from rationality in the betting market literature is that of the favourite-longshot bias (FLB). Prevalent over many decades and in many jurisdictions around the world, the bias is the phenomenon whereby returns to bets are such that the chances of high probability outcomes (favourites) are underestimated, while low probability outcomes (longshots) are overestimated.

1.3.1. Evidence from the laboratory

Preston and Baratta (1948) provided early laboratory evidence of the bias. They were concerned that 'rational' theories of behaviour could not universally explain peculiarities in the way people approached 'wagering games' (in which participants are required to bet on an uncertain outcome). They hypothesized that players apply to outcomes a scale of 'psychological' probabilities not necessarily identical to their mathematically correct probabilities. In order to investigate this possibility, they carried out games with undergraduate students and faculty members (the latter were more experienced in mathematics, statistics, and psychology). The game required the participants to compete against each other, bidding for the chance to win a given prize with a given probability. They found that the players tended to pay too generously for outcomes with low probabilities and not bid high enough for outcomes with high probabilities. The indifference point, where the psychological and mathematical probabilities coincided, was found to be about 0.20. Moreover, even the faculty members displayed the bias (although to a lesser extent than the undergraduates), despite in many cases appearing to actively employ mathematics when forming their decisions. The experimental findings of Preston and Baratta have since been confirmed in numerous laboratory experiments (e.g., Yaari, 1965; Rosett, 1971; Lichtenstein et al., 1974; Piron and Smith, 1994).

1.3.2. Evidence from betting markets and explanations

The first evidence of the FLB in betting markets was from Griffith (1949). He was inspired by the laboratory evidence of Preston and Baratta (1948), but keen to test the results outside of the laboratory. Employing racetrack data, Griffith found that horses with low and high probabilities of winning were systematically over- and under-valued, respectively, a result consistent with that of Preston and Baratta, with a similar indifference point of about 0.20. McGlothlin (1956) replicated (and expanded upon) Griffith's study with a larger data set, and found similar results.

In the decades that followed, a significant body of evidence for the bias emerged in betting markets around the world (e.g., in the USA: Ali, 1977; Asch et al., 1982; in the UK: Dowie, 1976; Vaughan Williams and Paton, 1997; in Australia and New Zealand: Tuckwell, 1983; Gandar et al., 2001)[4]. The emphasis in the research then shifted towards attempting to explain the origins of the bias. As a result, a broad range of explanations have been offered, including the 'bragging rights' associated with holding a winning longshot ticket (Thaler and Ziemba, 1988) or the additional excitement derived from longshot betting (Bruce and Johnson, 1992). Henery (1985) suggests that bettors may discount a fixed proportion of their losing bets, leading them to believe that longshot bets are more attractive. Alternatively, the bias may arise from particular characteristics of the market itself, such as the cost of obtaining information and transaction costs (Hurley and McDonough, 1995) or the defensive pricing policies adopted by bookmakers (Shin, 1991). In this paper we simply provide an overview of the significant debates concerning the origins of the FLB from the perspective of bettors' decision behaviour; for more comprehensive explorations, see surveys by Thaler and Ziemba (1988), Sauer (1998), Vaughan Williams (1999), Jullien and Salanié (2008), and Ottaviani and Sørensen (2008).

[4] Exceptions have been reported in the horserace betting markets in Hong Kong (Busche and Hall, 1988; Busche, 1994), the market at one racetrack in the US (Swidler and Shaw, 1995), and exchange betting markets in the UK (Smith et al., 2006).

1.3.3. Do bettors love risk or do they misestimate probabilities? Expected utility theory vs. prospect theory

One strand of the FLB literature in particular merits attention because it has led to an important intellectual debate concerning the relative merits of two prominent competing theories for explaining decision making in wider fields: *expected utility theory* and *prospect theory*. The building block for this debate is the 'representative bettor'. Weitzman (1965) introduced Mr. Avmart, a fictitious person who represents the 'social average' of all bettors. Weitzman's (p.26) innovation was to infer the preferences of the 'most typical' bettor from the population of bettors:

> instead of concentrating on individuals and trying to derive utility generalizations from their experimental behavior, more nearly the converse approach was attempted. A plethora of data concerning the collective risk actions of parimutuel bettors was employed in investigating utility aspects of the behavior of a hypothetical member of the group.

Weitzman was concerned primarily with constructing Mr. Avmart's utility of wealth curve (the mathematical representation of preferences over various monetary outcomes and the basis of expected utility theory). He found that the FLB in the data was best explained by a convex utility of wealth curve, indicating that the average bettor is locally risk-loving (i.e., preferring the riskier, low probability outcomes). Quandt (1986) extended the analysis by showing that the bias is the natural result of equilibrium in a market where the average bettor is risk-loving. The findings of Ali (1977) and Hamid et al. (1996) also supported this hypothesis.

However, there are alternative scenarios that can explain the biased decisions of the representative bettor. So, for instance, Golec and Tamarkin (1998) showed that the FLB can arise if bettors are risk-averse in general but with a preference for skewness of returns. An alternative explanation stems from Preston and Baratta's (1948) supposition that 'psychological' probabilities assigned to uncertain outcomes are systematically biased. If this is the case, the

FLB can be explained solely with reference to bettors' systematic misestimation of probabilities, and bettors need not be locally risk-loving. This is formalized in Kahneman and Tversky's (1979) *prospect theory* (later extended and renamed cumulative prospect theory; see Tversky and Kahneman, 1992). The important feature of prospect theory for this discussion is that objective probabilities are transformed into subjective decision weights that allow for biases in the estimation of probabilities.

Hence, there are now two broadly competing sets of theories regarding the explanations for the bias in terms of the representative bettor: are bettors unbiased in their estimation of probabilities, but risk-loving, or are bettors risk-neutral, but biased in their estimation of probabilities? Unfortunately, there is no straightforward answer. As Yaari (1965, p.278) comments:

> at first blush it seems as though one cannot, by looking at empirical data, choose between the two hypotheses (distortion of utility versus distortion of probability) because utility and probability are two purely theoretical components of an integral decision process. Thus, the two hypotheses are empirically indistinguishable, and choosing between them is a matter of taste.

However, some researchers have made progress in this regard. Golec and Tamarkin (1995) noted that risk-love cannot explain the relatively unfair returns for the low risk, low return side bets offered by some bookmakers. Instead, they suggested that overconfidence better explains the FLB (a conclusion consistent with bettors overestimating small probabilities). Jullien and Salanié (2000) found that prospect theory better explains the bias for standard bets, although computational limitations of this approach restricted their analysis. Bradley (2003) adapted the approach of Jullien and Salanié by accounting for bet size and found an even better fit to the data.

More recently, Snowberg and Wolfers (2010) set out to test the competing theories using a novel approach and a large dataset of all the horseraces run in North America from 1992 to 2001 (over 865,000 races). They first estimated the parameters of the two models by fitting them to standard 'win' bets. They then examined compound exotic bets, such as the 'exacta', a bet that two horses will finish a race in first and second place in a specific order. Snowberg and Wolfers reasoned that bettors would bet in the same manner in the exotic and win betting pools, so the same models should apply for each bet type. Accordingly, they used the fitted models to predict expected market prices in the exotic betting pools and compared their predictions with the actual prices on offer. They found that the misestimation of probabilities model predicted exotic bet prices more accurately than the risk-love model, and concluded that, with respect to the representative bettor, prospect theory explained the FLB more effectively than expected utility theory.

An important issue in this debate is the validity of the assumption that bettors' decisions can be averaged by the representative bettor. In the third section of this paper, we show that the distinction between different types of bettor, on the basis of the quality of the information they hold or how they handle this information, is crucial to fully understanding another bias in betting behaviour. Sobel and Raines (2003) demonstrated this by differentiating between 'serious' and 'casual' bettors. They identify serious bettors as those that attend the racetrack on week nights, bet larger sums, and bet to a greater extent on more complicated types of bet. On the other hand, casual bettors attend primarily on weekends and bet smaller sums on simpler types of bet. Sobel and Raines found evidence of the FLB, but the bias was significantly reduced in those races that involved a higher proportion of serious bettors.

1.3.4. The late-race effect

A curious element of the nature of the FLB is its apparent tendency to vary in a systematic manner over the course of a day's betting activity. In horserace betting markets, in particular, it has been found that the extent of the bias appears to increase significantly in the last race or last few races of the day, a phenomenon that has become known as the 'late-race effect'. Early evidence of this pattern was uncovered by McGlothlin (1956), who found that, in the last race of the day, bettors underbet favourites to a greater extent than in any other race. He suggested that bettors might avoid bets on favourites in the last race because winning such bets would not recoup earlier losses (the track take ensured that most bettors would finish the day out of pocket). Rather, they preferred to bet on longshots, hoping for a lucky win in order to end the day in profit.

Over time, as more evidence of the late-race effect emerged, it was explained in terms of the risk-loving attitudes of the representative bettor. For example, Ali (1977), who found a greater degree of the FLB in the last race than the first two races of the day, posited that this demonstrated that bettors became more risk-loving as the day progressed. Similarly, Asch, Malkiel, and Quandt (1982) replicated McGlothlin's (1956) results, although, in their study, the extent of the bias was greater in the last two races of the day. Metzger (1985) also found evidence of the effect, but only if the first race of the day was excluded from the analysis. The late-race effect soon passed into betting lore, with Kopelman and Minkin (1991) describing how an avid racing enthusiast known as 'Gluck' espoused the rule: 'The best time to bet the favourite is in the last race'. Kopelman and Minkin's analysis confirmed that there was a sound economic basis for Gluck's rule.

More recent evidence has thrown the existence of the late-race effect into question. Johnson and Bruce (1993) found that in UK betting shops, bettors tended to place more bets on favourites in the last race, and suggested that this might be due to a 'break even' effect, whereby bettors seek to recover their losses by betting on outcomes that have at least a moderate chance of occurring. This hypothesis is supported by evidence that decision makers tend to exhibit loss aversion after a series of prior losses (Thaler and Johnson, 1990). Similarly, Brown, D'Amato, and Gertner (1994) observed a greater prevalence of the FLB in the last race of the day than in earlier races, but the difference was not statistically significant. Sobel and Raines (2003) found a slight increase on betting on longshots in the last two races of day. However, they also found that the general trend over the latter half of the evening was for bettors to begin to prefer favourites and shun longshots. Finally, Snowberg and Wolfers (2010) found no significant difference in the extent of the FLB in the last race of the day (in a dataset of over 850,000 races), suggesting that the late-race effect has now been eliminated.

From the contrasting evidence discussed above, it appears that bettors' increasing risk-love over a day's betting cannot fully explain the late-race effect. Johnson and Bruce (1993) consider that their converse result (a decreasing FLB in the last race) could be explained by a 'break-even' effect. However, a similarly plausible explanation is used by other authors to explain the opposite effect (an increasing FLB in the last race). Furthermore, it is not clear that expected utility theory is an adequate explanation. As Thaler and Ziemba (1988, p.171) ask, "why should a reduction in wealth increase the tendency for risk seeking?" Camerer (2001) points out that expected utility theory cannot explain why the same bettor leaves the racetrack one day, arrives again the next, and adopts a completely different risk attitude. Thaler and Ziemba propose that the effect can be explained by 'mental accounting', whereby bettors partition their wealth into separate accounts, and do not attempt to recoup losses in one account with funds in another. So, the late-race effect could be explained by bettors opening a mental account at the beginning of the day and closing it at the end, with an increasing desperation to break even as the day progresses (Camerer, 2001). Finally, the relative paucity of evidence for the effect in recent years could be attributed to a learning effect among bettors, as those who are aware of the effect are able to arbitrage it away should it reappear.

In summary, the FLB, while proving to be an interesting riddle for researchers, admirably demonstrates the value of naturalistic environments, betting markets in particular, in the study of decision making. While some potentially unrealistic simplifications (such as the representative bettor) must sometimes be made, the quality and quantity of betting market data has enabled the development of a large body of research on the nature of preferences and perceptions of risk under uncertainty.

1.4. Anchoring and herding

Betting market research has largely focused on the FLB, but some studies have investigated whether bettors make biased decisions in other ways. In particular, anchoring and herding represent biased behaviour whereby the decision maker alters their decision to account for external stimuli. Thus, when employing the anchoring and adjustment heuristic, decision makers unnecessarily alter their judgements to reflect an initially-provided estimate. Herding arises when decision makers neglect their own information and alter their judgements to reflect those of others. This section details the findings of these studies.

1.4.1. Anchoring and adjustment

Laboratory research suggests that, when making a numerical estimate, individuals, in an attempt to simplify the decision making process, tend to start from an initial value and make 'adjustments' upwards or downwards from it (e.g., Tversky and Kahneman, 1974). However, this often results in a bias whereby the decision is 'anchored' on the initial estimate, and adjustments are not sufficient. This is known as the anchoring and adjustment heuristic. For example, Tversky and Kahneman asked participants pairs of questions such as:

> (a) Is the percentage of African countries in the United Nations higher or lower than 25?
> (b) What do you think the exact percentage is?

They found that the figure given in (a) (i.e., 25 in the above example) significantly influenced the participants' responses to (b), even when the figure was randomly generated by spinning a wheel of fortune in the participants' presence. Higher/lower random numbers were associated with higher/lower estimates.

Anchoring has mainly been studied in controlled laboratory conditions. The few studies that have been conducted in naturalistic environments (e.g., amongst auditors: Bhattacharjee and Moreno, 2002, and real estate agents: Northcraft and Neale, 1987) have generally concluded that anchoring does seem to occur in information-rich real-world settings. However, these studies have used questionnaires or artificial problems. Consequently, the advantages of studying anchoring in betting markets are that participants are making estimates that matter to them, in a familiar real-world environment, without the use of questionnaires or artificial problems, and that they do not alter their normal behaviour (because they do not know they are being observed).

In the first study to investigate whether bettors anchor their judgements excessively, Liu and Johnson (2007) were primarily concerned with whether or not participants in horserace betting markets employed factors relating to previous performance of horses, jockeys, and trainers as anchors. For example, if a jockey had won his or her previous race, do bettors overestimate the chance that he or she will also win the current race? Previous finishing positions are not anchors in the *traditional* sense, since bettors are not specifically required to make direct comparisons between initial values and final judgements. Rather, this study attempted to find evidence of *basic* anchoring, where decision makers can be influenced by anchor values even when not asked to consider them directly (Wilson et al., 1996). Liu and Johnson investigated, using betting market data from Hong Kong, various explanatory variables that represent possible anchors (such as whether the horse won its previous race). However, the only significant explanatory variable was one that summarised a horse's finishing position over its career; this variable showed that bettors tend to *under*estimate horses that have a strong finishing record. Consequently, it appears that bettors tend to ignore some useful information relating to the horses' potential (or are unable to effectively employ such a complicated variable). However, the key finding was that no other explanatory variables were significant, indicating that bettors do not anchor their judgements on previous performances.

It is possible that Liu and Johnson's (2007) results fail to identify the anchoring that does occur in betting markets since any bias created by the anchoring of most bettors could be arbitraged away by the remainder of bettors. For instance, it is well known (e.g., Benter, 1994) that large betting syndicates, attracted by the unusually large betting volumes and strict regulation in Hong Kong (which helps to eliminate malpractice and insider trading) use sophisticated computer models to make considerable profits in this market.

Johnson, Schnytzer, and Liu (2009) extended the analysis of Liu and Johnson (2007) in two ways. First, noting that bettors in Hong Kong often spend considerable time reviewing race results, they expected that barrier position (the stall position from which the horse starts the race) would be a significant anchor for bettors. Second, decision makers with a higher level of expertise tend to be less susceptible to anchoring effects (e.g., Northcraft and Neale, 1987), so they expected that more experienced bettors would be less prone to anchoring. They found that bettors as a whole did not anchor excessively on barrier position over all their data but bettors overestimated the advantage offered by a good barrier position in one of the two racetracks under investigation. However, they found that expertise significantly reduced the extent of anchoring displayed by bettors (they used early and late betting as a proxy for inexpert and expert bettors, respectively). In summary, the two anchoring studies conducted in betting markets indicate that anchoring in real world environments may be a more complex phenomenon than has been found in laboratory studies, suggesting that further research may be required to fully understand its influence on decisions in real world environments.

1.4.2. Herding

Herding occurs when decision makers neglect their own information and adjust their actions to be more representative of the actions of others. Early theoretical models rationalized herding behaviour as information cascades, where decisions are made sequentially by different agents who each hold their own private information (e.g., Banerjee, 1992; Bikhchandani et al., 1992; Avery and Zemsky, 1998). The validity of the information is inherently uncertain and, as a result, individuals may be rational in disregarding some of their private information when the information held by other agents appears to conflict with their own. Hence, strictly speaking, herding behaviour in itself may not be 'biased' decision making. However, a biased outcome results from the combined effect of herding by multiple participants. In particular, this behaviour can lead to expected returns differing significantly from their 'rational' value.

Laboratory studies have generally found that participants display herd behaviour, but to a lesser extent than theoretical models predict, although the evidence has been inconclusive (Spiwoks et al., 2008). In betting markets, herding might be expected because there is a belief that certain bettors have access to privileged information. It has been found that betting on a horse or team that subsequently attracts a high degree of betting interest during the course of the market (known as a 'market mover' or 'plunger') is, on average, profitable (e.g., Crafts, 1985). The problem, of course, is that it is difficult to identify such opportunities before the fact, and this is where bettors with access to privileged information can gain an advantage. Bettors with superior information are often referred to as 'insiders' in the literature because of the presumption that their information is not in the public domain (e.g., a racehorse owner may have knowledge of secret training programs). However, there are also some bettors who use only publicly available information, but expertly combine all the information in such as a way as to form highly accurate opinions of the competitors' chances; these bettors are often referred to as 'informed' bettors. The presence of insiders and informed bettors in betting markets is widely reported (e.g., Crafts, 1985) and, consequently, herding behaviour may ensue when 'uninformed' bettors interpret a significant price movement as a signal that a competitor is being backed by insiders or informed bettors, and alter their bets accordingly.

The first study that investigated whether bettors herd is that of Camerer (1998). He tested whether bettors might respond to privileged information signals by placing large early bets in pari-mutuel pools at US racetracks and recording subsequent betting patterns. The purpose of this field test was to investigate whether markets could be manipulated. However, by observing the reactions of bettors to the temporary bets (Camerer subsequently cancelled the early bets), Camerer was also able to observe whether or not bettors displayed herding behaviour. He began by placing temporary bets of $500. He found that, while his bets temporarily distorted the odds, prices returned to their expected levels after cancelling his bets, indicating that bettors were not responding to the fake signals. Following this, Camerer increased his bet size to $1000 and targeted smaller racetracks and 'maiden' races (for horses that have never won a race). He detected a weakly significant herding effect whereby bettors were more likely to respond in the maiden races. However, overall the results still resolutely showed that bettors did not display herding behaviour. There remains an important caveat: although Camerer's bets made up of about 7% of the pool in the second study, they still may not have been large enough to induce herding. In a later study, Law and Peel (2002) argued that the apparent lack of herding in Camerer's (1998) study probably arose because, while the bets were sufficiently large to temporarily distort the markets, there was little incentive for bettors to herd on the initial price movement, since pari-mutuel bettors cannot lock in profits. To counter this, they conducted an empirical test for herding in UK bookmaker markets for horseracing. They argued that since the returns to a bet with a bookmaker are known at the time of bet placement, bettors might be more likely to herd in these markets. They noted that while an initial price movement could be due to informed trading, a further price movement may result from further informed trading or herding. Using the Shin (1993) measure of the degree of insider (or informed) trading, they were able to identify those large price movements that resulted from the trading of those with access to privileged information (the Shin measure increased over the duration of the market) or from herding (the Shin measure decreased). Law and Peel were particularly interested in those horses that opened at shorter odds than forecasted that then attracted significant betting interest. Significant positive returns of 10.2% could be made by betting on horses with these characteristics whose odds plunged as a result of informed trading; returns were significantly negative otherwise, at -10.9%. Consequently, Law and Peel were able to demonstrate that herding led to biased prices, with negative (positive) returns being reported when price movements were due to herding (informed betting).

Schnytzer and Snir (2008) noted that a horse that is not attracting bets that suddenly attracts a high degree of betting interest is likely to be overestimated due to herding. However, early plunges in odds suggest trading by bettors with privileged information, so they hypothesised that, due to the limited budgets of insiders, "a short time later, when the odds on those runners are lengthened again, those insiders are either unable or unwilling to place bets of sufficient significance to affect prices, even when the odds on those runners have drifted back to initial levels or even further" (p.3). This may arise because informed traders place most of their bets early in the market to secure profits. Schnytzer and Snir considered two possible situations: either odds increase early in the market and then decrease, or odds decrease early in the market and then increase. In the former, the late betting interest on the horse is considered to be evidence of herding, since the horse attracted little interest in the early market, and the final odds are expected to overestimate the horse's winning chances. In the latter, the early plunge followed by a lack of betting interest in the late market is considered evidence of cash-constrained informed betting, and the final odds are expected to underestimate the horse's chances. Investigating their hypothesis in bookmaker markets in the UK and Australia, their results demonstrated that, for horses attracting early but not late betting interest, positive returns of 15.3% were possible. On the other hand, only highly negative returns (as low as -27.2%) were possible for horses that lacked interest in the early market but were the subject of herding in the late market. These results confirmed that bettors herd and that this can lead to highly biased outcomes.

In summary, studies of anchoring and herding in betting markets have offered mixed conclusions. Camerer (1998) was unable to induce herding behaviour with his 'fake' signals but other studies have found evidence of significant herding by bettors when insider trading is prevalent. However, it appears that bettors do not anchor their judgements to the extent that has been reported in the laboratory. This may result from the fact that bettors are making decisions in an environment with which they are familiar (cf. naive subjects in unfamiliar laboratory settings) and in which they have learned (e.g., through repeated trial and improvement) to handle appropriately the redundant information and decision-relevant cues. Equally, while many bettors may herd to a significant extent, the actions of informed bettors, who arbitrage on the herding behaviour of others, may serve to suppress the observable effects of herding.

1.5. The gambler's fallacy and the hot hand fallacy

The gambler's fallacy and the hot hand fallacy both involve a misunderstanding of the nature of randomness. The application of these fallacies often results in systematically biased behaviour. The gambler's fallacy is defined as the belief that an event's probability of occurring is reduced after that event has occurred, even if the event is independent from one trial to the next (Rabin, 2002). Laplace (1825, p.92) gave the following examples from lotteries and coin tossing:

> when one number has not been drawn in the French lottery, the mob is eager to bet on it. They fancy that, because the number has not been drawn for a long time, it, rather than the others, ought to be drawn on the next draw. . . . It is, for example, very unlikely that in a game of *heads or tails* one will get *heads* ten times running. This unlikeliness, which surprises us even when the event has happened nine times, leads us to believe that *tails* will occur on the tenth toss.

The gambler's fallacy is the conviction that the coin, which is known, objectively, to be fair, is more likely to land heads than tails after the 'streak' of nine tails. This belief is demonstrated in laboratory experiments where participants are asked to invent a random sequence, such as repeated tosses of a coin. The results show that people tend to produce sequences containing too many alternations in the outcome relative to genuine randomness (Falk and Konold, 1997). The representativeness heuristic has been proposed as an explanation: the gambler believes that small samples must be representative of the population, so if unexpected sequences occur, a correction is expected (Tversky and Kahneman, 1971). As Tversky and Kahneman (1974, p.1125) note: "chance is commonly viewed as a self-correcting process in which a deviation in one direction induces a deviation in the opposite direction". Since nine tails in a row is an extremely unlikely event, the observer committing the gambler's fallacy expects that the next toss should be heads, in order to make the sequence of ten tosses seem less unusual. A commonly-cited example of this phenomenon is that of the Monte Carlo casino where, in a roulette game in 1913, black occurred 26 times in a row. During this streak, customers bet increasing amounts on red, and the casino profited as a result (Lehrer, 2009).

The hot hand fallacy involves mistaken convictions that run contrary to the gambler's fallacy. In particular, this fallacy involves a belief that if a player or team is on a winning (or losing) streak, this streak will continue longer than should be expected in a random sequence. So, in a game where the objective is to obtain tails on the toss of a coin, a gambler who has achieved the unlikely feat of landing tails nine times in a row believes that they are on a 'hot streak', and therefore expects that the coin has a greater probability of showing tails than heads on the next toss.

Gilovich, Vallone, and Tversky (1985) found that many basketball players and fans believed that a player would be more likely to score on a shot if they had scored (cf. missed) on their previous shot. However, they found no evidence to support this claim in either real games or controlled shooting experiments. The hot hand has been attributed to the illusion of control, which is the misplaced perception that gamblers have an element of control over random events (Langer, 1975). In fact, it has been shown that some gamblers believe that luck is separate from chance, and that their good fortune allows them to operate outside the laws of probability while they are on winning streaks (Wagenaar and Keren, 1988). Gilovich, Vallone, and Tversky (1985) suggested that, as with the gambler's fallacy, bettors may be employing the representativeness heuristic. In this case, long runs are deemed too unusual for the representative sequence, so bettors infer that the sequence generating process is no longer random (e.g., a basketball player who shoots an usually high run of on-target shots is said to be 'in the zone', or a roulette table or die is assumed to be biased). It is possible that, while a general belief in the hot hand may be misplaced, an accurate belief in the hot hand in specific instances motivates people to believe in its universality (see Bar-Eli et al., 2006, for many examples of genuine hot hand effects). The remainder of this section details the findings of studies that have investigated the two fallacies in naturalistic environments.

1.5.1. Evidence of the gambler's fallacy in betting markets

Clotfelter and Cook (1993) undertook one of the early studies using real betting data to investigate the gambler's fallacy. The US state of Maryland runs a 'daily numbers' draw lottery, where a 3-digit number between 000 and 999 is picked at random, and the bettor wins if they select this number. Clotfelter and Cook found that betting volumes on a number decreased in the days after the number was drawn, before returning to original levels after 84 days. It was postulated that bettors could be reducing their bets on numbers that had been drawn previously because they thought that that number was less likely to appear again. However, Clotfelter and Cook were unable to eliminate a 'wealth effect' from their data: bettors who regularly bet a particular number might stop betting altogether because they had achieved their financial goals. This could lead to a natural reduction in betting volumes on a winning number in the days and weeks after its appearance. A more significant caveat with Clotfelter and Cook's study was noted by Terrell (1994): the Maryland lottery has fixed payouts (winners are always paid $500 on a one-dollar bet), so choosing numbers based on the gambler's fallacy does not reduce the expected return to the bettor.

Croson and Sundali (2005) studied 18 hours of roulette play in a real casino, during which over a hundred players placed thousands of bets. They found evidence of the gambler's fallacy after streaks of around 5 or more similar outcomes (e.g., 5 red numbers in a row). However, Croson and Sundali (p.200) pointed out a similar concern to that existing in the Clotfelter and Cook (1993) study: "since the house advantage on (almost) all bets at the wheel is the same, there is no economic reason to bet one way or another (or for that matter, at all)."[5]

These studies highlight an important issue: while the gambler's fallacy is anecdotally known to be a common belief among gamblers, it does not always result in biased behaviour. For example, in roulette, the returns to bets on each outcome are independent of the bets placed by the customers. Therefore, the decision of which outcome to bet on is irrelevant. The gamblers in the Monte Carlo casino were not necessarily wrong to bet on red rather than black (although they might have bet more than they could afford). In such cases, it is plausible that belief in the fallacy only adds to the excitement of the game.

[5] Croson and Sundali also found evidence of the hot hand fallacy: 80% of bettors quit playing after losing a bet while only 20% quit after winning. Moreover, bettors tended to place more bets after winning than after losing.

In circumstances where acting on the fallacy results in a systematic bias that leads to a lower expected return for the bettor, it might be expected that the fallacy would be eliminated (e.g., by a learning process). However, there are a number of examples of the gambler's fallacy resulting in a systematic bias. These studies have necessarily needed to be creative in order to identify situations where one might expect evidence of the gambler's fallacy. For example, Metzger (1985) found evidence that horserace bettors tend to believe that streaks of favourites and longshots winning should cancel out. So, if a series of longshots wins, they bet more on favourites, and vice versa. Terrell and Farmer (1996) thought that bettors at greyhound racing events might believe that the starting positions of the winning dogs should be more random than it appears. Thus they might underestimate the winning chances of a dog starting in a given position from winning if the winner of the previous race also started from that position. Their calculations revealed that this was the case, with a positive return of $1.09 per dollar bet for a strategy of betting on dogs starting from the same position as the winner of the previous race. Terrell (1998) extended the study of Terrell and Farmer (1996) with a larger dataset, but only found significant evidence of the fallacy in one of the two years in their data.

Terrell (1994) conducted a similar investigation to Clotfelter and Cook (1993) but in a pari-mutuel New Jersey lottery, where payouts are shared between all the bettors who choose the winning number. Hence, if many gamblers avoid numbers that have recently appeared, the expected return to these gamblers is reduced. As expected, the extent of the gambler's fallacy was lower in this case. However, there was still a tendency to avoid numbers that had recently appeared. Terrell also found that if the results of Clotfelter and Cook were converted to a pari-mutuel system, there would be frequent occurrences when the payout would exceed $500, giving a positive expected return to bettors. This is not the case in New Jersey, so bettors appear to bet more evenly to avoid foregoing the increased potential winnings, and this diminishes the potential to exploit the fallacy. An alternative explanation for the results is that bettors simply prefer not to bet on a recently-seen number, in the same way that they prefer certain numbers (such as 777). Similarly, Papachristou (2004) found only marginal evidence of the gambler's fallacy in the pari-mutuel lottery in the UK.

1.5.2. Evidence of the hot hand fallacy in betting markets

As indicated above, the hot hand fallacy is also a mistaken perception of randomness. However, as with the gambler's fallacy, this mistaken belief does not necessarily impose economic penalties. Camerer (1989) examined the economic significance of the hot hand fallacy by investigating whether this mistaken belief is represented in gamblers' betting decisions. He categorized basketball teams based on their current winning or losing streak (in games), and then compared the actual results with the point-spreads offered by bookmakers[6]. If bettors believe in the hot hand, point-spreads will overestimate the chances of teams currently on winning streaks against the spread, while underestimating the chances of teams on losing streaks. The results showed that the performance of teams on winning streaks is worse than predicted by point-spreads, and teams on losing streaks perform better than predicted. However, the results were only marginally statistically significant.

Brown and Sauer (1993a, p.1377) highlighted the importance of the following critical assumption in Camerer's (1989) study: "the hot hand is belief in a myth". Camerer was effectively testing two alternatives: either bettors believe in a mythical hot hand, or they do not. However, there is evidence that genuine hot hand effects exist (Bar-Eli et al., 2006). Consequently, there is a third alternative: bettors believe in a genuine hot hand[7]. In this case, while bettors will move point-spreads to account for the hot hand effect, so teams' performance levels will also change. Brown and Sauer considered all three alternatives in basketball point-spread markets, but found only mixed results. They could not reject the hypothesis that the hot hand is real and that bettors

---

[6] The point-spread market is a betting market in which a bet wins if the home team wins by a specified margin of points (the point-spread) or, if the point-spread is negative, the home team loses by less than the point-spread (this is known as the team winning 'against the spread').

[7] There is a fourth alternative - that bettors are unaware of a genuine hot hand effect - but this hypothesis is not tested by Brown and Sauer.

correctly account for it, but they could also not reject the hypothesis that bettors believe in a mythical hot hand.

In a further study on the hot hand in point-spread markets for basketball, Oorlog (1995) found strong evidence against the hypothesis that gamblers believe in the hot hand. They devised a number of betting strategies to account for possible hot hand effects but none were profitable. Avery and Chevalier (1999) investigated US football betting markets, and also found a small bias as a result of the hot hand fallacy, but, again, the magnitude of the effect was small.

Further mixed evidence for the hot hand fallacy was provided by Durham, Hertzel, and Martin (2005). They found that point-spreads over-/under-estimated US college football teams on short winning/losing streaks against the spread, which is consistent with the hot hand fallacy. However, the point-spreads suggested that bettors expected longer winning or losing streaks to end rather than continue. Similarly, Paul and Weinbach (2005) found that betting against basketball teams on short winning streaks was profitable, while betting against teams on longer winning streaks was not. Moreover, they found no hot hand effect for teams on losing streaks, and suggested that this might be because bettors derived additional utility from betting on teams on winning streaks.

1.5.3. The paradox of the hot hand and gambler's fallacies

An important consideration is that the hot hand and gambler's fallacies appear at first to be opposite effects. While bettors may believe that long runs in the results of players or teams will continue (the hot hand), they simultaneously believe that long runs should end (the gambler's fallacy). This begs the question: how can we explain two apparently opposite effects?

One proposed explanation for both fallacies is the representativeness heuristic (Tversky and Kahneman, 1971), in which decision makers believe that sequences should be representative of the generating process. Decision makers apply the 'law of large numbers' too readily, i.e., they believe in the 'law of small numbers'. That is, while the relative frequencies of outcomes approximate the generating process in the long run, people believe that this should also be the case in the short run. So, the gambler's fallacy is explained because people believe that unusually long streaks are not representative, and so predict an alternation to make the sequence more representative. The hot hand is explained because people tend to over-infer from short sequences in a random process and decide that there is some underlying non-random process generating the sequence (Rabin, 2002).

It is potentially problematic to explain opposite phenomena with the same principle. However, a solution was provided by Ayton and Fischer (2004; see also Burns and Corpus, 2004). They tested whether the type of random process employed to generate the result was consequential in whether decision makers displayed the hot hand or the gambler's fallacy. They hypothesized that, when outcomes reflect human performance, people believe in the hot hand, whereas, when outcomes reflect inanimate mechanisms, people believe in the gambler's fallacy. This might explain why winning streaks of basketball and roulette players are perceived to exhibit long run tendencies, whereas outcomes of roulette games and lotteries are not. They conducted an experiment where they asked participants to play a simulated roulette-style game. Participants were required first to choose between red and blue, and were then asked to rate their confidence in their prediction. The results confirmed that while people are more likely to predict an alternation after a long run of either colour, they are also more confident in their own ability after a long run of successful predictions. Ayton and Fischer (p.1374) concluded that while the sequences of outcomes (red or blue) and predictions (win or lose) are each identical independent processes, "the two sequences are psychologically perceived quite differently; subjects *simultaneously* exhibited both . . . the hot hand and the gambler's fallacy". In a second experiment, they found that participants were more likely to attribute random sequences with low/high alternation rates to human performance/inanimate mechanisms. This line of experimentation goes some way to unravel the problematic nature of explaining two apparently opposite effects with the same heuristic.

In summary, there is evidence from a diversity of naturalistic betting environments that the decisions of bettors are consistent with the gambler's fallacy. However, the extent of the fallacy is reduced when it results in biased decisions, suggesting that bettors are sensitive to its economic significance. Research examining the hot hand fallacy in betting markets has been inconclusive. None of the above studies found irrefutable evidence that bettors believe in the hot hand and that market odds are biased in accordance with this belief. If there is a hot hand effect in markets, it is generally so small as to be economically insignificant.

1.6. Conclusion

The theme of this paper has been that, while many biases in decision making have been demonstrated in laboratory-based studies, there are many reasons for suggesting that these findings may not translate to the real world. Betting markets offer a valuable naturalistic setting in which to explore biased decision making, because participants are making decisions in a situation that is more representative of the environments in which day-to-day decisions are made. We have argued that bettors display significantly less biased judgements in their natural domain than those of naive participants in laboratory experiments. To support this view we have cited a number of examples related to rationality and calibration of subjective probability judgements. Furthermore, we have shown that there is only mixed evidence that bettors anchor their judgements, engage in herding behaviour, or believe in the hot hand or gambler's fallacies. Even the FLB, which has been the focus of the majority of research in betting markets, is no longer observable in some markets.

The primary conclusion of this paper is that, while systematic biases reported in the laboratory have been found in naturalistic betting markets, the extent and generality of these biases in these real world environments is often significantly less. The context of the decision task, the incentives offered, the lack of scrutiny involved, and the experience of the decision makers all contribute to an explanation for this conclusion. Another consideration is the importance of aggregation. It is costly and ethically challenging to obtain betting market datasets from which it is possible to discern individual biases. In a more typical dataset, individual biases may be eliminated by aggregation of the opinions of a diverse range of bettors. Moreover, even a systematic bias that is attributed to a large portion of the betting population can be reduced by the unbiased actions of a wealthy few, as there is always a strong economic motivation to capitalize on the biases of others.

A drawback to the heuristics and biases approach to decision making in general is highlighted by our discussion of the hot hand and gambler's fallacies. There is initially a problem with explaining two apparently opposite biases with the same heuristic, although subsequent research has clarified that there are two separate situations when people will use either of these fallacies. On the other hand, it can be impossible to narrow down multiple explanations for one bias to the single explanation that is most valid. Thus, there have been a wide range of explanations proposed for the FLB. Similarly, the hot hand fallacy could be explained by the illusion of control, or by the representativeness heuristic, or by extrapolation of genuine hot hand effects. As Wagenaar (1988, p.115-116) argues, the heuristics and biases approach

> does not specify rules telling us which heuristic will be applied in a given situation. Even worse, from the individual differences among gamblers, it is obvious that several heuristics could be chosen in one and the same situation, and that these heuristics lead to opposite behaviors. . . . There are so many heuristics, that it will be virtually impossible to find behaviours that cannot be accounted for.

Hence, while there is some evidence of biased behaviour in betting markets, explaining its prevalence is another matter altogether.

There are further issues associated with betting market research that may lead one to question the generalizability of the conclusions drawn from such studies. For example, bettors may be unrepresentative of the wider public since they are predominantly older males (Dipboye and Flanagan, 1979), and there may be some self-selection effects (indeed, it is not obvious as to why some people gamble and some do not; see Rachlin, 1990). We must also retain some scepticism about generalizability from betting markets to other economic settings (Levitt and List, 2007). Just as laboratory research should recognize that generalizability of findings is limited, future research into biased decision making in betting markets should acknowledge that laboratory experimentation is often the first available evidence that heuristics are being employed or biased outcomes are occurring. Without either the theoretical background or the controlled elegance of laboratory research, naturalistic research might be confounded by the vast array of potential variables involved and the often unintuitive nature of real-world decision making. The way forward appears to be a tandem approach with betting market studies being informed by results from laboratory experiments, and the latter being designed to examine the causes of phenomena that the former highlight. In this manner, the true nature and real world characteristics of behavioural biases may be revealed.

2. Noise, herding, and the efficiency of market prices: insights from markets for state contingent claims

Abstract

We develop new insights into unresolved issues related to the role of noise traders and the nature and effect of herding in financial markets by examining an electronic exchange market for state contingent claims. We find that noise trading is associated with increased market efficiency, and attribute this to informed traders being attracted by the improved liquidity that noise trading creates. We find evidence of differing 'buy' and 'sell' signal-induced herding in the later, more active stages of the market. We demonstrate that this results in an economically significant inefficiency; strategies designed to trade against the herd show substantial positive returns.

2.1. Introduction

An important concern in the financial markets microstructure literature has been the role of information and noise in market efficiency. Does noise trading result in excessively volatile, inefficient markets, in which added risks limit the possibility of arbitrage by informed investors? Or is noise essential in providing liquidity to informed investors in order that markets are efficient? A related issue is how to reconcile the apparently irrational behaviour of herding with the efficient markets hypothesis. If herding *is* rational, why do some of its worst consequences, asset bubbles and crashes, seem so irrational? Answering these questions has proved difficult using traditional financial market data. Consequently, we examine these issues by employing data from a market for state contingent claims, which offers considerable advantages for this research.

In Black's (1986) model, noise and information are contrasted, but each is essential for liquid markets. Noise and informed trading are complementary, since noise trading is regarded as trading that is *not* based on information. For Black, noise traders are entirely irrational. However, some studies have argued that noise trading can be rational because of its potential to make positive returns (Hong and Stein, 1999). It has been predicted (Black, 1986; De Long et al., 1990) and, to an extent, verified empirically (Campbell and Kyle, 1993) that a consequence of noise trading is increased volatility in market prices, a result that has important implications for policy making (e.g., Shleifer and Summers, 1990). However, the true role that noise plays in market efficiency remains the subject of much debate. On the one hand, it has been argued that noise may result in volatility in excess of the variations justified by underlying fundamental information (Shiller, 1981, 1990). In this case, noise is detrimental to market efficiency because of its destabilizing effect on long-run equilibrium values (De Long et al., 1990, Shleifer and Summers, 1990, Shleifer and Vishny, 1997). Noise introduces risks for informed traders, such as the risk that the market remains inefficient longer than the informed trader can remain liquid. Consequently, because informed investors are risk averse, they limit their arbitrage; thus, noise is seen as contributing to price inefficiency.

On the other hand, noise may be essential for generating liquid, and thus efficient, markets (Black, 1986). The seminal models of Grossman and Stiglitz (1980) and Kyle (1985) each predict that increased noise trading brings forth more informed trading. In Kyle's model, noise does not destabilize prices when informed traders are risk neutral. In Grossman and Stiglitz's model, the increased levels of noise and informed trading cancel each other out, so prices are stable. More recently, in the experimental market of Bloomfield, O'Hara, and Saar (2009), noise trading is shown to harm market efficiency, but only when prices are extreme. When prices are not extreme, noise traders help to make prices *more* efficient by providing liquidity for the informed traders.

It is difficult to resolve this issue in regular financial markets where uncertainty is always present, resulting in difficulties in fully measuring their efficiency. In Black's (1986, p.529) model, "noise is what makes our observations imperfect. It keeps us from knowing the expected return on a stock or portfolio … It keeps us from knowing what, if anything, we can do to make things better". As Shleifer and Summers (1990, p.22) note, identifying noise trading is tricky, "since price changes may reflect new market information which changes the equilibrium price". To overcome this problem, we examine a market that has a defined end point at which all uncertainty in the relation between prices and fundamental information is resolved. Specifically, we examine an electronic exchange market for state contingent claims, a horserace betting market. "In its simplest formulation, the market for bets in an *n*-horse race corresponds to a market for contingent claims with *n* states in which the *i*th state corresponds to the outcome in which the *i*th horse wins the race" (Shin, 1993, p.1142). These markets also offer the opportunity of quantifying the proportion of market activity attributable to informed trading (using a model developed by Shin, 1993). Consequently, we are able to examine variations in market efficiency with respect to the levels of noise trading (as the complement of the level of informed trading) in the market.

While noise is trading based on anything but information, a different, but related phenomenon is that of herding, which is trading based on the perceived information of other traders. Specifically, herding occurs when market participants neglect their own private information and adjust their actions to be more representative of those of other traders. They do this in the belief (perhaps mistaken) that other traders are more informed than themselves. The combined activity of many herding traders can result in extraordinary changes in asset values over a short period, possibly leading to bubbles, crashes and bank runs (Devenow and Welch, 1996). While the consequences of herding are irrational at the aggregate level, herding may be rational at the individual level. Theoretical models have rationalized herding as 'information cascades', where decisions are made sequentially by different agents who each hold their own private information (e.g., Banerjee, 1992; Bikhchandani et al., 1992; Avery and Zemsky, 1998). There is uncertainty over the validity of price signals and, consequently, it may be rational for agents to disregard some of their private information when that held by others appears (as revealed by their actions) to conflict with their own. In fact, in Hong and Stein's (1999) model, momentum traders can earn positive profits, provided they trade early enough in the 'momentum cycle'.

While herding behaviour has a theoretically sound basis, empirical evidence for the phenomenon in financial markets is inconclusive (Lakonishok et al., 1992; Wermers, 1999; Sias, 2004). Similarly, mixed results have been found in laboratory-based studies, in which both the decisions and the information on which they are based are observable (e.g., Cipriani and Guarino, 2005; Spiwoks et al., 2008). The common finding from these studies is that participants do herd, but to a lesser extent than theoretical models predict.

We address the unresolved issues raised above and provide important evidence which furthers understanding of the role of noise traders and the nature and effect of herding in financial markets. In particular, we first demonstrate that markets are both more volatile and more efficient when there is a greater degree of noise trading, supporting the hypothesis that noise trading improves market efficiency, perhaps by providing liquidity to informed traders. Second, we find that herding behaviour is prevalent in the market and leads to greater inefficiency than previous studies have suggested. We show that trading strategies designed to capitalize on mispricing caused by herding can earn significant abnormal returns, with initial capital rising by 95% on around 500 trades (a rate of return on turnover of over 10 percent). In addition, we are able to identify inefficiences in over 33 percent of the 1514 separate markets considered, indicating that the prevalence of herding is a significant issue. In addition, our results help resolve the conundrum of why previous herding evidence is so mixed, as we are able to measure the extent of herding at different stages of the market and, separately, for 'buy' and 'sell' signals. Specifically, we show that it is possible to make abnormal returns by trading on herding activity because market participants overestimate the information contained in large price movements in the later stages of the market, when there is little time for the inefficiency to be corrected. Furthermore, we show that the extent of herding is greater following 'sell' (cf. 'buy') signals.

This paper is organized as follows. In section 2.2, we describe the advantages betting markets offer for gaining insights into noise and herding in financial markets. In section 2.3, we develop the hypotheses, and in section 2.4, we outline the data and methods employed. We present the results in section 2.5, discuss them in section 2.6, and conclude in section 2.7.

2.2. Noise and herding in betting markets

Betting markets are valuable settings in which to explore behaviour in financial markets (Sauer, 1998). They share many characteristics with other financial markets, including the complexity and interdependence of factors which influence an asset's value, ease of entry, and a large number of participants who have access to a range of information (Vaughan Williams and Paton, 1997). In addition, as indicated above, betting markets (based on events such as political elections or horseraces) are markets for state contingent claims (Shin, 1993).

Furthermore, there is an important reason for believing that insights regarding noise trading and herding in financial markets may be more forthcoming when studying betting markets. In an efficient market, we should expect market prices to precisely reflect revealed fundamental information. However, in financial markets, prices are never entirely derived from the current fundamental information; rather, prices represent the current expectation of future prices. Hence, even if current fundamentals were fully known, there remains some uncertainty in prices. Betting market data enable us to overcome this concern. In particular, markets for an event (e.g., a race) close at a pre-defined end point. Bets are then settled, with all bettors receiving unambiguous payoffs. Consequently, in these markets there is a time when all uncertainty is resolved; the underlying fundamental information is revealed, in the sense that a winner is declared. This is repeated often, with several thousand separate markets per annum.

Exploration of noise trading and herding in betting markets also offers advantages over laboratory enquiry. In particular, betting involves uncertain and dynamic information, time stress, and rewards and penalties that matter to decision-makers: features that Orasanu and Connolly (1993) argue are only present in real-world decision contexts. Anderson and Brown (1984) confirm that risk-taking behaviour in high-stakes, real-world contexts is difficult to reproduce in laboratory settings. In addition, caution must be exercised when inferring from laboratory-based studies (which often involve non-experts making decisions in alien domains: Johnson and Bruce, 2001) the behaviour of experts such as those populating regular financial and betting markets. Consequently, betting markets appear to offer an ideal environment in which to develop insights into noise trading and herding.

Despite these advantages, relatively few studies have investigated these topics in betting markets. Brown and Sauer (1993b) found that, in a basketball betting market, the noise component in price variation was small relative to that associated with unobserved fundamentals, i.e., "noise is news" (p.1208). Those studies investigating herding in betting markets have generally shown that, as expected, betting on events (e.g., a horse winning a race) that subsequently attract a high degree of betting interest (known as a 'plunge') is profitable (Crafts, 1985; Schnytzer and Shilony, 1995). The problem, of course, is identifying such opportunities before the fact, and this is where bettors with access to privileged information can gain an advantage. Herding behaviour may ensue when a plunge is regarded by noise traders as a signal that the horse is being backed by those with privileged information. However, Camerer (1998) found that 'fake' privileged information signals (he placed large early bets in US pari-mutuel pools) failed to cause herding behaviour. This may have been because the fake bets were not large enough for other bettors to perceive them as genuine informed bets. Alternatively, bettors had little incentive to herd on the initial price movement, since, in pari-mutuel markets, payoffs are only known at the market close.

Law and Peel (2002) examined occasions when genuine privileged information resulted in significant price movements in UK bookmaker markets. They found that positive returns were obtainable by betting on horses that plunged as a result of informed trading, whereas returns were negative otherwise. Notably, it is the absence of herding which leads to the inefficiency, i.e., bettors fail to recognize genuine signals of privileged information. However, plunges resulting from trading by agents with privileged information were rare.

Schnytzer and Snir (2008) developed a model of cash-constrained informed traders in a bookmaker market. If mispricing becomes apparent early in the market, informed traders bet to take advantage, and herd betting by noise traders may ensue, causing a large price movement. However, there may be occasions when the price then returns to inefficient levels; at this point the informed traders have no cash remaining to exploit the inefficiency, which remains in the final prices. Consistent with the model, they find that positive returns can be made by betting on horses for which there has been a significant early plunge, but a later reversal in price. However, the set of such horses is again very small, so it is still unclear whether the results represent a genuine inefficiency. Their study is reminiscent of Hong and Stein's (1999) model, in which 'newswatchers' are cash-constrained and so underreact to their private information. This enables momentum traders to initially profit from the newswatchers' revealed information, but later a herding effect is created as momentum traders follow each other's trades rather than those of the newswatchers.

To develop important new insights, our study differs from these previous studies in a number of ways. First, we adopt a method (unlike that of Brown and Sauer, 1993b) that enables us to distinguish between noise and informed trading, allowing us to reveal the effect of noise trading on market efficiency. Second, we employ data from a betting exchange, where prices are derived entirely from the relative levels of supply and demand. This avoids the difficulty of interpreting lowering prices in bookmaker markets (employed in Law and Peel, 2002) as evidence of herding, since these price movements may result from bookmakers artificially lowering prices and may not be related to changes in bettors' demand. Finally, previous betting market studies have been conducted in markets (bookmaker or pari-mutuel), where 'assets' can only be 'bought'. We examine betting exchanges, which more faithfully represent wider financial markets. Importantly, in these markets participants can both buy and sell assets (i.e., 'back' or 'lay' a contestant to win or lose, respectively), allowing us to assess the relative importance of 'buy' and 'sell' trading signals.

2.3. Betting exchanges, Shin *z*, and hypotheses

2.3.1. Betting exchanges

We employ data from Betfair, the largest exchange betting market in the world by traded volume, with horserace betting revenues of £103 million in

2011 (Betfair, 2011). We consider 'win' markets, in which bettors must predict which horse will win (or, alternatively, which horses will lose). The odds for horse i in race j, are the best odds $\overline{R}_{ij}$ at which it is currently possible to back the horse to win. This represents the return to a £1 winning bet (e.g., a winning £1 bet with odds of 3.00 returns £3 for a profit of £2[8]). However, as is typical of exchanges, Betfair generally take a commission of 5% on net winnings[9]. Consequently, the effective odds, which we use in our analysis, are given by $R_{ij} = 1 + 0.95(\overline{R}_{ij} - 1)$. It is standard in the betting markets literature to make a distinction between 'odds' R_{ij} and 'price' $r_{ij} = 1/R_{ij}$, and we adopt this convention. The odds-implied probability q_{ij} of horse i winning race j, with n_j runners, is then given by

$$q_{ij} = \frac{r_{ij}}{\sum_{s=1}^{n_j} r_{sj}}. \tag{2.1}$$

While the horses' true winning probabilities are not knowable explicitly, each race j results in a vector of outcomes $(y_{1j}, y_{2j}, \ldots, y_{n_j j})^T$, where y_{hj} = 1 for the winning horse h and y_{ij} = 0 otherwise. If markets are efficient, then, over many races, odds-implied probabilities should approximate true winning probabilities as realized by race results.

2.3.2. The Shin measure of informed trading

Shin (1993) developed a means (known as Shin z) of measuring the proportion of market participation that can be attributed to traders with privileged information. His model describes a game based around a horserace, consisting of an expected profit-maximizing market maker (bookmaker) and a randomly selected bettor who is either perfectly informed (i.e., they know

---

[8] Exchange odds are expressed inclusive of unit stake and are often referred to as 'decimal' odds; this is in contrast to bookmaker markets where odds are expressed as, say, 2/1 for the equivalent of exchange decimal odds of 3.00.

[9] At the time the data used in this study was collected, the commission structure on Betfair was considerably more complicated than this, with a lower base commission rate applied to high volume bettors, and an additional charge applied to consistent winners. Thus our assumption of 5% commission on average is an estimate (and the true average commission rate is of little consequence to the results).

precisely the winner of the race) or a noise trader. The model predicts that, since the bookmaker is not perfectly informed, they will depress odds on longshots (the horses with the least chances of winning the race) relative to those on favourites in order to protect themselves from the possibility of large losses from an informed trader, who is in possession of superior information. Although in Shin's original model, informed traders are perfectly informed, Fingleton and Waldron (1999) relaxed this assumption, showing that it is equivalent to suppose that the precision of the informed trader's information can vary, and that the Shin z value is equal to the level of informed trading times the degree of precision. Hence, we can assume a more general situation in which a range of different types of informed traders operate, but that the level of influence they have in the market is likely to vary in tandem. The Shin z value itself is directly derived from final bookmaker prices and has been used extensively in betting market studies in order to investigate claims relating to the level of informed trading (e.g., Vaughan Williams and Paton, 1997; Smith et al., 2006). An explanation of the method used to derive Shin z is given in Appendix 1.

If the proportion of traders holding priviliged information is low in a market, then the proportion of traders whose information is shared by other market participants (in the case of Shin's model, shared with the bookmaker) is high, and vice versa. If this shared information is already incorporated into market prices (through odds-setting by the bookmaker), any further trading by participants with shared information is uninformed and is hence noise trading. Thus, we use the complement of Shin z to measure the degree of noise trading in the markets. As in Smith, Paton, and Vaughan Williams (2009), we note that Shin's model predated the advent of betting exchanges, and so the assumption that prices are set by a monopoly bookmaker is no longer valid. However Shin's model can be adapted for exchange markets with a few reasonable assumptions. For example, instead of a monopoly bookmaker, we can assume that there is an oligopoly of 'big players' who act as market makers. The motivation behind this idea is that it is well known that there are well-informed traders controlling a vast share of the wealth that is traded in betting exchange markets. These traders are subject to lower commission rates due to their large historical traded volumes, and this enables them to maintain their dominant status. With these appropriate modifications to Shin's model, Shin z can again be employed to assess the degree of noise trading in the market.

2.3.3. Hypotheses

We first test the predictions of Black (1986) and De Long, Shleifer, Summers, and Waldmann (1990) that, in the absence of new information, noise trading increases short term volatility. Previous tests of these predictions (e.g., Campbell and Kyle, 1993) are rare and those that have been conducted are only marginally conclusive. Second, as indicated above, there is some debate as to whether noise trading has a net positive or negative effect on efficiency. On the one hand, noise trading moves prices away from efficient levels, and informed traders, who can restore efficiency, may fail to do so because they are risk averse (Shleifer and Summers, 1990; Shleifer and Vishny, 1997). However, when prices are not extreme, noise trading can provide liquidity to informed traders, enabling them to arbitrage away inefficiency (Bloomfield et al., 2009). Hence, our first hypothesis is:

1. Increased noise is associated with increased market price (*i*) *volatility,* (*ii*) *efficiency*.

In betting exchanges, the demand for bets on a particular outcome is directly represented in the current market price. Thus, if a large unidirectional price movement results in a price differing from that expected (given the complete set of fundamental information), then that price movement is evidence of herding.

It is well established that bets placed in the later stages of a betting market are more informative than early bets (Asch et al., 1982; Gandar et al., 2001)[10]. One explanation is that the timing of bets is variably incentivized depending on the quality of the bettors' information (Ottaviani and Sørensen, 2005). More informed bettors have an incentive to bet late to avoid revealing their private information to other bettors. In addition, liquidity is lower in the early stages of

---

[10] While not reported here, we have verified that this is also the case in our exchange market data.

exchange betting markets, so bettors incur additional transactional costs in the form of wider bid-ask spreads, and are unable to place large enough bets to compensate them for revealing their information. Consequently, herding on price movements that are believed to be signals of informed trading are likely to be more commonplace in the later stages of the market (when there may be insufficient time for informed traders to arbitrage away the resulting inefficiency). Moreover, on-course bookmaker prices are posted online just 10 minutes before the start of each race, at which point there is usually a considerable adjustment in off-course and online prices due to the information contained in the on-course prices (Schnytzer and Snir, 1995). Thus, we should expect herding to be more prevalent in this final window of betting before the race starts. These considerations motivate our second hypothesis:

2. Bettors display herding behaviour, but to a greater extent in the later stages of the market than in the early stages.

In betting exchanges, as in other financial markets, 'buy' or 'sell' signals may provide different information signals. Consequently, we investigate whether bettors' herding behaviour differs depending on the *direction* of large price movements. This investigation is motivated by the fact that betting exchanges facilitate the laying of 'known losers': horses which are deliberately prevented from running to their potential (Marginson, 2010). This practice could benefit horse owners who know that their horse will lose. Despite rules which forbid such behaviour, its prevalence is the subject of much debate, suggesting that bettors might be more likely to interpret 'sell' signals as genuine informed trading. Consequently our third hypothesis is:

3. Bettors herd to a greater extent on 'sell' (lay) signals than 'buy' (back) signals.

Herding leads to inefficiency if the deviation in market prices from fundamental information is sufficiently large such that an arbitrage opportunity arises. We believe this is likely to be the case, so this motivates our fourth hypothesis:

4. Herding presents an inefficiency, such that it is possible to make positive returns by betting against those who herd.

2.4. Data and methods

2.4.1. Data

The data employed are sequences of odds for 62,124 horses running in 6058 races in the UK and Ireland from August 2009 through August 2010. The data were downloaded in real-time using the Betfair API and consist of the odds on the exchange for each horse in each race (as indicated above, the odds are the best price at which it is possible to back the horse to win[11]). The data were collected at 1-minute intervals throughout the duration of the market, from 9:00 a.m. on the morning of the race, through to race start time (resulting in over 8.4 million data points). We segment the market on each race into four time periods, depending on the amount of time left before the race start, in order to determine the prevalence and direction of herding over different periods of the market. While markets are often activated on the evening before the race, or earlier for the most popular events, the vast majority of participation in markets takes place on the day of the race, so segment 1 begins at 9:00 am and ends at the race start time. The most active stage of the market then begins 30 minutes before the race start, since this is the typical length of time between races at each racetrack, so is the period when most participants direct their attention to the race. We divide this period of time up into three segments: segments 2 and 3 end at the race start time and begin 30 minutes before the race and 15 minutes before the race, respectively, and segment 4 begins 30 minutes before the race and ends 15 minutes before the race. Hence segment 1 lasts at least 4 hours, depending on the race start time, segment 2 lasts 30 minutes, and segments 3 and 4 last 15 minutes.

For our analysis of noise trading, we split the dataset into 'high noise' and 'low noise' subsets, consisting, respectively, of those races with below-/above-median Shin z (with 3029 races in each set). We compare the levels of informed trading and the efficiency of market prices between the two sets. For our

[11] Instead of 'back' prices we could have used 'lay' prices or the mid-point of 'back' and 'lay' prices. Similarly, we make a minor assumption that odds are equally valid as prices whatever the stake limit. Neither of these considerations have more than a very minor effect on our analysis.

analysis of price movements, we split the full dataset into a training set of the first 75% of races (4544: 47,196 horses), and a holdout set of 25% of races (1514: 14,928 horses). We estimate conditional logit models (McFadden, 1974) on the training set, in order to determine whether bettors exhibit herding behaviour. We use these models to predict horses' winning probabilities in the holdout set and construct betting strategies based on these probabilities to test the market efficiency implications of any observed herding. Hence, our conclusions about market efficiency, which are based only on the holdout set, can be relied upon, because they are out-of-sample and thus are not influenced by fitting our models on the training set.

2.4.2. Measures of trend and volatility of odds

We generate an *odds curve* for each horse in each segment of the market for a given race, using the method of Johnson, Jones, and Tang (2006). That is, for each horse *i* in race *j*, and for each market segment *k*, we have a sequence S_{ijk} of L_{ijk} pairs of times $t_{ijk}(l)$ and odds $R_{ijk}(l)$, i.e., $S_{ijk} = \{[t_{ijk}(1), R_{ijk}(1)], \ldots, [t_{ijk}(L_{ijk}), R_{ijk}(L_{ijk})]\}$. We record price changes so that, for each time in the sequence, the odds are different from the preceding time. Consequently, for any time *T*, where $t(l) \le T \le t(l+1)$, $R = R(l)$ (here, and in the following, we drop the subscripts *i*, *j* and *k* when their use is not required). The first/last pair is the first/last time in the segment along with the first/final odds recorded. The final odds recorded in segment 1 (the full duration of the market) are the odds at which the horse started the race (or 'starting price'); this special case is used to calculate the final odds-implied probability, which is given by $q_{ij}(L_{ij}) = [1/R_{ij}(L_{ij})] / \sum_{s=1}^{n_j} [1/R_{sj}(L_{sj})]$. Finally, we rescale all the sequences so that $t(1) = 0$, $t(L) = 1$, and $R(L) = 1$. The result of this procedure is that each odds curve is a piecewise continuous step function $\Phi(t)$ on the interval [0, 1], such that $\Phi(1) = 1$. From the odds curve, we measure underlying trends in the odds. Specifically, the trend μ is estimated as the slope of the ordinary least squares regression line fitted to the pairs in *S*, constrained to pass through (1, 1), i.e., (2.2) $$Y(t) = 1 + (t-1)\mu,$$

and is therefore given by

(2.3) $$\mu = \frac{\sum_{l=1}^{L}[R(l)-1][t(l)-1]}{\sum_{l=1}^{L}[t(l)-1]^2}.$$

A trend variable is estimated for each horse in each race for each of the four segments. Further, because bettors might infer differing information from 'lay' and 'back' bets, increasing or decreasing prices may be interpreted differently. Consequently, we derive two trend variables, $\mu^+ = \max(\mu, 0)$ and $\mu^- = |\min(\mu, 0)|$, for each horse in each segment (i.e., eight trend variables for each horse in each race). Hence, for horse *i* in race *j*, and for market segment *k*, $\mu_{ijk}{}^+ = \max(\mu_{ijk}, 0)$ and $\mu_{ijk}{}^- = |\min(\mu_{ijk}, 0)|$.

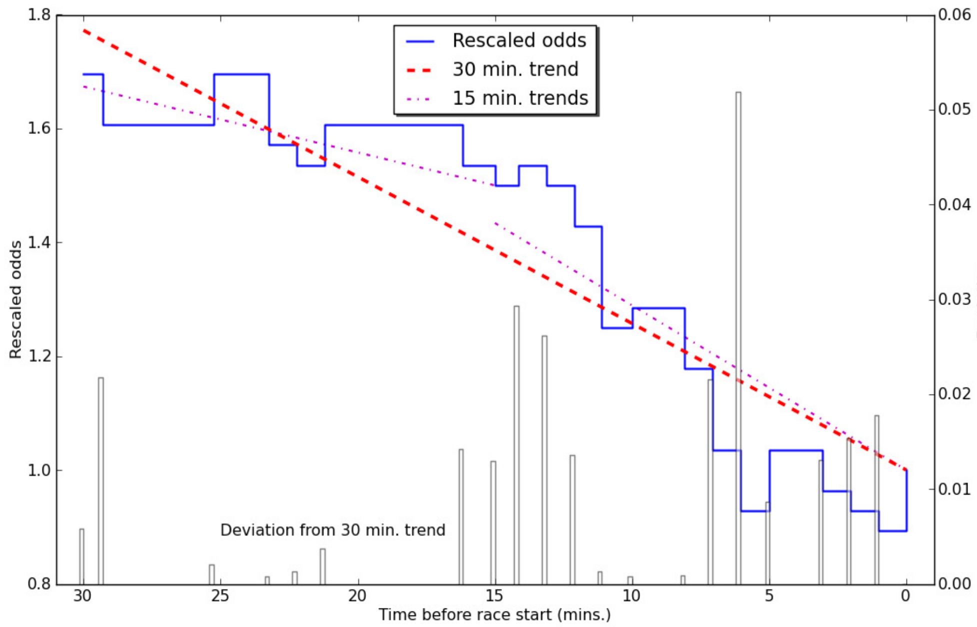
Since racetrack betting markets are based on a significant amount of information that is revealed in real time (e.g., the condition of the horses, the weather), it is reasonable to expect, *ex ante*, that prices will fluctuate around an underlying trend. The trend represents the bettors' collective opinion of the horse's chances at the close of the market, relative to their chances at the opening of the market. So, in order to obtain a meaningful measure of volatility, we calculate, for each horse, in each race, the trend μ_{ij} in the odds curve, and then the volatility σ_{ij} is given by the variance of the odds around the regression line in (2.2)[12], i.e.,

(2.4) $$\sigma_{ij} = \frac{1}{L-1}\sum_{l=1}^{L-1}\{R_{ij}(l) - Y_{ij}[t_{ij}(l)]\}^2 = \frac{1}{L-1}\sum_{l=1}^{L-1}\{R_{ij}(l) - 1 - [t_{ij}(l)-1]\mu_{ij}\}^2 .$$

To illustrate, Figure 2.1 shows an example of an odds curve with the 30 minute trend, along with the deviation for each $l = 1, 2, \ldots, L-1$. The volatility measure discussed above relates to each horse in turn. However, we are primarily interested in volatility on a race-by-race basis. Consequently, we use the mean of σ_{ij} over all the horses in each race, $\overline{\sigma}_j = (1/n_j)\sum_{i=1}^{n_j}\sigma_{ij}$.

[12] A range of other measures of volatility were tested, but the results were the same.

Figure 2.1. Example of the least squares regression method for determining the trend and volatility of the odds curve for a horse.



2.4.3. The conditional logit model and herding

The conditional logit (CL) model (McFadden, 1974) has been employed in many betting market studies (Asch et al., 1984; Bolton and Chapman, 1986; Benter, 1994; Sung and Johnson, 2010). It allows us to estimate the winning probability of each horse, taking into account competition between horses in the race. Formulation of the CL model begins with an estimate of horse *i*'s ability to win race *j*,

(2.5) $$W_{ij} = \sum_{l=1}^{m} \beta(l) x_{ij}(l) + \varepsilon_{ij},$$

where $\beta(l)$, for $l = 1, \ldots, m$, are the coefficients that determine the importance of the variables $x_{ij}(l)$, and ε_{ij} is an independent error term. If the independent errors are identically distributed according to the double exponential distribution, the estimated winning probability for horse *h*, p_{hj}, is given by

(2.6) $$p_{hj} = \Pr(W_{hj} > W_{ij}, i = 1,2,\ldots,n_j, i \neq h) = \frac{\exp[\sum_{l=1}^{m} \beta(l) x_{hj}(l)]}{\sum_{i=1}^{n_j} \exp[\sum_{l=1}^{m} \beta(l) x_{ij}(l)]}.$$

The coefficients $\beta(l)$ are estimated by maximizing the joint probability of observing the results of all the races in the dataset; this is achieved by maximizing the log-likelihood (LL) of the full model (i.e., one including all independent variables in which we are interested):

(2.7) $$\ln L(full) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij} ,$$

where $y_{ij} = 1$ if horse i won race j and $y_{ij} = 0$ otherwise, and N is the total number of races in the dataset. For this study an appropriate measure of the predictive accuracy of the model is Maddala's (1983)[13] pseudo-R^2, given by

(2.8) $$R^2 = 1 - \exp\{(2/N)[\ln L(naive) - \ln L(full)]\} ,$$

where $\ln L(naive)$ is the LL of the naive model (where each horse in a race is assigned the same probability of winning):

(2.9) $$\ln L(naive) = \sum_{j=1}^{N} \ln(1/n_j) .$$

The standard normal test statistic $z(l) = \beta(l)/S.E.[\beta(l)]$ is used to test if variable coefficients are significantly different from 0, i.e., variables add predictive power to the model. An additional test to justify augmenting simpler models with additional variables utilizes the likelihood ratio (LR) test statistic $2[\ln L(full) - \ln L(naive)]$, which is χ^2 distributed with degrees of freedom equal to the number of additional variables.

In our analysis, the first variable in the CL models will always be log of final odds-implied probability, i.e., $x_{ij}(1) = \ln[q_{ij}(L_{ij})]$. If the estimated value of the coefficient of this variable, $\beta(1)$, is equal to one when there are no other variables in the model, this implies that there is no favourite-longshot bias (FLB), where FLB is the widely-reported phenomenon whereby favourites/longshots are under-/over-bet (e.g., Dowie, 1976). The greater the value of $\beta(1)$, the greater is the degree of the FLB (Bacon-Shone et al., 1992). However, previous studies have indicated that betting exchanges display little, if any, FLB, (Smith et al., 2006), suggesting that $\beta(1) = 1$. Whatever its value,

---

[13] We use Maddala's pseudo-R^2 rather than McFadden's (1974) more popular definition, because McFadden's R^2 has the unfortunate property of varying with the average number of horses in each race, which is not the case for Maddala's pseudo-R^2. Our results would be the same using the McFadden pseudo-R^2, but if we used this measure we could not account for variations in market efficiency due to the differing numbers of runners in the high and low noise sets.

having developed a model incorporating an appropriate value of $\beta(1)$ (i.e., having adjusted for any FLB), the pseudo-R^2 of a single-variable CL model is an appropriate measure of the predictive accuracy of market prices, and thus the market efficiency.

To compare the effects of increased noise levels on efficiency, we estimate single-variable CL models on the high and low noise subsets, and compare the models' pseudo-R^2s. We estimate the distributional properties of the pseudo-R^2s using a bootstrap method (Efron, 1979). For each of the high and low noise sets (each of 3029 races), we repeat 1000 times a random sampling of 3029 races, with replacement, and fit a single-variable CL model to each sample. The random sampling is stratified so that the proportions of handicap races (where horses are allocated weights to carry based on their previous performances) are approximately equal in the samples and the full dataset. This controls for potential effects on the accuracy of odds-implied probabilities from the greater complexity involved in handicaps (Johnson and Bruce, 1998). The sample means, $\mu(R_H^2)$ and $\mu(R_L^2)$, and variances, $s^2(R_H^2)$ and $s^2(R_L^2)$, of the resulting sets of pseudo-R^2s are used to derive a standard normal test statistic,

(2.10) $$z[\mu(R^2)] = \frac{\mu(R_H^2) - \mu(R_L^2)}{\sqrt{s^2(R_H^2) + s^2(R_L^2)}}.$$

This is used to test part (ii) of our first hypothesis, that high market noise is associated with increased market efficiency (similar test statistics are derived for the levels of FLB and volatility).

Herding occurs when bettors alter their actions to be more representative of the actions of others. It may be rational for individual bettors to rely on the private information of bettors' that they believe are more informed than themselves. However, a multitude of simultaneously herding bettors can lead to significant movements in prices that cannot possibly be fully accounted for by the underlying objective information of a handful of informed bettors. Hence, when a horse's odds decrease/increase significantly (i.e., resulting from bets on that horse to win/lose), and that price movement is not fully attributable to a genuine increase/decrease in that horse's chances of winning, then the horse's final odds will imply a probability that is greater/lower than the horse's true winning probability, i.e., the odds will be too low/high. Therefore, the second variables we employ in the CL models are the two trend variables that we derived previously for each market segment. Higher values for the trend variables imply steeper price changes. If large price movements correspond to herding, the coefficients of the second variables should be significantly different to zero (and the corresponding LL ratio tests should also be significant): when odds increase/decrease, a significantly positive/negative variable implies herding, i.e., horses whose odds-implied probabilities decrease/increase over time win more/less often than implied by the odds. Determination of the extent to which bettors herd over the different time periods when odds increase or decrease allows us to test our second and third hypotheses.

2.4.4. Betting strategies

If bettors herd to the extent that final odds-implied probabilities are not in line with true winning probabilities, it should be possible to find profitable betting opportunities. To investigate this possibility, we estimate a CL model involving a combination of variables relating to herding behaviour, estimated on the training set of 4544 races. We use this model to predict winning probabilities for horses running in the 1514 holdout races. If herding results in inefficiency, betting strategies based on the model should be profitable in the long term and involve relatively low risk. Considering each holdout race *j* in turn, with initial wealth £1000 and current wealth W_j, we use the estimated probabilities as the basis for the following betting strategies:

1. *Level staking*: For each horse *i*, if $p_i > r_i$, bet 1% of current wealth on horse *i*. Therefore, if a bet is to be placed, the size of the bet is £$W_j/100$.
2. *Proportional staking*: For each horse *i*, if $p_i > r_i$, bet an amount such that the profit from a win, after commission, is 10% of W_j, i.e., bet size is £$W_j/10(R_i - 1)$. An advantage of this strategy in assessing inefficiency is that returns are not unduly influenced by 'lucky' wins on horses with very high odds (Schnytzer and Snir, 2008).

3. *Kelly staking*: The Kelly strategy (Kelly, 1956) assigns bet sizes x_i over all n horses in the race to maximize the log of expected wealth after the race, $G(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} p_i \ln F_i$, where $F_i = 1 + 0.95(x_i \overline{R}_i - \sum_{s=1}^{n} x_s)$ if $x_i \overline{R}_i > \sum_{s=1}^{n} x_s$, and $F_i = 1 + x_i \overline{R}_i - \sum_{s=1}^{n} x_s$ otherwise (since 5% commission is only paid if bets result in an overall profit). The x_i are estimated using numerical optimization. The Kelly strategy is optimal in the sense that it maximizes the asymptotic rate of growth of wealth and minimizes the expected time to reach a pre-defined wealth target (Breiman, 1961). However, since recommended bets may be very large, the volatility of returns from a full Kelly strategy over the 1514 holdout races may not result in a positive overall return.
4. *Half Kelly staking*: Some authors (e.g., Benter, 1994) recommend a fractional Kelly strategy, whereby bet sizes are a fixed proportion (in this case, a half) of those recommended by the full Kelly strategy. This is sub-optimal in that it no longer maximizes the asymptotic growth rate of wealth. However, fractional Kelly strategies are less risky, and, over medium-length time horizons, may result in a higher expected return as a percentage of the total amount bet (MacLean et al., 2010).

The above strategies all entail a zero probability of ruin, assuming that arbitrarily small bets can be placed. A non-intuitive property of the Kelly strategy is that it may recommend bets on outcomes with negative expected returns. However, this is only optimal if our estimates for the true winning probabilities, p_i, are accurate. If this is not the case, over-betting will occur (MacLean et al., 1992). To allow for inaccuracies in our estimates, we adapt the Kelly strategies so that no bets are placed on horses for which the expected return is negative, i.e., $p_iR_i < 1$ (Hausch et al., 1981). Second, the strategies might recommend large bets on horses with a high probability of winning, so a single unfortunate loss may skew the overall returns. Similarly, skewed returns may result from a fortunate win on a horse with a low winning probability. We therefore restrict single bet sizes to a maximum of 10% of current wealth. We assess the performance of the betting strategies using the following measures:

1. *Rate of return*: the ratio of the profit (or loss) earned to the total amount bet.
2. *Risk-adjusted return*: the risk-adjusted return, given by $R/[Var(R)]^{1/2}$, where R is the rate of return and the variance is estimated using a bootstrap procedure, by sampling with replacement from the holdout set 1000 times and calculating returns on each sample.
3. *Expected final wealth*: $W_0\prod_{j=1}^{N} X_j$, where N is the number of races bet on and, for race j, X is the expected increase in wealth factor $\sum_{i=1}^{n} p_i \ln \widetilde{F}_i$. Here, $\widetilde{F}_i = 1 + 0.95(\widetilde{x}_i\overline{R}_i - \sum_{s=1}^{n}\widetilde{x}_s)$ if $\widetilde{x}_i\overline{R}_i > \sum_{s=1}^{n}\widetilde{x}_s$, and $\widetilde{F}_i = 1 + \widetilde{x}_i\overline{R}_i - \sum_{s=1}^{n}\widetilde{x}_s$ otherwise, where $\widetilde{x}_i$ is the fraction of current wealth bet on horse i after any restrictions are imposed (MacLean et al., 1992).
4. *Probability that final wealth is above x% of initial wealth*: this is given by $1-\Phi\left\{\dfrac{(1/N)\ln(x/100) - E[\ln X_j]}{\sigma[\ln X_j]}\right\}$, where Φ is the standard normal cumulative distribution function (MacLean et al., 1992).

2.5. Results

2.5.1. Noise, volatility, and efficiency

The high and low noise sets consist of the races in the dataset that, respectively, have Shin z less than (mean Shin z = 0.0089) and higher than (mean Shin z = 0.0129) the median (0.0105). Further characteristics of the data in the high and low noise sets, as well as their 1000-bootstrapped samples, are summarized in Table 2.1.

CL models, with log of final odds-implied probability as the single predictor variable, are estimated for the high and low noise datasets, as well as for each of the bootstrap samples. The coefficient of the single variable, $\beta(1)$, is not significantly different from 1 in any case (z = 0.25, p = 0.8037 and z = 0.94, p

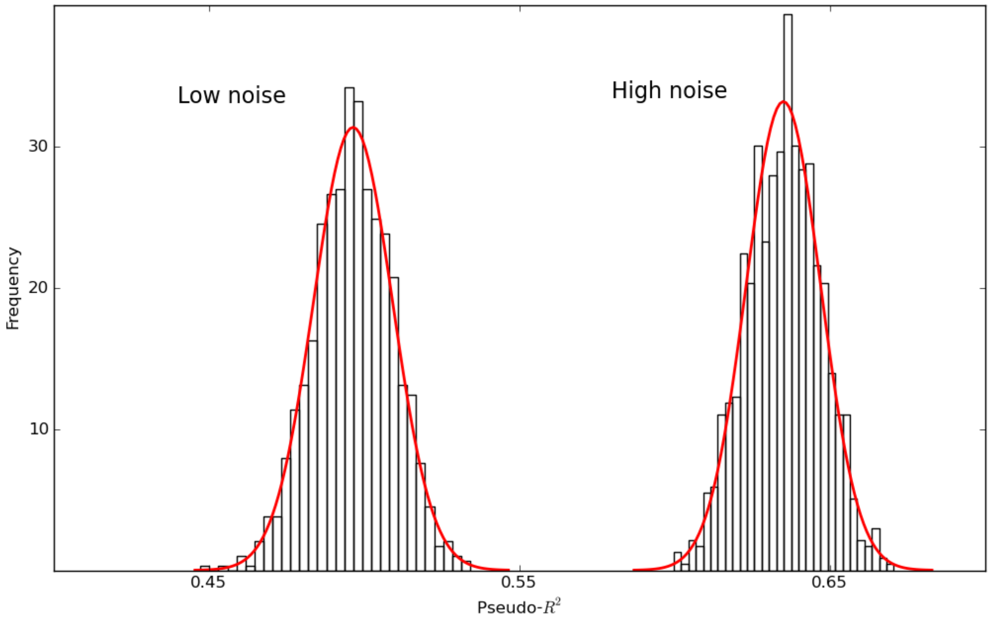
= 0.3453 for the original high and low noise sets, respectively; z = 0.20, p = 0.8421 and z = 0.99, p = 0.3236 for the high and low noise bootstrapped datasets), indicating, as expected, the absence of FLB. Furthermore, we find that mean pseudo-R^2 values for the high and low noise bootstrapped sets (see Figure 2.2) confirm that prices are on average both more volatile (high: mean volatility 0.0286; low: mean volatility 0.0200; z = 11.84, p = 0.0000) and more accurate in predicting winners when market prices are noisier (high: mean pseudo-R^2 = 0.6351; low: mean pseudo-R^2 = 0.4965; z = 7.91, p = 0.0000). Consequently, the results support both parts of our first hypothesis, that increased noise is associated with greater market volatility and efficiency.

Table 2.1. A description of the data: races with high and low noise.

	All data	High noise (low Shin z) set	Low noise (high Shin z) set		Bootstrapped high noise set	Bootstrapped low noise set
Number of races	6058	3029	3029	Number of races	3029	3029
Number of horses	62,124	39,555	22,569	Mean number of horses	39,511	22,598
Mean number of horses per race	10.3	13.1	7.5	Mean number of horses per race	13.0	7.5
Number of handicaps	3255	1747	1508	Mean number of handicaps	1628.2	1628.5
Proportion of handicaps	0.537	0.577	0.498	Mean proportion of handicaps	0.538	0.538
Mean Shin z	0.0109	0.0089	0.0129	Mean Shin z	0.0089	0.0128
Level of FLB $\beta(1)$	1.014	1.001	1.026	Mean level of FLB $\beta(1)$	1.005	1.027
				z(β)	-0.64 (0.035)	
[β(1)-1] / $S.E.$[β(1)]	0.78 (0.017)	0.25 (0.022)	0.94 (0.028)	[β(1)-1] / $S.E.$[β(1)]	0.20 (0.023)	0.99 (0.027)
Mean volatility $\bar{\sigma}$	0.0242	0.0280	0.0204	Mean volatility $\bar{\sigma}$	0.0286	0.0200
				$z(\bar{\sigma})$	11.84** (0.001)	
ln$L$(naive)	-13,670.0	-7702.9	-5967.3	Mean ln$L$(naive)	-7699.7	-5974.3
ln$L$(full)	-11,125.0	-6224.6	-4900.1	Mean ln$L$(full)	-6172.0	-4934.8
Pseudo-R^2	0.5684	0.6232	0.5057	$\mu(R^2)$	0.6351	0.4965
				$z[\mu(R^2)]$	7.91** (0.018)	

** denotes significance at the 1% level in a 2-tailed test.

Figure 2.2. Histograms of pseudo-R^2 from 1000-bootstrap samples taken from high noise and low noise sets (fitted normal curves shown).



2.5.2. Herding

The results of estimating models (using the training set of 4544 races) including indicators of possible herding behaviour in the four time segments of the market (separated by whether prices increase or decrease) are presented in Table 2.2. The coefficient of log of final odds-implied probability in the CL model where this is the only variable (Model 0), is significantly different from zero (z = 51.45, p = 0.0000). In addition, the model's LL is -8337.6, confirming that the odds, as expected, add significant predictive power over the naive model (LL = -10,307.7). Models 1 to 8 include a second variable that describes a trend in prices over time. In Models 1 and 2, which assess the predictability of the trend over the full duration of the market, the coefficient of the second variable is not significantly different from zero (z = 0.92, p = 0.3576 and z = -0.13, p = 0.8966, respectively). These results suggest that large price movements over the full duration of the market do not necessarily result in odds-implied probabilities differing from true winning probabilities, i.e. herding is not apparent when considering the full duration of the market.

Table 2.2. Conditional logit results using indicators of herding behaviour.

Model	Variable	Coefficient $\beta(l)$	$z(l) = \beta(l)/$ $S.E.$ [$\beta(l)$] ($S.E.$)	ln$L$	LR test vs. Model 0	Pseudo-R^2
Naive	-	-	-	-10,307.7	-	-
0	$\ln q_{ij}(L_{ij})$	1.015	51.45** (0.020)	-8337.6	-	0.5798
1	$\ln q_{ij}(L_{ij})$	1.025	45.68** (0.022)	-8337.2	0.82	0.5799
	$\mu_{ij1}{}^{+}$	0.058	0.92 (0.063)			
2	$\ln q_{ij}(L_{ij})$	1.016	50.28** (0.020)	-8337.6	0.02	0.5798
	$\mu_{ij1}{}^{-}$	-0.004	-0.13 (0.028)			
3	$\ln q_{ij}(L_{ij})$	1.033	48.48** (0.021)	-8335.1	4.92*	0.5803
	$\mu_{ij2}{}^{+}$	0.182	2.23* (0.082)			
4	$\ln q_{ij}(L_{ij})$	1.019	51.12** (0.020)	-8336.8	1.59	0.5800
	$\mu_{ij2}{}^{-}$	-0.049	-1.25 (0.039)			
5	$\ln q_{ij}(L_{ij})$	1.031	48.82** (0.021)	-8335.3	4.54*	0.5803
	$\mu_{ij3}{}^{+}$	0.215	2.14* (0.100)			
6	$\ln q_{ij}(L_{ij})$	1.017	51.16** (0.020)	-8337.3	0.56	0.5799
	$\mu_{ij3}{}^{-}$	-0.038	-0.74 (0.051)			
7	$\ln(Q_{ij})$	1.019	50.75** (0.020)	-8337.2	0.78	0.5799
	$\mu_{ij4}{}^{+}$	0.117	0.91 (0.129)			
8	$\ln q_{ij}(L_{ij})$	1.014	51.35** (0.020)	-8330.5	14.27**	0.5812
	$\mu_{ij4}{}^{-}$	-0.471	-3.56** (0.132)			
9	$\ln q_{ij}(L_{ij})$	1.026	48.52** (0.021)	-8329.1	16.89**	0.5814
	$\mu_{ij3}{}^{+}$	0.165	1.63 (0.102)			
	$\mu_{ij4}{}^{-}$	-0.442	-3.32** (0.133)			

* and ** denote significance at the 5% and 1% level in a 2-tailed test, respectively.

However, considering the last 30 minutes (Models 3 and 4) and the last 15 minutes of the market (Models 5 and 6), the coefficients of the second variable are significant when odds increase (Model 3: z = 2.23, p = 0.0258; Model 5: z = 2.14, p = 0.0324), but not when odds decrease (Model 4: z = -1.25, p = 0.2150; Model 6: z = -0.74, p = 0.4592). Therefore, large price movements in the later stages of the market do result in odds-implied probabilities differing from true-winning probabilities, but only when odds increase, i.e., bettors herd on increasing odds in the late stages, but not on decreasing odds. Finally, Models 7 and 8 are based on the period between 30 and 15 minutes from the race start. Here we find the opposite effect, i.e., bettors herd on decreasing odds (Model 8: z = -3.56, p = 0.000) but not on increasing odds (Model 7: z = 0.91, p = 0.3682).

These results are all supported by LR tests vs. Model 0: only Models 3, 5, and 8 add significant predictive power over odds alone (Model 3: χ_1^2 = 4.92, p = 0.0266; Model 5: χ_1^2 = 4.54, p = 0.0331; Model 8: χ_1^2 = 14.27, p = 0.0002). Consequently, the results support our second hypothesis, that herding behaviour is only evident in the later stages of the market. There is mixed evidence to support our third hypothesis that bettors herd to a greater extent on 'sell' signals than 'buy' signals. We find that bettors herd to a greater extent on 'sell' signals in the last 15 minutes of the market (which is the most active betting period), but herd to a greater extent on 'buy' signals in the period between 30 and 15 minutes from to the race start.

2.5.3. Economic significance of herding

We estimate Model 9 using the training data. This model includes two variables to account for the herding we observed on increasing odds in the last 15 minutes, and on decreasing odds in the 30 to 15 minute period prior to the race start (μ_{ij3}^+ and μ_{ij4}^-), The results are presented in Table 2.2. We employ this model to estimate winning probabilities in the holdout sample and develop betting strategies to exploit any mispricing. The results are presented in Table 2.3 and Figure 2.3.

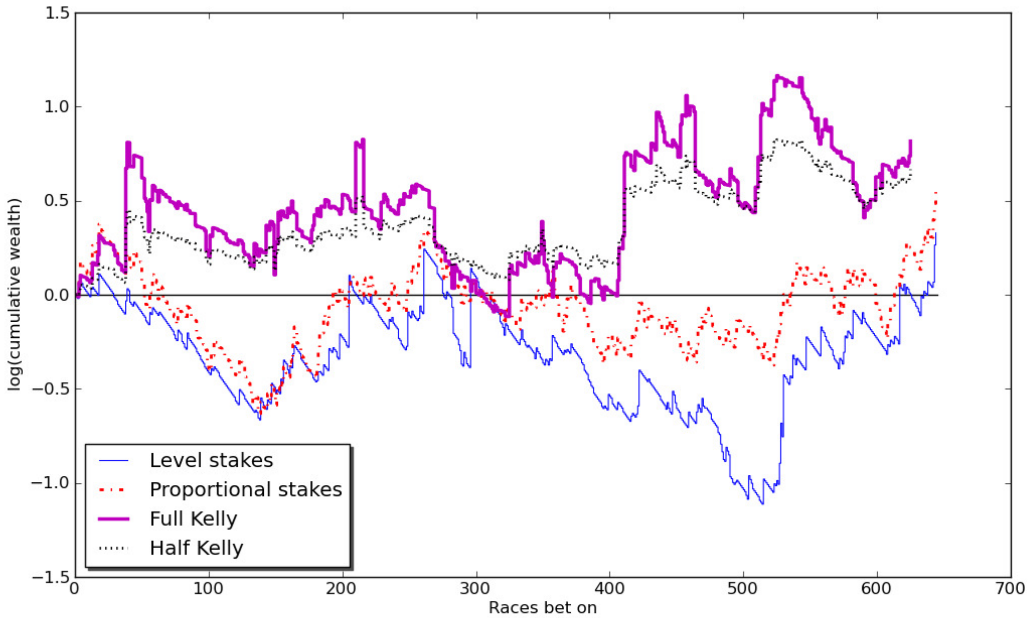
The results show that a strategy of betting against the herd is profitable for all betting strategies. However, the level stakes strategy (rate of return: 5.20%) and the proportional stakes strategy (6.49%) spend a significant portion of the holdout period betting at a loss relative to initial capital (see Figure 2.3 for cumulative wealth for each strategy). On the other hand, the full Kelly (6.16%) and half Kelly (10.39%) strategies rarely drop below initial capital. The greatest monetary accumulation is achieved with the full Kelly strategy, with initial capital increasing by over 126%.

Table 2.3. Results of betting strategies on the holdout set using probabilities estimated from Model 9.

Strategy		*Level stakes*	*Proportional stakes*	*Full Kelly*	*Half Kelly*
Number of races bet on		546	546	532	532
Total number of bets		644	644	625	625
Number of winning bets		111	111	111	111
Total amount bet (£)		7479.9	11759.0	20495.0	9114.3
Final capital (£)		1389.2	1762.9	2262.2	1946.9
Profit or loss (£)		389.2	762.9	1262.2	946.9
Rate of return R (%)		5.20	6.49	6.16	10.39
Risk-adjusted return		0.24	0.48	0.53	0.90
Expected final wealth (£)		1448.4	1637.0	1922.5	1387.6
Probability that final wealth is above x% of initial wealth	50	0.93	0.69	0.48	0.33
	100	0.94	0.74	0.54	0.40
	150	0.76	0.64	0.55	0.49
	200	0.86	0.64	0.47	0.35

However, it is also the riskier of the two Kelly strategies, with 20.5 times the initial capital bet over the course of the holdout period (cf. just 9.1 times for the half Kelly strategy). For the half Kelly strategy, initial capital increases by over 94%. Consequently, the risk-adjusted return is greatest for the half Kelly strategy, with a value of 0.90. Similarly, the full Kelly strategy has the highest expected final wealth and the highest probability of doubling wealth (0.49) but also the lowest probability of retaining at least half of initial wealth (0.76). In summary, the positive returns identified for the various betting strategies, including a sizeable return of 10.39% from our preferred strategy (half Kelly), provide support for our fourth hypothesis, that herding represents an economically significant inefficiency.

Figure 2.3. Log of cumulative wealth relative to initial wealth from betting strategies.



2.6. Discussion

2.6.1. Noise, volatility, and efficiency

Our finding that markets associated with greater noise trading are more both volatile and more efficient contributes evidence to the debate concerning the roles of noise and information in financial markets. We find that noise increases short term volatility, an empirical result that confirms the theoretical predictions of Black (1986) and De Long, Shleifer, Summers, and Waldmann (1990). Previous empirical evidence in this regard has been only marginally conclusive. For example, Campbell and Kyle (1993) found that noise was helpful in explaining historical stock price volatility, but that its importance depends on the interest rate assumption. Our results, from the largest betting exchange market in the world, support the hypothesis that noise trading is associated with an increase in market efficiency. This is the first time such a conclusion has be drawn from an empirical financial market study. The results support Black's (1986) assertion that noise is essential for liquid, efficient markets and they support the theoretical predictions of Grossman and Stiglitz (1980) and Kyle

(1985), that increased noise need not necessarily destabilize markets. Indeed, we would go one step further, and suggest that, provided prices are not extreme, noise actually makes prices more efficient because the improved liquidity allows informed traders to arbitrage away the inefficiency. This conclusion echoes that of Bloomfield, O'Hara, and Saar's (2009) experimental study. The most prominent inefficiencies in financial markets are, of course, bubbles and subsequent corrections, where asset prices are pushed well above their fundamental values before plummeting when the bubble bursts. Arguably these volatile market periods are the times that markets are at their least efficient, since crises can occur even when economic fundamentals are sound (Cipriani and Guarino, 2008). We have demonstrated, conversely, that noise trading and the ensuing volatility can be an important tool for price discovery.

It is important to examine why our result that noise can increase market efficiency contrasts with the arguments of Shiller (1990), De Long, Shleifer, Summers, and Waldmann (1990), and Shleifer and Vishny (1997), among others, that noise is detrimental to market efficiency because of its destabilizing effect on long-run equilibrium values. In fact, the reason for this apparent contradiction appears to lie not in the destabilizing effect of noise trading, per se, but in the limitations to arbitrage. Our results are drawn from a market where noise traders do not introduce added risks for informed traders, limiting their arbitrage. In particular, in regular financial markets, an arbitrageur faces the risk that noise traders continue to keep prices away from fundamental values for an extended period of time, potentially forcing arbitrageurs to liquidate at a loss (De Long et al., 1990). In our study, this risk is not present, because traders in betting markets need not trade their assets in order to realise returns. Instead, returns from state contingent claims are ensured in the initial trade, and traders may simply hold their assets until the market closes and receive their contingent return. Thus, even if prices are noisy, arbitrageurs can effectively guarantee returns at the moment of the initial trade, and need not worry about the future direction of the market. Consequently, our results demonstrate the inherent value of noise trading, in that, when noise trader risk does not limit the arbitrage of informed traders, market efficiency increases as noise trading increases.

The above discussion still begs the question, if noise trading is apparently irrational in the sense that it is loss-making, why does noise trading persist in markets? In regular financial markets, the motivation for noise trading is often assumed to be some portfolio-based requirement such as hedging or liquidity trades (Bloomfield et al., 2009). In betting markets, it is likely that, despite financial losses, noise traders gain utility from the act of gambling itself (Vaughan Williams and Paton, 1998) or are locally risk-loving in the appropriate domain (Friedman and Savage, 1948). There is some evidence that this may also be true of traders in regular financial markets outside of betting. For example, Gao and Lin (2011) present evidence that even institutional investors see lotteries as a substitute for financial market trading. Whatever the motivations of noise traders in regular financial markets, it is apparent from our results in a large exchange betting market that tighter controls on speculators and institutional noise traders, in an effort to reduce more general risks, may serve to increase liquidity risks for other traders.

2.6.2. Herding

While noise trading may have a positive effect on overall market efficiency, this is not the case when herding occurs. In fact, we found that, under certain conditions, herding has a detrimental effect on efficiency. In particular, some price movements are too large relative to the underlying fundamental information, and are such that final market prices can differ significantly from true winning probabilities. We infer from this that traders herd on price movements under certain conditions, pushing prices to inefficient levels. Moreover, the results of our modeling of these price movements show that the larger the price movement, the greater the inefficiency (i.e., larger price movements correspond to greater disparities between final odds-implied probabilities and true winning probabilities). However, this behaviour only becomes significant in the later stages of the market. Large price changes over the full duration of the market do not generally lead to inefficiencies in final market prices. This is not unexpected, since there is a lengthy period during which any inefficiencies induced by herding can be corrected. Moreover, much of the information pertaining to horses' chances is revealed on the day of the race. For example, information concerning results of previous races, jockey changes, and horses' condition and behaviour may not be revealed until the market on a race has opened (Bruce and Johnson, 1995). Therefore, prices *are expected* to change before the final stages of the market (resulting from revealed fundamental information) and herding is therefore unlikely to take place as a result of early stage market price changes. Consequently, it appears that inefficiencies resulting from herding are more likely to occur when (i) there is little time remaining to correct the inefficiency, and (ii) when traders perceive price movements as evidence of trading by those with privileged information. In fact, this conclusion chimes well with classic cases of herding in regular financial markets, such as that evidenced in the South Sea Bubble (Dale et al., 2005).

Our finding that, in the later stages of the market, herding patterns are asymmetric, serves to confirm our prediction that noise traders' perceptions of the actions of informed traders are key to the prevalence of herding behaviour. In particular, while previous studies of herding in betting markets (Law and Peel, 2002; Schnytzer and Snir, 2008) have focused on bookmaker markets, where bettors may only back their preferred horse to win (leading to a reduction in its odds), our study examines a betting exchange, where bettors may also lay horses to lose (leading to an increase in their odds). There is little qualitative difference between backing/laying a horse one thinks will win/lose. Consequently, the differences we observe in herding behaviour (indicated below) must be due to differences in the bettors' perceptions of 'buy' and 'sell' signals. In particular, we argue that this stems from their belief that those traders with privileged information will trade at different times, depending upon whether they believe a horse will win or lose a race.

We find that bettors do not herd on decreasing odds in the last 15 minutes of the market (or in the last 30 minutes, if price changes are considered over the whole 30 minute period). This finding is consistent with the literature on herding in financial markets, which has found little conclusive evidence that investors display herding behaviour. It suggests that the average bettor does not consider a late 'plunge' to be a signal containing valuable information, or, at least, bettors realize that, by the time the plunge has happened, the information is assimilated in the price. Alternatively, the bets placed in the last 15 minutes by informed traders could cancel out the bets of herding traders. This is consistent with the literature, which suggests that strategies of simply betting on horses whose odds decline sharply ('plungers') are not profitable once the price change has occurred (Crafts, 1985; Bird and McCrae, 1987).

On the other hand, we find that bettors *do* herd on plungers that occur early in the betting (in the period between 30 minutes and 15 minutes before race start). In this case, bettors herd to such an extent that further price movements, which happen in the last 15 minutes of the market, are insufficient to restore efficiency. This might be explained in several ways: (i) cash-constrained informed bettors bet early, but they may not have the funds to correct prices for a second time, should prices revert to inefficient levels (Hong and Stein, 1999; Schnytzer and Snir, 2008), or (ii) bets placed in the 30 to 15 minute market segment are generally those of less informed bettors (who might be more likely to herd), since more informed bettors benefit from placing their bets later so as not to divulge their own information (Ottaviani and Sørensen, 2005). In either case, it appears that uninformed bettors perceive that odds that decline sharply in the period 30 to 15 minutes before the race start result from the actions of informed traders (presumably believing that any fundamental information would have been discounted in prices in the earlier stages of the market). It has been found in empirical studies of financial markets that, not only do the trades enacted by informed traders move prices towards efficient levels, uninformed traders are able to detect the informed trading via the volume and direction of the informed trades (e.g., Meulbroek, 1992). Hence, herding will or will not occur depending on (i) uninformed traders' *perception* of the degree of influence over market prices held by traders with privileged information, and (ii) the *actual* degree of influence informed traders have. Consequently, herding will occur only if less informed investors believe that market price movements are currently reflecting the opinions of more informed investors. On the other hand, the extent of herding will be reduced if informed investors have sufficient market power to restore prices to efficient levels.

The perceptions of uninformed traders also appear to play a part when considering whether bettors herd on increasing odds. Large increases in odds in the last 30 or 15 minutes prior to market close often lead to situations where the odds are too high (i.e., the horse is relatively under-valued). This suggests that bettors herd on increases in odds (by laying unfavored horses) even in the late stages of the market. It seems, therefore, that 'sell' signals are treated differently to 'buy' signals. A 'sell' signal is taken seriously even late in the market. This may arise because bettors perceive that it is more likely that individuals with access to privileged information (e.g., horse owners) lay horses to lose (rather than back them to win) late in the market, since it is easier for them to predict (and/or influence) that their horse will lose (Marginson, 2010). It may be perceived that they are more likely to do this later in the market when positive information concerning the prospects of other runners has been fully discounted in prices. Indeed, it is not even necessary that this practice of laying known losers is prevalent, provided bettors perceive that it is.

There is also some evidence from regular financial markets that investors treat 'sell' and 'buy' signals differently. For example, Wermers (1999) found that the level of herding by mutual funds was greater when selling stocks than buying them, particularly if those stocks were shares of small companies with low past returns. This can be explained by mutual funds' particular aversion to small stocks, i.e., they would be more likely to sell past losers in small stocks than buy past winners. Similarly, the model developed by Epstein and Schneider (2008) implies that investors react more strongly to bad (cf. good) news. Our results highlight the importance of understanding potential differences in the manner in which buy and sell signals are perceived by uninformed traders.

Finally, our results demonstrate that considerable inefficiency can be caused by herding. In particular, we not only find that prices are often out of line with true winning probabilities after large price movements, but that trading strategies can be constructed that show consistent positive returns from betting against the herd. Such a strategy is based on a model that accounts for likely differences in noise traders' perceptions of the actions of informed traders at different times in the market. In fact, we find that a half Kelly strategy with some restrictions provides a substantial rate of return (10.39%) over the holdout sample; the return is sufficiently large to compensate for potential variation in returns and the model risk involved. Previous studies of herding in betting markets (Law and Peel, 2002; Schnytzer and Snir, 2008) have demonstrated that positive returns can be made by avoiding following the herd, but these approaches offer very few betting opportunities. On the other hand, our results show that it is possible to develop a strategy to profit by betting against the herd and that this strategy provides a significant number of betting opportunities (betting in over 33 percent of markets). Our results clearly demonstrate that herding is of considerable economic importance, and should be accounted for in more advanced forecasting models.

2.7. Conclusion

This study is the first to study noise trading and herding in an electronic betting exchange (akin to electronic exchanges in regular financial markets). We find evidence that increased noise trading in markets is associated with an increase in efficiency, and we attribute this to informed traders being attracted to the resulting increase in liquidity. We also find that bettors herd, but only under certain conditions. In particular, herding is concentrated in the later, more active stages of the market. In addition, while herding occurs on both 'buy' and 'sell' signals, it does so differently at various times in the market.

Our findings contribute new evidence to the literature on information cascades and herding, where results of empirical studies have been inconsistent. We find support for the theoretical models of herding, in that the asymmetry of information held by bettors is clearly important as an initial condition for subsequent herd behaviour. Herding is rational at the individual level for less informed traders when they are aware that more informed traders may be participating in the market. However, at the aggregate level, herding results in prices departing from efficient levels, particularly when the market has insufficient time to correct the resulting mispricing, or when informed traders are not actively participating. We demonstrate that the inefficiency which remains offers the prospect of abnormal returns for those who seek to capitalize on the herding behaviour of others. Most importantly, we find that herding is extremely prevalent, with inefficiencies in over one third of the markets that we examine. The implications of our findings are that, in wider financial markets, regulations should be considered that minimize the impact of herding, and particular attention should be given to situations where uninformed traders may incorrectly believe that there are traders with privileged information operating. Furthermore, markets that involve contingent returns at a fixed point in time (such as the markets examined here) should always be allowed sufficient time to reach efficient levels.

Our results also cast new light on the relationship between noise and efficiency in financial markets, a relationship that has been difficult to determine in previous studies because of uncertainty in the link between fundamental information and prices. The data we employ overcome this problem, and we are able to measure the degree of informed trading, hence enabling us to observe a positive correlation between noise and efficiency. Our main finding that noise trading, volatility, and efficiency of final market prices all move in tandem has important policy implications for all financial markets. For example, our results add weight to arguments that regulatory measures to protect investors from the destabilizing effects of noise are self-defeating. Of course, the operations of betting markets themselves are often restricted or banned outright on the basis that gamblers should be saved from themselves. But when they are in operation, no one would suggest that the involvement of noise traders should be limited. However, in conventional financial markets, so long as there is a social cost to unwitting participants in market volatility, the actions of speculators will always be under scrutiny. Our study sheds new light on the potential value and possible costs that such traders can bring to a financial market, and suggests that focusing on innovative means of reducing the risks to arbitrageurs, rather than discouraging speculators, may be the best approach to achieving efficient markets.

3. The favourite-longshot bias in competing betting markets

Abstract

This paper provides an explanation for the enduring presence of the favourite-longshot bias (FLB) in some betting markets and its absence in others. We develop a theoretical model that suggests the bias may result from competition between bookmakers and with betting exchanges, combined with bettors' greater demand elasticity with respect to favourites. Further, we propose that the FLB will be eliminated when informed traders dominate and transaction costs are low. We confirm the model's predictions by analysing how the bias develops throughout the active market in 6058 races run in the UK and Ireland from August 2009 through August 2010.

3.1. Introduction

The favourite-longshot bias (FLB) is a phenomenon in betting markets reported over many decades and in many jurisdictions, whereby market prices deviate systematically from their fundamental value; favourites are under-valued while longshots are over-valued (*USA:* Weitzman, 1965; Ali, 1977; Snyder, 1978; Asch et al., 1982; Snowberg and Wolfers, 2010; *UK:* Dowie, 1976; Vaughan Williams and Paton, 1997; Bruce and Johnson, 2000; Sung and Johnson, 2010; *Australia and New Zealand:* Tuckwell, 1983; Gandar et al., 2001). However, a few studies have found no evidence of the bias (Busche and Hall, 1988; Busche, 1994; Swidler and Shaw, 1995) and there is mixed evidence of whether the extent of the bias is substantial enough to result in weak form inefficiency (Sung and Johnson, 2010).

Many studies have sought to explain the enduring presence of the FLB and its absence in some markets, but little consensus has been reached (see Jullien and Salanié, 2008, for a recent review). The existing accounts alternatively link the origin of the FLB to supply- or demand-side factors, related to characteristics of the market or the bettors, respectively. However, these alternative accounts fail to provide a satisfactory explanation for the presence/absence of the FLB in the two major competing types of horserace betting market in the UK and Ireland (and in other jurisdictions, such as Australia): bookmakers and betting exchanges. The rapid growth of betting exchanges, and an intensification of competition through internet–based betting in general, makes such an explanation an important objective. We seek to achieve this by exploring two aspects of these parallel markets: competition and informed trading. The markets for horserace betting in the UK are increasingly competitive, with many different operators and a wealth of information regarding prices available to bettors. We develop a model to investigate the optimal pricing decisions of bookmakers when the betting public are able to rapidly compare prices. We also argue that informed trading has a significant effect on reducing the degree of the FLB in the markets, but only when transaction costs are low. We use the predictions of our model to develop hypotheses, which we test empirically by analysing how the bias develops over the course of the markets for 6058 races run between August 2009 and August 2010, requiring the analysis of over 5.5 million market prices in total.

Our results confirm that three factors contribute to the existence of the FLB: the pricing decisions of bookmakers, the availability of information, and the level of transaction costs. First, we show that, because of (i) competition between bookmakers and with exchanges, and (ii) bettors' demand for competitive prices on favourites, bookmakers' optimal pricing decisions necessarily lead to the FLB. Second, we show that the FLB is present in exchange prices in the early stages of the market and it is not eliminated, because of higher transaction costs in the form of wider bid-ask spreads. Finally, we draw upon models of prediction markets (Gjerstad, 2005; Wolfers and Zitzewitz, 2006b) to suggest that, when informed traders dominate, any FLB in betting exchange prices is likely to be short-lived, and we find that this is the case.

The remainder of this paper is organized as follows. Section 3.2 provides a brief overview of alternative explanations for the FLB. In section 3.3 we develop a model of the FLB in competing markets. In section 3.4, we derive hypotheses and introduce the data and method employed to test the hypotheses. The results are presented in section 3.5 and discussed in section 3.6. We draw conclusions in section 3.7.

3.2. The origins of the favourite-longshot bias

3.2.1. Supply-side explanations

Some authors have argued that market ecology must be taken into account when seeking explanations for the FLB. Shin (1991, 1992, 1993) modelled price-setting in bookmaker markets as a game between a profit-maximizing bookmaker and a randomly chosen bettor. The model assumes that the bettor is likely to be a noise trader, but could be an insider whose superior knowledge allows them to bet on the winning horse, to the bookmaker's cost. Shin's model can explain the FLB in bookmaker market prices, provided one accepts that knowledgeable insiders are more likely to bet on longshots than favourites. Although some of the assumptions in Shin's model are unrealistic, similar conclusions have been reached where the assumptions are relaxed. For example, Schnytzer and Shilony (2005) found that bookmakers should raise prices on longshots more than favourites in order to defend themselves against insider knowledge, without assuming that insiders know which horse will win the race, or that insiders are more likely to bet on longshots. Peirson and Smith (2010) extend the Shin model while relaxing the assumptions that insiders know which horse will win the race, and that the amount bet by insiders is fixed and not related to the odds on offer. Their model demonstrates that bookmakers should increase prices on those horses where there is a higher probability of inside information being employed.

Transaction and information costs have also been identified as possible causes of the FLB. Hurley and McDonough's (1995, 1996) model implies that the FLB would not exist in pari-mutuel markets without transaction costs, and that the extent of the bias should increase as transaction costs increase. However, this theory is not supported by their empirical investigation. Terrell and Farmer (1996) suggest that the FLB results from costly information. This is supported by Vaughan Williams and Paton (1997), who find a lower degree of

FLB in bookmaker market prices in higher grade handicap races (where it is assumed information is more widely available and thus is less costly to obtain). Sobel and Raines (2003) also found a greater level of the FLB in races that, because there is less information available to inform decisions, attract lower betting volumes. Smith, Paton, and Vaughan Williams (2006) compared the effect of altering the level of transaction costs on the level of the bias in bookmaker markets and exchanges, which typically have higher and lower transaction costs, respectively. They found that there was significantly more bias in bookmaker market prices. However, Smith (2010) noted an exception to the predictions of the transaction costs model: there is little FLB in UK pari-mutuel market prices, yet these markets involve relatively high transaction costs.

3.2.2. Demand-side explanations

Two broad approaches have emerged that seek to attribute the FLB to factors associated with the decision-making processes of bettors. On the one hand, it has been suggested that bettors have unbiased expectations, but are risk-loving (Weitzman, 1965; Quandt, 1986; Hamid et al., 1996). Alternatively, bettors are risk-neutral, but have biased expectations (Henery, 1985; Chadha and Quandt, 1996). The former approach was originated by Weitzman (1965), who suggested that the bias must be explained by hypothesising a convex utility of wealth function for the average bettor (i.e., the average bettor is risk-loving). Quandt (1986) extended this model to show that the bias is a natural consequence of equilibrium in a market where bettors are risk-loving. Variants of the model have been developed, including replacing bettors' risk-loving nature with their desire for skewness (Golec and Tamarkin, 1998) or the extra utility they gain from long-odds betting (Thaler and Ziemba, 1988). In sum, these studies argue that it is bettors' motivations that cause the FLB, not bias in their expectations. The second broad set of demand-side explanations for the FLB is based on prospect theory (Kahneman and Tversky, 1979). In this case, bettors are risk-neutral, but misestimate probabilities, attaching moderate winning probabilities to horses which are more likely to win or lose. Snowberg and Wolfers (2010) compare these two sets of competing demand-side theories by testing the implications for the pricing of compound events (e.g., exactas, where the bet is to choose both the winner and the runner-up in the correct order) and find evidence in support of the latter, biased-expectations approach.

3.2.3. Competing markets

The search for a full explanation of the FLB anomaly is further complicated by the presence of different types of betting market, with varying rules, costs and participants. In the UK, bookmaker and exchange markets account for most betting activity, with 94% of horserace betting turnover (over £5.7 billion) in the year to March 2010 in the UK (cf. £356 million for the pari-mutuel operator, the Tote) (Gambling Commission, 2010). These markets operate in parallel in the UK, Ireland, and a number of other jurisdictions. There are three significant differences between the two types of market. First, in bookmaker markets, individuals are allowed to bet on their preferred contestant (e.g., a horse in a race) at the advertised odds. In exchanges, individuals can either bet on their preferred contestant to win, or, alternatively, can lay a contestant to lose (i.e., offer to match bets placed by other bettors on this contestant). Second, in bookmaker markets, odds are set by the bookmaker, whereas the prices in exchanges are a strict representation of supply and demand, and are reached as an implied consensus of all the market participants. Finally, in exchanges, participants typically pay only a small commission on their net winnings (e.g., 5%), whereas bookmaker transaction costs (implicit in the over-round; see later) are typically significantly higher (e.g., 18%). For a full explanation of betting exchanges, see Smith and Vaughan Williams (2008). Both types of market involve a number of competing operators, the latest odds are readily available to bettors through the internet, and bets can be easily and rapidly placed with exchanges or with any chosen bookmaker using a mobile phone, from wherever the bettor is located (e.g., racetracks, home, high street betting shops). This is important, because it has been found that demand becomes increasingly price-sensitive and frictionless when internet price search is available (Ellison and Ellison, 2009).

A few recent studies have examined the degree to which the FLB is present in betting exchanges and their relation to other types of market, and have shown that betting exchanges are significantly more efficient, with a lower degree of the FLB (Smith et al., 2006) and greater predictive accuracy of market prices (Smith et al., 2009; Smith, 2010; Franck et al., 2010). However, these studies focus on market prices at one point in time (early or final prices). In addition, many of the studies discussed above explain the FLB in terms of the average bettor (e.g., Weitzman, 1965) and fail to account for clear differences in the behaviour of informed and uninformed bettors (e.g., Shin, 1993; Sobel and Raines, 2003). Hence, we will define uninformed bettors as those who display non-neutral risk tendencies or are biased in their expectations; informed bettors are risk-neutral and have unbiased expectations. The latter behaviour is to be expected of bettors who have developed expertise through repeated practice and extensive study. We also consider insider traders to be informed bettors since, as Schnytzer and Shilony (1995) demonstrated, inside information is a significant predictor of race outcomes.

3.3. A model of competing markets

3.3.1. Bookmaker markets and competition

We develop a model to explain the FLB in bookmaker markets by considering two competing markets, a bookmaker and an exchange, both of which offer prices on all horses running in a single race with n runners. Traders buy contracts on horse i, which pay out £1 if the horse wins the race, and which cost £q_i and £r_i in the bookmaker and exchange markets, respectively. Consequently, the implied bookmaker market odds (it is odds that are normally quoted by bookmakers rather than prices) of horse i are the reciprocal of the price minus one, i.e., $1/q_i - 1$, and represent the profit on a winning £1 bet. Similarly, the implied exchange market odds of horse i are given by $1/r_i - 1$ and represent the profit on a winning £1 bet less 5% commission (which is the typical amount). The over-round is the sum of purchase prices across all the horses in the race minus one, and represents the average transaction cost to a bettor. A lower over-round, therefore, allows a more competitive set of prices. The bookmaker and exchange odds-implied probabilities of horse *i* winning the race are given by $q_i/(1+B_b)$ and $r_i/(1+B_e)$, respectively, where $B_b = \sum_{i=1}^{n} q_i - 1$ and $B_e = \sum_{i=1}^{n} r_i - 1$ are the bookmaker and exchange over-rounds, respectively, and we assume that the exchange is able to offer lower transaction costs, i.e., $B_b > B_e$.

The task for the bookmaker is two-fold: to estimate the true probabilities p_i for each horse to win the race, and to set their own prices so as to maximize their overall profit from the race. Since the true probabilities are unknown, we assume that the bookmaker's best estimates for them are simply the exchange odds-implied probabilities, i.e., $p_i = r_i/(1+B_e)$. Later, we will show empirically that this is appropriate. Finally, we make the approximation that the q_i are continuous on the interval (0, 1). Considering a small time interval, the bookmaker's goal is to maximize their expected returns from overall bets taken in this interval. We allow this interval to be small because the bookmaker can update their prices at the end of the time interval, and we suppose that they are able to do this very frequently. Since this interval is small, we make the restriction that B_e, B_b, and the r_i are constant in this period. Then the bookmaker aims to maximize their expected returns G $(q_1, \ldots\ldots, q_n)$ over this time interval,

$$(3.1)\; G(q_1,\ldots,q_n) = \sum_{i=1}^{n} f(q_i, r_i)\left[(1-p_i) - p_i\left(\frac{1}{q_i} - 1\right)\right] = \sum_{i=1}^{n} f(q_i, r_i)\left(1 - \frac{r_i}{(1+B_e)q_i}\right),$$

subject to the over-round condition $\sum_{i=1}^{n} q_i = 1 + B_b$. The demand curve $f(q_i, r_i)$ is the amount bet on horse *i*, when the bookmaker and exchange prices are q_i and r_i, respectively. We expect that when the bookmaker and exchange prices are equal, the demand in the two markets will be identical. Consequently, we normalize the demand curve so that it satisfies $f(r,r) = r/(1+B_e)$. If the bookmaker price is set above/below the exchange price, the demand will fall/rise. The rate at which the demand changes with respect to changes in bookmaker prices, relative to exchange price *r*, depends on the price elasticity of demand at *r*. Although we do not know the exact shape of *f*, it represents the amount by which bettors are either discouraged/attracted by uncompetitive/favourable prices, and we imagine that its shape could be empirically derived given knowledge of actual bets taken by bookmakers. Assuming that f is continuously differentiable on (0, 1), we at least know that a sensible demand curve would require $f'(q) < 0$ on this interval. In addition, f, strictly speaking, would need to be non-negative, but we need not impose this. As an example, the form of a linear demand curve with elasticity which can vary with r is given by

(3.2) $$f(q,r) = \frac{r}{1+B_e} - g(r)(q-r),$$

where $f'(q,r) = \partial f / \partial q = -g(r)$ defines the price elasticity of demand at r.

For fixed r_i, B_b and B_e, this is a constrained optimization problem to maximize $H(q_1,\ldots,q_n)$ where H is given by

(3.3) $$H(q_1,\ldots,q_n) = G(q_1,\ldots,q_n) - \lambda(\textstyle\sum_{i=1}^{n} q_i - 1 - B_b),$$

where λ is a constant. The solutions are given by the system of equations

(3.4) $$\partial H / \partial q_i = f(q_i,r_i)\frac{r_i}{(1+B_e)q_i^{\,2}} + f'(q_i,r_i)\left(1 - \frac{r_i}{(1+B_e)q_i}\right) - \lambda = 0,\ i = 1, \ldots, n,$$

$$\textstyle\sum_{i=1}^{n} q_i = 1 + B_b.$$

We do not seek to solve this system of equations, but note a sufficient condition for the FLB in the bookmaker prices: for two horses j (a longshot) and k (a favourite: $r_j < r_k$), with odds-implied probabilities equal across the exchange and bookmaker markets ($q_j = (1+B_b)r_j/(1+B_e)$, $q_k = (1+B_b)r_k/(1+B_e)$), the marginal increase in expected returns for an increase in price is greater for the longshot ($\partial H/\partial q_j > \partial H/\partial q_k$); furthermore, the greater the difference, the greater is the level of FLB. So, denoting $\hat{q}_i = (1+B_b)r_i/(1+B_e)$,

(3.5) $$\frac{1}{r_j} f(\hat{q}_j,r_j) - \frac{1}{r_k} f(\hat{q}_k,r_k) + \frac{1+B_b}{1+B_e} B_b[f'(\hat{q}_j,r_j) - f'(\hat{q}_k,r_k)] > 0.$$

We have now the following as a sufficient condition for the FLB:

(3.6) $$f'(xr_j,r_j) > f'(xr_k,r_k),$$

for all x such that $1 \le x \le (1+B_b)/(1+B_e)$.

If this condition is satisfied, then so is (3.5) (for a proof, see Appendix 2). Moreover, the difference $\partial H / \partial q_j - \partial H / \partial q_k$ is increasing in both B_b (holding $(1+B_b)/(1+B_e)$ fixed) and $(1+B_b)/(1+B_e)$, i.e., the level of FLB as imposed by optimal bookmaker prices should increase both with bookmaker over-round and the level of competition between the bookmaker and the exchange[14]. Also, if there is no competition ($B_b = B_e$), then there is still FLB (provided (3.6) holds), and further, if there is no over-round ($B_b = 0$) then there is no FLB (for proofs of each of these propositions, see Appendix 2). Condition (3.6) holds if bettors' price elasticity of demand is greater for favourites than for longshots, i.e., if bettors are driven away from betting at uncompetitive prices more rapidly for favourites than for longshots (or if bettors are attracted by favourable prices more strongly for favourites than for longshots). Provided this is the case, it follows that bookmakers are driven to post competitive prices on favourites or risk losing business to their competitors, even if, as a result, they must offer 'poor value' prices on longshots. Finally, it should be noted that the competition need not be an exchange; the analysis is identical if the competition comes from another bookmaker.

In the case of a linear demand curve with variable elasticity (equation 3.2), condition (3.6) is satisfied if $g(r_k) > g(r_j)$, for example, if $g(r_i) = Ar_i$ for some $A > 0$. Table 3.1 shows the optimal q for various values of r, A, B_b and $(1+B_b)/(1+B_e)$ in a race with two runners, and indicates that the level of FLB increases with elasticity, over-round and competition. Figure 3.1 shows how this curve varies for different values of r. For a linear demand curve, there is FLB in the optimal bookmaker prices, provided elasticity is greater for favourites than for longshots. This would also be the case for non-linear demand curves.

---

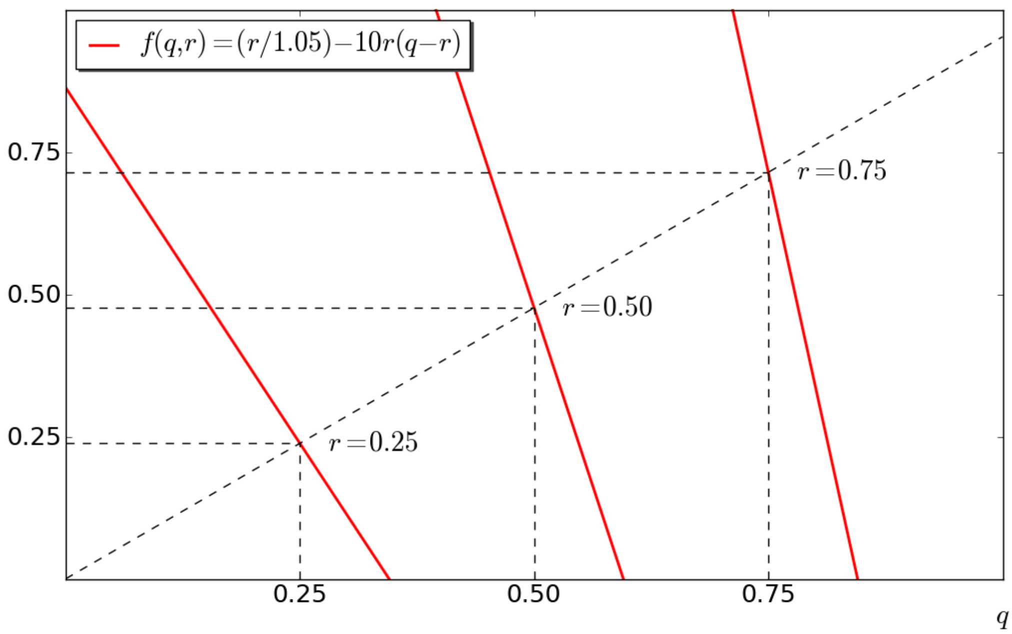
[14] Smith and Vaughan Williams (2010) show empirically that the level of FLB in bookmaker markets fell after the introduction of exchange markets. Although the correlation between increased competition and reduced FLB is inconsistent with the prediction of our model, it was also the case that bookmaker over-rounds fell during the same period, which could account for the reduced FLB.

Table 3.1. A model of competing markets: the favourite-longshot bias in a two-horse race.

A = 10, B_b = 0.18, $(1+B_b)/(1+B_e)$ = 1.18/1.05		r = 0.700, B_b = 0.18, $(1+B_b)/(1+B_e)$ = 1.18/1.05		r = 0.700, A = 10, $(1+B_b)/(1+B_e)$ = 1.18/1.05		r = 0.700, A = 10, B_b = 0.18	
R	$\hat{q}$	A	$\hat{q}$	B_b	$\hat{q}$	$(1+B_e)/(1+B_b)$	$\hat{q}$
0.100	0.163	1	0.690	1.13	0.678	1.18/1.05	0.676
0.200	0.246	5	0.679	1.17	0.676	1.18/1.04	0.674
0.300	0.331	10	0.676	1.21	0.675	1.18/1.03	0.672
0.400	0.417	50	0.672	1.25	0.675	1.18/1.02	0.671
0.500	0.503	100	0.671	1.29	0.675	1.18/1.01	0.669
0.600	0.590	500	0.671	-	-	1.18/1.00	0.667
0.700	0.676	1000	0.671	-	-	-	-
0.800	0.761	-	-	-	-	-	-
0.900	0.846	-	-	-	-	-	-

$\hat{q} = (1+B_b)/q(1+B_e)$, where q denotes the optimal bookmaker price for the horse, as estimated from the model in section 3.3.

Figure 3.1. Linear demand curves with variable elasticity $g(r)$ = 10r.



In summary, provided we accept some reasonable assumptions, and a sensible demand curve, this model demonstrates that when bettors' price elasticity of demand is greater for favourites than for longshots, the optimal pricing decision of the bookmaker leads to FLB in their prices, and this FLB increases with the level of competition between the bookmaker and their competitors. We note that our model is consistent with the argument of Levitt (2004), who showed that the optimal pricing policy of bookmakers is to distort prices to reflect the biases of bettors.

3.3.2. Betting exchanges and informed trading

Now we consider an exchange market, where traders can both buy and sell contracts on horse *i* for £r_i. As Wolfers and Zitzewitz (2006b) and Gjerstad (2005) demonstrate, if we assume that traders' beliefs and wealth levels are heterogeneous and that the difference between buy and sell prices (the bid-ask spread) is zero, and denote an individual trader's belief of the probability of horse *i* winning as p_i, and the trader's wealth as *w*, each drawn from a distribution $F(p_i, w)$, then the equilibrium price is a 'wealth-weighted average' of the beliefs of all traders,

(3.7) $$r_i = \frac{1}{\overline{w}}\int_{-\infty}^{\infty} p_i w dF(p_i, w),$$

where $\overline{w}$ is the average wealth level across all traders. Now, suppose that traders are either informed or uninformed. We assume that the informed traders know the true probability p_i of horse *i* winning the race and have combined wealth *X*. The uninformed traders have beliefs p_{iy} drawn from distribution $F(p_{iy})$ and combined wealth *Y*. Uninformed traders do not know the true probability, so $E(p_{iy}) \neq p_i$. Without loss of generality, suppose $X + Y = 1$. Consequently, from (3.7),

(3.8) $$r_i = Xp_i + Y\overline{p}_{iy}.$$

This is the 'wealth-weighted average' of beliefs of the informed bettors and the uninformed traders. A consequence is that if informed or uninformed traders dominate the market, the price will be close to p_i or $\overline{p}_{iy}$, respectively.

Now we relax the assumption that bid-ask spreads are zero, i.e., contracts may only be bought for £ $(r_i + t)$ and may only be sold for £ $(r_i - t)$, for some $t > 0$. Now, when supply meets demand, we have that

(3.9) $$\int w(p_i - r_i)dF(p_i < r_i - t, w) + \int w(p_i - r_i)dF(p_i \geq r_i + t, w) = 0 .$$

It is clear from (3.9) that higher transaction costs decrease both supply and demand. In addition, as Wolfers and Zitzewitz (2006a) note, transaction costs increase the proportion of trading done by the traders with the noisiest observations, as the trades contributing to supply and demand in (3.9) result from those traders with beliefs further away from the median belief. This causes prices to deviate from objective probabilities further, a result which is consistent with Hurley and McDonough (1995, 1996) and Sobel and Raines (2003) in that transaction costs increase inefficiency in market prices. We now employ the models developed in this section to derive testable hypotheses.

3.4. Hypotheses, data, and method

3.4.1. Hypotheses

Equation (3.8) approximates the mechanism by which prices are set on betting exchanges and suggests that the FLB can only arise on an exchange as a result of some proportion of wealth belonging to uninformed traders, who under-/over-estimate the winning chances of favourites/longshots. Furthermore, the extent of any FLB is exacerbated by two factors: a lack of informed trading, and higher transaction costs. Markets for horse races typically begin on the evening before the race and informed traders are likely to be more in evidence in the later stages of the market for the following reasons. First, significant information may only emerge in the later stages of the market, including details concerning results of previous races at the race meeting, non-runners, jockey changes, and, even later in the market, the horses' condition and behaviour (Bruce and Johnson, 1995). All these details can impact the way the race is run, so bettors may benefit from waiting to discern this information. Second, trading is light in the early stages of the market, enabling only relatively small bets to be placed. Consequently, informed traders are unlikely to be prepared to give away their information cheaply to other bettors, as would be the case if they placed what would have to be relatively small bets in this early market (Asch et al., 1982). Third, the lighter trading in the early stages of the market is likely to lead to higher over-rounds (since there is a greater divergence of opinions) and this is likely to deter informed traders who are looking for value. As liquidity increases, and opinions become less divergent, we expect spreads to narrow. Indeed, as discussed above, we expect Y to decrease and X to increase over the period of the market, until, in (3.8), $r_i \approx p_i$ for all horses in the race (i.e., prices will be accurate and not include the FLB). This motivates our betting exchange hypotheses:

In betting exchanges,

1. prices approach true winning probabilities over time,

2. the FLB is eliminated in the later stages of the market.

However, we have argued above that the FLB is likely to persist in bookmaker prices because they are competing with other bookmakers or exchanges, and bettors demand competitive prices on favourites (equation 3.6). The models developed above predict that bookmakers' prices on favourites will be very close to their respective exchange prices. In addition, because of higher over-rounds, the prices in bookmaker (cf. exchange) markets on longshots are likely to be significantly higher. Furthermore, bookmaker prices cannot be said to follow a 'wealth-weighted average' of beliefs. Consequently, if the FLB is present in bookmaker markets, there is no obvious mechanism by which it might be eliminated by informed traders. In particular, informed traders are likely to be deterred by the significantly higher over-rounds in these markets. In addition, bookmakers impose restrictive limits on bet size on the most informed and wealthy traders (Smith et al., 2009), or may refuse to do business altogether. The combined effect of less informed trading and more FLB is likely to mean that bookmaker prices do not predict race results as accurately as exchange prices. However, we would expect their accuracy to increase as more information becomes available in the later stages of the market. This motivates our bookmaker hypotheses:

In bookmaker markets,

3. prices approach true winning probabilities in the later stages of the market, but not as quickly as exchange prices,
4. the FLB is present at all stages of the market,
5. prices on favourites closely match exchange prices.

3.4.2. Data

The data employed are odds and finishing positions for 62,124 horses running in 6058 races in the UK and Ireland from August 2009 through August 2010. In particular, we collected matched bookmaker and exchange prices on each horse throughout the duration of the market on each race, from 4 hours before the race start and at intervals up until the start of the race (3 hours, 2 hours, 1 hour, 30, 15, 10, and 5 minutes). The data were downloaded using the Betfair API and directly from the bookmakers' websites. The exchange prices are those of the largest UK betting exchange by traded volume, Betfair, and we use the best prices at which it is possible to back the horse to win, in order to make a fair comparison with equivalent prices in bookmaker markets[15]. The bookmaker prices are recorded as the mean prices offered by a broad cross section of nine leading bookmakers; the data, therefore, includes over 5.5 million price points. The number of runners per race ranges from 2 to 30, with a mode of 9.

We focus on 'win' bets, so the finishing positions were recorded as 1 for a winner and 0 otherwise. We also captured the mean betting volume (amount traded) on Betfair. We find that this trading volume is concentrated in the last moments before the race starts; with an average of 57.5% of total volume matched in the last 5 minutes of the market (see Table 3.2 in section 3.5.1). We suspect that a similar pattern may be true in bookmaker markets, but this data is not available. There is a strong positive correlation between exchange betting volume and accuracy in each market (exchange: corr. = 0.80, p = 0.0046; bookmaker: corr. = 0.91, p = 0.0004; see section 3.5.1).

---

[15] We do make a minor assumption in that prices are equally valid whatever the stake limit at that price. This will have the effect of slightly understating the over-round and the level of FLB, but the inaccuracy will be small enough to be negligible.

3.4.3. Method

To measure the extent of the FLB in bookmaker and exchange prices, we use a conditional logit (CL) modelling approach (McFadden, 1974), which has been employed in several betting market studies (Figlewski, 1979; Asch et al, 1984; Bolton and Chapman, 1986; Benter, 1994; Sung and Johnson, 2010). The CL model enables us to estimate, based on previous race results, the objective probability of a particular horse winning a particular race, given a set of horse-related variables, whilst taking into account the competition in the race. With price as the single variable, the CL model is an effective method for estimating the level of FLB, and is formulated as follows.

Define an estimate of horse *i*'s ability to win race *j* as

(3.10) $$W_{ij} = \beta \ln q_{ij} + \varepsilon_{ij},$$

where β is the parameter that determines the importance of the log of the price q_{ij} for horse i in race j and ε_{ij} is an independent error term. McFadden (1974) shows that if the independent errors are identically distributed according to the double exponential distribution, then the probability of horse i winning race j is given by

(3.11) $$p_{ij} = \Pr(W_{ij} > W_{kj}, k = 1,2,\dots,n_j, k \neq i) = \frac{\exp(\beta \ln q_{ij})}{\sum_{k=1}^{n_j} \exp(\beta \ln q_{kj})} = \frac{{q_{ij}}^{\beta}}{\sum_{k=1}^{n_j} {q_{kj}}^{\beta}},$$

where n_j is the number of horses in race *j*.

The parameter β is estimated by maximizing the joint probability of observing the results of all the races in the sample. This is achieved by maximizing the log-likelihood

(3.12) $$\ln L = \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij}$$

where $y_{ij} = 1$ if horse *i* won race *j* and $y_{ij} = 0$ otherwise, and N is the total number of races in the sample. If the estimated value of β is one, this implies that the odds-implied probabilities are equal to the true probabilities, as realized by the race results. If this estimated value is greater than one, this implies that the FLB is present; the greater the value of β, the greater is the degree of FLB (Bacon-Shone et al., 1992). To test whether the FLB is present, we employ the standard normal test statistic $z = (\beta - 1)/S.E.(\beta)$.

The log-likelihood also gives rise to a natural measure of the predictive accuracy of the bookmaker and exchange prices, which can be compared through time as the market evolves. This is the McFadden pseudo-R^2, which is given by

(3.13) $$R^2 = 1 - \frac{\ln L}{\ln L_n},$$

where $\ln L_n$ is the log-likelihood of the naive model, where each horse in a race is assigned an identical probability of winning:

(3.14) $$\ln L_n = \sum_{j=1}^{N} \ln(1/n_j) \ .$$

In order to test if the predictive accuracy of prices increases over time, and if the predictive accuracy of exchange prices is greater than that of bookmaker prices, we compare pseudo-R^2 values over time. However, it is not straightforward to apply a measure of precision to values of pseudo-R^2, because their distributions are complex and depend on unknown parameters. Here, we adapt the method of Hu, Shao, and Palta (2006) and estimate the *asymptotic distribution* of the pseudo-R^2s, i.e., the expected distribution as the number of races tends to infinity. For more details, see section 5.4.2.

Finally, in order to compare the actual prices offered by the exchange and the bookmakers, we follow Ali (1977) and rank horses by whether they are the favourite, the second favourite, and so on (by exchange prices). An approximate standard normal test statistic to compare exchange prices r with bookmaker prices q is given by

(3.15) $$z = \frac{r-q}{SE(r,q)} = \frac{r-q}{\sqrt{\frac{r(1-r)+q(1-q)}{N}}} \ .$$

3.5. Results

3.5.1. Predictive accuracy

First, we examine the predictive accuracy of exchange and bookmaker prices throughout the duration of the market. The results are presented in Table 3.2. Comparing the predictive accuracy of prices at the time the race starts (starting prices) with prices available throughout the market, we find that the pseudo-R^2 of CL models including price as the only independent variable (see equation 3.11) increases over time in both markets (see Figure 3.2), although this finding is not statistically significant over the last hour of the market duration. Similarly, exchange prices are consistently more accurate than bookmaker prices at predicting race winners, but the difference is not statistically significant at any time point.

Figure 3.2. Predictive accuracy of exchange and bookmaker odds through time, with betting volume.

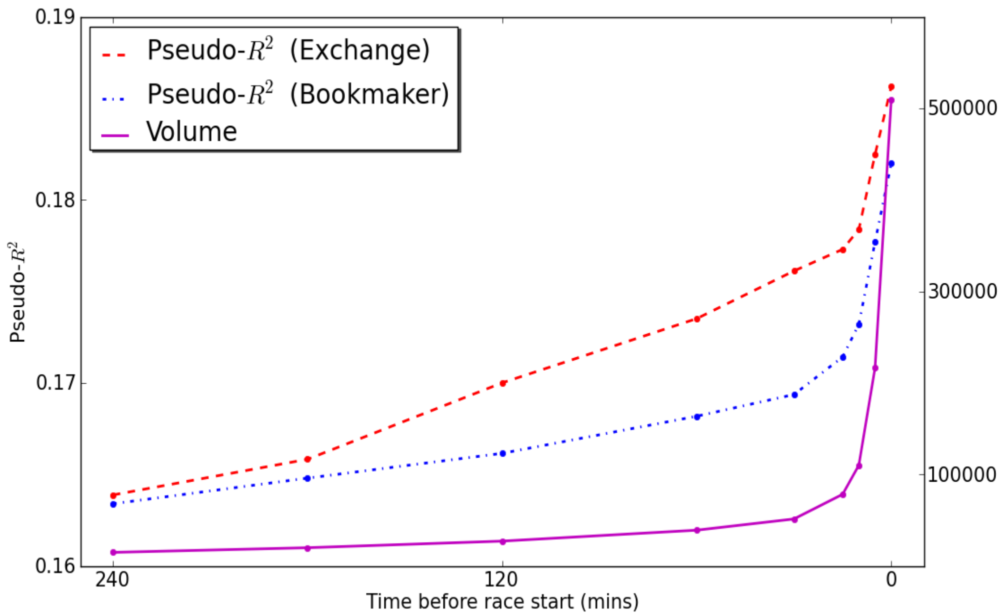


Table 3.2. Betting volume and predictive accuracy in exchange and bookmaker odds over time.

Time until race start (mins)	Cumulative exchange betting volume *vol* (£ 000) (% of final volume)	Exchange pseudo-R^2 $R_e^2(t)$	$z_e(t)$	Bookmaker pseudo-R^2 $R_b^2(t)$	$z_b(t)$	$z(t)$
240	14.7 (2.9)	0.1639	2.84**	0.1634	2.29*	0.06
180	19.8 (3.9)	0.1658	2.59**	0.1650	2.12*	0.13
120	27.0 (5.3)	0.1700	2.06*	0.1661	1.95*	0.48
60	39.0 (7.7)	0.1735	1.62	0.1682	1.70*	0.67
30	51.2 (10.0)	0.1761	1.28	0.1694	1.55	0.85
15	78.1 (15.3)	0.1773	1.37	0.1714	1.30	0.74
10	109.9 (21.6)	0.1784	1.00	0.1732	1.08	0.65
5	216.3 (42.5)	0.1825	0.48	0.1778	0.53	0.60
START	509.5 (100.0)	0.1862	-	0.1820	-	0.52
	Corr(*vol*, *Pseudo-R^2*)	0.79**		0.90**		

* Significant at the 5% level, ** 1% level (1-tailed test).
Standard errors are 0.008 to 3 decimal places in all cases so are omitted.

$$z_e(t) = \frac{R_e^2(START) - R_e^2(t)}{\sqrt{s^2(R_e^2(START)) + s^2(R_e^2(t))}}, \quad z_b(t) = \frac{R_b^2(START) - R_b^2(t)}{\sqrt{s^2(R_b^2(START)) + s^2(R_b^2(t))}},$$

$$z(t) = \frac{R_e^2(t) - R_b^2(t)}{\sqrt{s^2(R_e^2(t)) + s^2(R_b^2(t))}}$$

3.5.2. FLB and over-round

Eight CL models (as per equation 3.11) were developed for both the bookmaker and the exchange markets, respectively, incorporating prices available at the eight different time periods before the race start. The estimated parameters in these models were used to compare the degree of FLB in these markets at different times before the race start (see Tables 3.3 and 3.4 for the exchange and bookmaker markets, respectively; Table 3.4 also includes the difference in the level of FLB between bookmaker and exchange markets).

Table 3.3. The FLB in exchange odds over time.

Time until race start (mins)	FLB β	SE (β)	$\frac{\beta-1}{SE(\beta)}$	$\frac{\beta-\beta_{START}}{SE(\beta,\beta_{START})}$	Over-round B_e
240	1.080	0.0186	4.31**	2.630**	0.113
180	1.072	0.0185	3.92**	2.328*	0.101
120	1.056	0.0181	3.12**	1.718	0.079
60	1.052	0.0181	2.89**	1.552	0.070
30	1.047	0.0180	2.60**	1.331	0.064
15	1.042	0.0180	2.34*	1.143	0.064
10	1.036	0.0178	2.02*	0.909	0.063
5	1.028	0.0176	1.59	0.584	0.058
START	1.014	0.0173	0.79	-	0.057
Corr(β,B_e)	0.918**				

*: significantly different from 1.00 at the 5% level, **: 1% level (2-tailed test).

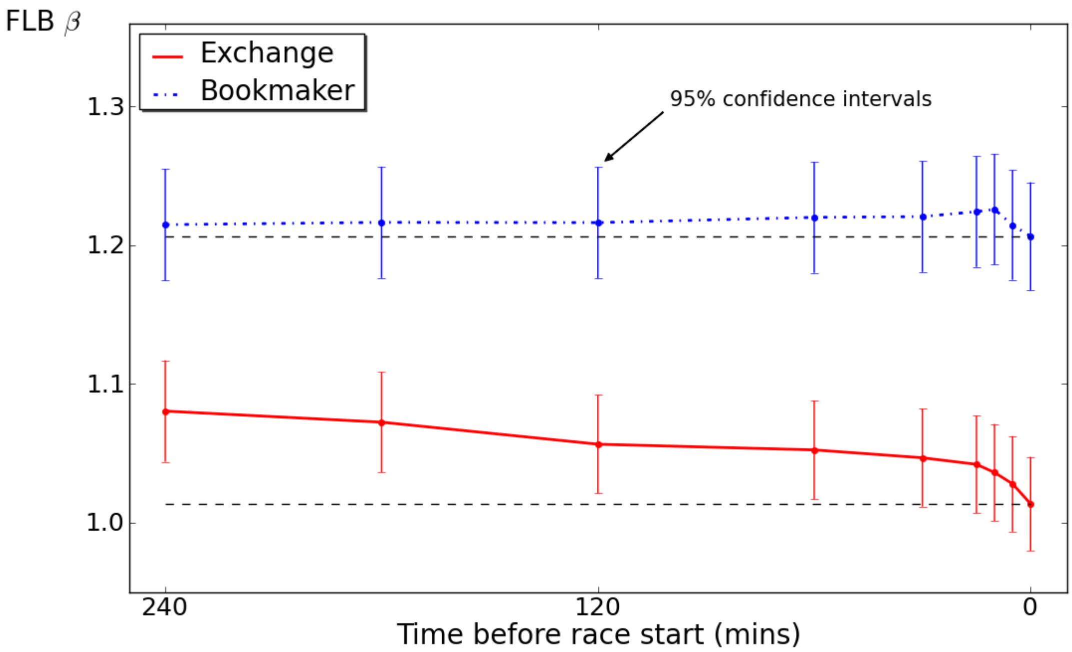
Table 3.4. The FLB in bookmaker odds over time.

Time until race start (mins)	FLB β	SE (β)	$\frac{\beta-1}{SE(\beta)}$	$\frac{\beta-\beta_{START}}{SE(\beta,\beta_{START})}$	$\frac{\beta-\beta_{EXCH}}{SE(\beta,\beta_{EXCH})}$	Over-round B_b
240	1.215	0.0205	10.47**	0.293	4.85**	0.198
180	1.216	0.0204	10.56**	0.351	5.22**	0.207
120	1.216	0.0204	10.58**	0.345	5.85**	0.214
60	1.216	0.0204	10.77**	0.478	6.14**	0.217
30	1.220	0.0204	10.82**	0.499	6.40**	0.217
15	1.224	0.0204	11.01**	0.624	6.71**	0.200
10	1.226	0.0204	11.09**	0.683	7.01**	0.190
5	1.214	0.0201	10.65**	0.286	6.98**	0.180
START	1.206	0.0198	10.40**	-	7.33**	0.181
Corr(β,B_b)	0.394					

*: significantly different from 1.00 at the 5% level, **: 1% level (2-tailed test).

The results show that there is FLB in exchange prices 30 minutes or more before the race start (the parameter is significantly greater than 1.00 at the 1% level: z = 2.60, p = 0.0096) when just 10% of final volume has been traded at this stage. Most importantly, the results show that there is no significant FLB in exchange prices in the later stages of the market (e.g., based on prices 5 minutes before the race start: z = 1.59, p = 0.1142). In addition, the FLB in exchange prices is significantly greater 3 hours before the race start than it is at race start (z = 2.33, p = 0.0204). On the other hand, there is FLB in bookmaker prices available at all times before and at race start (i.e., the parameter in the CL model is significantly greater than 1.00 at the 1% level based on all these sets of prices, e.g., at race start: z = 10.40, p = 0.0000). Moreover, when comparing the level of FLB in the bookmaker prices at different times throughout the market with that at race start, there is no apparent trend (see Figure 3.3). In addition, the FLB is greater in the bookmaker prices than in the exchange prices at all times (e.g., comparing the FLB in bookmaker vs. exchange prices at race start: z = 6.98, p = 0.0000).

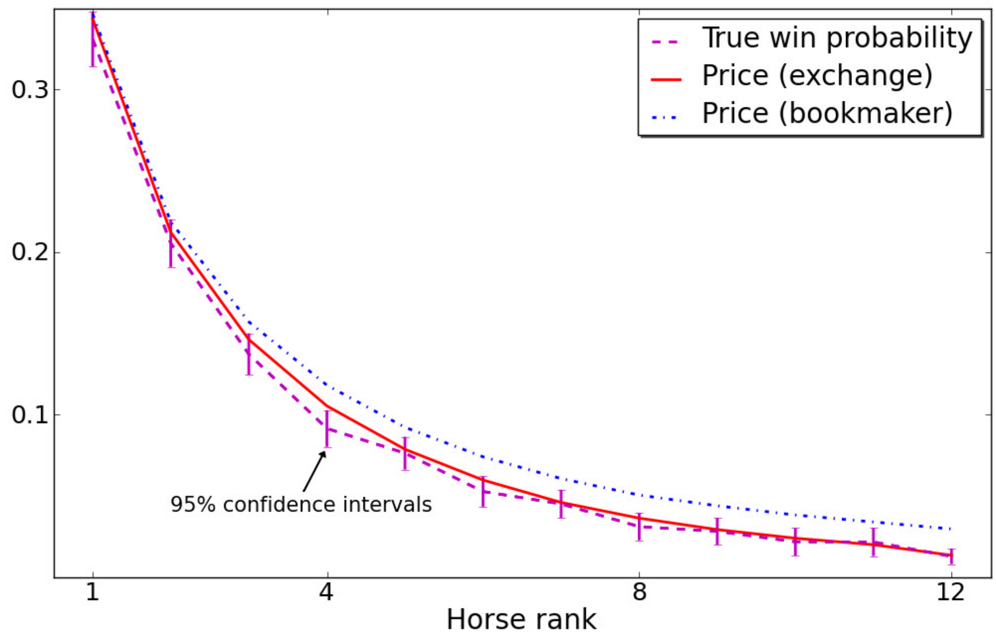
Figure 3.3. Level of FLB over time in exchange and bookmaker markets.



We find that bookmaker over-rounds are much greater than those on the exchanges (even after allowing for commission in exchange markets: see Tables 3.3 and 3.4), with bookmaker/exchange over-rounds ranging between 0.180/0.057 and 0.217/0.113 at different stages of the market. These results confirm that bettors face significantly higher costs in bookmaker (cf. exchange) markets. In addition, we find a strong positive correlation between the exchange FLB and over-round (corr. = 0.92, p = 0.0002), but no such clear relationship in bookmaker markets (corr. = 0.39, p = 0.1474). In fact, the exchange over-round decreases through the duration of the market, along with the level of FLB, whereas it remains fairly stable in bookmaker markets.

Finally, we compare starting prices in exchange and bookmaker markets for horses with greater and smaller chances of success (as predicted by their prices). In particular, the favourite is ranked 1, the second favourite is ranked 2, and so on. Horses which are ranked 12th or more are grouped together. The results of comparing exchange and bookmaker prices are presented in Table 3.5 and illustrated in Figure 3.4.

Figure 3.4. True winning probability and prices by horse rank.



We find that there is no significant difference in exchange and bookmaker prices for the first three ranks, but for horses ranked 4th favourite or greater, the exchange prices are significantly lower, representing better value for the bettor (e.g., for rank 4: z = 2.22, p = 0.0264). In sum, our results confirm hypotheses 2 and 4 that the FLB is eliminated in the later stages of the market for the exchange, but that the FLB in the bookmaker market is present throughout. Furthermore, we find support for hypothesis 5, that bookmaker and exchange prices for the favourites are similar, but bookmaker prices on longshots are significantly higher than exchange prices.

Table 3.5. Comparison of true winning probabilities and mean starting prices by horse rank in exchange and bookmaker markets.

Horse rank by exchange price	Number of horses	True winning prob. p	Exchange price r	Bookmaker price q	SE (r,q)	$\frac{r-q}{SE(r,q)}$
1	6058	0.3310	0.3434	0.3463	0.0086	0.34
2	6058	0.2053	0.2121	0.2184	0.0075	0.86
3	6055	0.1371	0.1461	0.1569	0.0065	1.67
4	6012	0.0913	0.1052	0.1180	0.0057	2.22*
5	5857	0.0761	0.0784	0.0920	0.0052	2.62**
6	5522	0.0525	0.0596	0.0739	0.0047	3.01**
7	5074	0.0449	0.0458	0.0605	0.0045	3.30**
8	4496	0.0309	0.0361	0.0504	0.0043	3.32**
9	3886	0.0280	0.0291	0.0437	0.0042	3.45**
10	3241	0.0216	0.0238	0.0381	0.0043	3.33**
11	2656	0.0215	0.0198	0.0338	0.0044	3.16**
12 or more	7209	0.0126	0.0133	0.0295	0.0024	6.74**

*: significantly different from 1.00 at the 5% level, **: 1% level (2-tailed test).

3.6. Discussion

Our first two hypotheses, related to exchange prices, were that they should approach the true winning probabilities over time, and that the FLB would be eliminated over time. We argued that this would occur because an increasing betting volume from more informed bettors is likely as the market develops, because of a reduction in spreads as liquidity increases (increased liquidity reduces the divergence of opinions allowing lower bid-ask spreads) and a reduction in uncertainty as new information is revealed. We found that exchange prices develop in the manner predicted. However, whilst the correlation between exchange betting volume (i.e., of matched bets) and accuracy is high, the increase in accuracy of exchange prices in the later stages of the market is not statistically significant, which suggests that at least some informed betting occurs early. We also found that there was a significant FLB in the early stages of the exchange market. This suggests that early-stage bettors are generally uninformed and may, as suggested in earlier studies, create the FLB because of their risk-loving tendencies (e.g., Weitzman, 1965; Quandt, 1986; Hamid et al., 1996) or biased expectations (Henery, 1985; Chadha and Quandt, 1996; Snowberg and Wolfers, 2010). However, we learn very little about exchange bettors as a population from the early-stage group because so little of the total betting volume is matched at this time (on average, only 2.9% of final volume is matched 4 hours before the race start). We find that the FLB in the exchanges is eliminated over time, suggesting that the exchanges are dominated by informed bettors who bet in a manner which eliminates any FLB. Clearly, there may be some bettors who display risk-loving attitudes or exhibit biased expectations but either their bets are matched by more informed bettors (taking advantage of the lower over-rounds in exchanges), or such bettors may bet with bookmakers.

Our third hypothesis was that bookmaker prices would approach the true probabilities over time, but not as quickly as exchange prices. Our finding that bookmaker prices are not as accurate a predictor of the race winner as exchange prices corroborates the findings of previous studies (Smith et al., 2009; Franck et al., 2010). Our results lead us to agree with their conclusions that this is an indication of the type of bettors that bet in these markets, and not a reflection on the bookmakers themselves. In particular, the higher over-rounds in bookmaker markets result in less competitive prices, driving informed betting towards the exchanges, and it is the bets of these informed traders which, in turn, lead to the improved accuracy in exchange prices.

Our finding, in support of hypothesis 4, that the FLB is present in final bookmaker prices, is not controversial; many previous studies have found a greater FLB in bookmaker markets than in parallel pari-mutuel markets (Bruce and Johnson, 2000; Peirson and Blackburn, 2003) and exchange markets (Smith et al., 2006). However, we discovered that the FLB is present in bookmaker prices *at all times* throughout the market, and that the level of FLB is not correlated with over-round, neither is there any trend in the level of FLB over the duration of the market. We are not aware of any study which has investigated the level of FLB over the duration of the market, and our finding that there is no correlation between over-round and the level of FLB contrasts with the view that higher transaction costs are the cause of a greater FLB in bookmaker markets (Smith et al., 2006). We also found evidence in support of our fifth hypothesis that prices of favourites are similar in the exchange and bookmaker markets (for horses ranked 1-3rd favourite).

Taken together, our findings support our argument that the FLB in bookmaker markets is largely the result of the bookmakers' pricing policy. Specifically, it appears that the FLB in these markets results from bookmakers deciding to price in this manner in the presence of specific conditions related to bettors' demand (that bettors' demand is more elastic for favourites than longshots). This represents both a supply- and a demand-side explanation for the phenomenon. We now consider three previous explanations for the FLB in turn, and discuss how our theory is consistent with these explanations, or at least how these explanations may be adjusted to be consistent with our theory. We label these explanations: demand-side explanations, the Shin pricing policy, and transaction and information cost explanations.

Demand-side explanations, as discussed above, generally focus on one of two sources for the bias: bettors are risk-loving (e.g., Weitzman, 1965), or bettors have biased expectations (e.g., Henery, 1985). In both cases, it is believed that the FLB directly results from bettors' decisions. For example, using an extremely large dataset of North American races Snowberg and Wolfers (2010) found that the mispricing of exotic bets was more consistent with the FLB being driven by biased expectations than risk-love. Their investigation was restricted solely to pari-mutuel market odds (which are the monopoly market in the USA). It could also be the case in bookmaker markets that bettors' preferences are consistent with them being risk-loving or subject to biased expectations. However, we found that an alternative explanation based on bookmakers' optimal pricing policy was more satisfactory: demand-side explanations do not adequately explain why there is a FLB in bookmaker prices at the outset of the market, when there has been little or no betting volume, and it does not explain why this FLB reduces through time. The current study has investigated the two major types of betting market in the UK (i.e., bookmakers and exchanges), neither of which are pari-mutuel, and found significantly different results, which are more consistent with supply-side factors being the cause of the FLB in bookmaker markets, and informed betting eliminating the FLB in exchange markets.

Shin's (1991, 1992, 1993) bookmaker pricing policy model suggests that bookmakers deliberately increase prices (or reduce odds) on longshots to protect their interests against insider traders. In Shin's model, as in later studies by Schnytzer and Shilony (2005) and Peirson and Smith (2010), insiders are defined as being more informed than anyone else, including the bookmaker. Shin assumes that insider trading is likely to be more associated with longshots than with favourites, and, consequently, he argues that bookmakers stand to lose more if they allow prices on longshots to be 'fair'. Rather, we have suggested that bookmakers increase prices on longshots in order to allow them to offer competitive (lower) prices on favourites, to address the demand preferences of bettors. Shin's model makes a number of simplifying assumptions (e.g., insider traders are perfectly informed traders) which we do not make in our model of bookmaker competition. In particular, the Shin model and other similar models rely heavily on the idea that there are insider traders operating in the market, and that these are feared by the bookmaker, resulting in their pricing policy. However, this view does not account for the reality of the betting market. In particular, bookmakers are able to refuse business or restrict betting from bettors who they believe to be insiders, and bookmakers have extensive intelligence systems which reduces the chance of 'unknown' insiders damaging their profits. In addition, we have argued that insider traders are far more likely to bet with the exchanges, particularly because of lower transaction costs. By contrast, our belief is that bookmakers are far more concerned with competition amongst themselves and with exchanges than with the activities of insiders. In particular, they are more interested in the average customer in a competitive market who is simply concerned with getting a competitive price. It is these who provide bookmakers with the majority of their betting volume. Under this assumption, our model demonstrates that there is an incentive for bookmakers to set their prices incorporating the FLB if bettors' demand is more elastic for favourites than for longshots, particularly if the bookmaker is competing with operators (e.g., exchanges) who are able to offer much lower transaction costs.

The transaction and information costs explanation for the FLB suggests that this phenomenon results from an increase in the cost of obtaining information or placing bets. This is consistent with some of our results. In particular, we found that the FLB was significantly higher in the bookmaker market, where transaction costs are significantly higher (and the cost of obtaining information does not differ for those betting with an exchange or with a bookmaker). Furthermore, we found that the level of FLB in the exchange was higher when the over-round was higher (i.e., in the early stages of the market). However, we found that the FLB in the bookmaker market was not significantly correlated with over-round. We believe that informed traders eliminate the FLB in both exchange and bookmaker markets at any time when over-rounds are not set at a level which prohibits betting at an acceptable price. However, the high over-rounds in the bookmaker markets deter informed traders, resulting in the FLB remaining throughout the duration of the market. Similarly, the FLB in the early stages of the exchange market is not eliminated because spreads are higher (as a result of lower liquidity) and there is a greater degree of uncertainty about future developments. Consequently, our results suggest that transaction costs are a factor which influences the amount of informed trading in the market which, alongside competition, affects the level of FLB.

3.7. Conclusion

Previous research has shown that the FLB has existed in a variety of jurisdictions over many decades. Whilst many studies have identified FLB in the UK bookmaker market, there is little evidence in the growing literature on betting exchanges that these markets also exhibit the FLB. Our study aimed to provide an explanation for this contrasting evidence. Consequently, we developed a model to explain the FLB in competing bookmaker markets and tested resulting hypotheses related to how the predictive accuracy, over-round, and the FLB develop over the duration of the betting market for a typical race. The model's predictions were confirmed using empirical data from the UK horserace betting market.

This study makes a number of important contributions. First, we have found further evidence of (i) FLB in bookmaker markets, and (ii) no FLB in exchange markets. Second, we confirm that predictive accuracy of exchange prices is largely superior to that of bookmaker prices. However, in the case of exchange markets, we have also uncovered significant relationships between the FLB and betting volume (or time remaining before race start, each of which is related to the level of informed trading), and between the FLB and over-round using a unique dataset consisting of matched bookmaker and exchange odds on each runner throughout the course of the market on each race. We also discovered that there were no such relationships in bookmaker markets. More importantly, we have developed a model which suggests that the optimal pricing policy for a bookmaker, who competes with other operators for betting on favourites, is to set prices which include the FLB. Our empirical results are largely supportive of the predictions of this model.

We have set our explanations for the FLB within the wider context of the ongoing debate about the cause of FLB in betting markets, and have shown that both supply- and demand-side explanations are important contributors to the bias. Transaction and information costs explanations are still relevant, but only in the sense that higher costs restrict informed betting, which, in turn, prevents the FLB being eliminated. Furthermore, we demonstrate that bookmakers' optimal pricing policy, arising as a consequence of competition between operators and bettors' demand for competitive prices on favourites, is an important contributor to the existence of the phenomenon.

4. New evidence for a prospect theory explanation of systematic decision making bias in a market for state contingent claims

Abstract

The favourite-longshot bias (FLB) is the widely-reported systematic bias in markets for state contingent claims, such as prediction and betting markets, whereby market prices under-/over-value favourites/longshots. We provide new and unique evidence to support the view that, where the bias exists independently of a market maker (e.g., in pari-mutuel betting markets), it is due to cognitive errors of traders rather than their preference for risk. This is achieved in two stages: first, we derive a model that shows that prospect theory, and not risk-love, predicts a 'strong favourite' effect, where the level of the FLB is reduced in events where the variance of odds for non-favourites is low. Then we test the predictions of the model by employing pari-mutuel market price data related to 2447 UK horseraces. An analysis using the conditional logit model verifies that the extent of the FLB is indeed reduced in races with strong favourites, as well as in handicap races. Furthermore, unlike previous attempts to confirm that traders' cognitive errors are the source of the FLB, our results are independent of parametric assumptions or assumptions about the choice set of the decision maker.

4.1. Introduction

The favourite-longshot bias (FLB) is the systematic bias reported in markets for state contingent claims, such as prediction and betting markets, whereby market prices are such that high-probability outcomes (favourites) are under-valued and low-probability outcomes (longshots) are over-valued. First discovered in a laboratory setting by psychologists Preston and Baratta (1948), and in the naturalistic setting of betting markets for horseracing by Griffith (1949), the FLB has been shown to be present in many jurisdictions and throughout many decades (e.g., Dowie, 1976; Ali, 1977; Snowberg and Wolfers, 2010), with a few studies finding contrasting evidence (e.g., Busche and Hall, 1988; Busche, 1994). In the context of markets for state contingent claims, the presence of the bias, and its absence in some settings, has been attributed to a variety of causes including, among others, the risk-loving nature of traders (Weitzman, 1965), errors in the estimation of probabilities (Henery, 1985), the pricing policies of bookmakers (Shin, 1993), and limited information of traders (Sobel and Raines, 2003). However, it is empirically difficult to discriminate between the various competing explanations because the decision making processes of individual market participants are not observable. Rather, the market prices result from the combined decisions of traders, and represent the market's subjective assessment of the probability of each outcome occurring; the FLB is observed when comparing these subjective probabilities with observed event outcomes (Griffith, 1949). Hence, competing theories of decision making in general (expected utility theory and non-expected utility models such as prospect theory) are observationally equivalent (Snowberg and Wolfers, 2010).

As a result of this difficulty, there is no standard method for assessing the relative strengths of the various hypotheses, and so there have been few attempts to do so. Those that have attempted this task have relied on parametric assumptions or assumptions relating to the choice set of the decision maker. In this study, we develop a new methodology for choosing between the hypotheses, which does not rely on these assumptions.

We first develop a model based on the representative agent that predicts a 'strong favourite' effect on the level of FLB. Specifically, we demonstrate that the level of bias in an individual event varies in a predictable manner depending on the traders' risk preferences. If the representative agent is risk-seeking, the model predicts an increased FLB in events where the variance of the odds on competitors other than the favourite is relatively low, *ceteris paribus*. Convesely, if the representative agent is risk-averse, the level of FLB is reduced or a reverse bias is predicted when the same variable is relatively low. This prediction is independent of whether probabilities enter the decision process linearly (as in expected utility) or nonlinearly (as in prospect theory). Hence, empirical tests can be conducted that distinguish between hypotheses that do and do not require the representative agent to be risk-loving. We also show that, in events where the rules are designed to equalize the competitors' winning chances (e.g., in handicap horse races), the model predicts an equivalent result to the strong favourite effect, i.e., an increased/reduced FLB when the representative agent is risk-seeking/averse.

We test both predictions of the model using a large set of data from betting markets for UK horseraces, and find strong evidence to support the hypothesis that in markets independent of a market maker (e.g., a bookmaker), the FLB is caused by the cognitive errors of traders, rather than a general preference for risk. Our paper contributes new and unique evidence to the existing literature that examines the relative merits of prospect theory, expected utility theory, and other hypotheses in explaining biases in naturalistic decision making contexts.

The paper proceeds as follows. In section 4.2, we review representative agent models of the FLB at the level of demand-side factors, i.e., explanations of the FLB based only on factors related to the decisions of traders. In section 4.3, we outline our model, and in section 4.4 we derive hypotheses and introduce the data and the method employed to test the hypotheses. The results are presented in section 4.5 and discussed in section 4.6. We draw conclusions in Section 4.7.

4.2. Representative agent models of the FLB in horserace betting markets

It is widely recognized that horserace betting markets offer a valuable naturalistic setting in which to explore decision making (Sauer, 1998). In particular, horserace bettors operate in a setting that involves uncertain and dynamic information, time stress, regular outcome feedback, and meaningful incentives. These features are only present in real world decision contexts (Orasanu and Connolly, 1993) and risk-taking in the high stakes betting context is not easily reproduced in comfortable laboratory settings (Anderson and Brown, 1984). Consequently, horserace betting markets appear to offer an ideal environment in which to explore decision making biases in a real world setting (Bruce and Johnson, 2003).

The FLB is one such bias, and has received much attention in the literature (for reviews, see Vaughan Williams, 1999; Ottaviani and Sørenson, 2008; Jullien and Salanié, 2008). A range of explanations for the phenomenon have been proposed. While some studies have proposed explanations related to supply-side factors, such as the pricing polices of bookmakers (Shin, 1993), most studies associate the FLB with the decisions of bettors. A simplified but useful categorization of these explanations is provided by Snowberg and Wolfers (2010, pp. 724-725): "each yields implications for the prices of gambles equivalent to stark models of either a risk-loving representative agent or a representative agent who bases her decisions on biased perceptions of true probabilities". So, bettors have unbiased expectations, but are risk-loving (e.g., Weitzman, 1965), or have biased expectations, but are risk-neutral or risk-averse (e.g., Henery, 1985). This categorization warrants attention because it addresses the relative merits of competing theories for explaining decision making in wider fields: specifically, expected utility (EU) theory and non-expected utility models such as prospect theory.

The former class of models originates from the proposition that, in order to explain the FLB, the representative agent must be risk-loving over the relevant part of the decision making domain. So, Weitzman (1965) introduced the 'representative bettor' Mr. Avmart, who represents the 'social average' of all bettors. Instead of concentrating on individuals, Weitzman inferred the preferences of the most typical bettor from the population in order to construct Mr. Avmart's utility of wealth curve (the mathematical representation of preferences over various monetary outcomes and the basis of EU theory). He found that the FLB in his data was best explained by a convex utility of wealth curve, indicating that the average bettor is locally risk-loving (i.e., the average bettor prefers the riskier, low probability outcomes). Quandt (1986) extended the analysis by showing that the bias is the natural result of equilibrium in a market where the average bettor is risk-loving. This theory was confirmed empirically by the EU models of Ali (1977) and Hamid, Prakash, and Smyser (1996). More generally, for the FLB to be explained with reference to the bettor's utility of wealth function, the bettor need not be monotically risk-loving over the whole decision making domain. Indeed, it is possible for bettors to be risk-averse in general, but with a preference for skewness of returns (Golec and Tamarkin, 1998; Walls and Busche, 2003). Alternatively, bettors may be risk-averse over some parts of the domain and risk-seeking over others (Cain and Peel, 2004). However, these alternatives are all equivalent to the risk-loving representative agent model (Snowberg and Wolfers, 2010).

A broad alternative classification of explanations for the FLB stems from Kahneman and Tversky's (1979) prospect theory (PT), in which the value of an outcome is defined relative to a reference point and bettors are risk-averse for gains and risk-loving for losses. Crucially, PT can explain some violations of EU theory, such as the Allais (1953) paradox, as well as explaining the FLB, because objective probabilities are transformed into subjective decision weights that allow for biases in the estimation of probabilities. If the assumptions of PT hold, then the FLB can be explained solely with reference to bettors' systematic misestimation of probabilities, i.e., bettors need not be locally risk-loving. While PT is a formal model of decision making under uncertainty, this alternative class of explanations can effectively include any model in which probabilities enter the decision objective function nonlinearly in order to explain the FLB. For instance, Sobel and Raines (2003) derived an alternative specification to the risk-love model that allows for the representative bettor to be risk-neutral. In this specification, the FLB can be explained by limited information of bettors, since either underreaction to new information or limited precision in the decision making process (or a combination of the two) results in relatively large/small probabilities being under-/over-estimated. Henery (1985) argues that bettors systematically discount a constant proportion of losses; since longshots lose more often, this leads bettors to over-estimate the winning chances of longshots. Effectively, the above explanations can be incorporated into the biased expectations class of models (Snowberg and Wolfers, 2010) because in each case it is the distortion of probabilities that explains the FLB, and no restriction must be made on the risk preferences of the representative bettor.

The literature discussed above suggests that there are two broadly competing sets of theories regarding the explanations for the FLB in terms of the representative bettor: bettors are unbiased in their estimation of probabilities, but risk-loving, or they are risk-neutral or risk-averse, but biased in their estimation of probabilities. These two explanations appear to be empirically indistinguishable, because it is not immediately apparent how to employ data related to actual decisions to differentiate between risk preferences and biased estimation of probabilities (Yaari, 1965). However, some researchers have employed innovative methods for doing so. In particular, Golec and Tamarkin (1995) attempted to test the two competing hypotheses using data related to alternative bets offered by bookmakers in addition to standard 'win' bets (bets that a horse will finish in first place). In this instance, these alternative bets were so-called 'teaser' bets on outcomes which were more likely to pay off than 'win' bets, but had a corresponding lower return. Since 'teaser' bets are relatively low-risk compared to 'win' bets, risk-love would predict that bettors demand an extra return to compensate them for the low risk. However, they found returns from the side bets were relatively unfair, which is a result that risk-love cannot explain. Instead, they suggested that overconfidence (which is consistent with bettors overestimating small probabilities) better explains the FLB. Jullien and Salanié (2000) and Bradley (2003) also offered support for the view that PT (cf. EU) better explains the FLB, although they relied on parametric assumptions about the functional forms of the utility, value, and probability weighting functions of the representative agent.

More recently, Snowberg and Wolfers (2010) set out to test the competing theories using a novel approach and a large dataset of all the horseraces run in North America from 1992 to 2001 (over 865,000 races). They first estimated the parameters of the two models (the EU model and the PT model) by fitting the models to standard 'win' bets. They then examined compound exotic bets, such as the 'exacta', a bet that two horses will finish a race in first and second place in a specific order. Snowberg and Wolfers reasoned that bettors would bet in the same manner in the exotic and win betting pools, so the same models should apply for each bet type. Accordingly, they used model predictions based on win bets to forecast expected market prices in the exotic betting pools. They found that the model based on misestimation of probabilities predicted exotic bet prices more accurately than the risk-love model. Snowberg and Wolfers concluded that, with respect to the representative bettor, PT explained the FLB more effectively than EU theory.

While the general consensus from the literature has been that non-expected utility (cf. EU) models better explain decision making in the naturalistic context of horserace betting, each study has conducted different empirical tests. Indeed, despite the availability of rich and detailed quantitative data on the decisions of traders in this context, the key difficulty in distinguishing between risk-love and biased expectations is a methodological one. Hence, there have been relatively few attempts to arrive at an empirical solution to the problem, and those that have been made have relied either on parametric assumptions (Jullien and Salanié, 2000; Bradley, 2003) or on the assumption that bettors' decision making models are identical over different choice sets (Golec and Tamarkin, 1995; Snowberg and Wolfers, 2010). The purpose of this study, therefore, is to address this problem with a new and alternative methodology that does not rely on these assumptions. In the next section, we develop a model that leads to a new empirical test between the alternative hypotheses, based on the risk preferences of the representative agent, independent of the probability weighting function.

4.3. Representative agent models: the 'strong favourite' effect

Here we demonstrate that two simple models, based on the representative agent, result in alternative predictions of a 'strong favourite' effect. Specifically, they predict that the level of FLB is higher or lower in races where, *ceteris paribus*, the variance of odds for non-favourites is relatively low. The first model is based on PT, in which the representative agent is risk-averse for gains: this model predicts that the FLB will be decreased in races with the 'strong favourite' condition. Conversely, our second model is based on EU and so the representative agent must be risk-loving in order to explain the FLB. This model predicts that the FLB will be increased in races with the strong favourite condition.

Following a number of FLB studies (e.g., Ali, 1977; Jullien and Salanié, 2000; Bradley, 2003; Snowberg and Wolfers, 2010), we make the assumption of the representative agent, i.e., we assume that: (i) each bettor is identical and bets an equal amount x under any circumstances, (ii) bettors bet on at most one horse in a race (with n runners), (iii) each horse is bet by at least one bettor, and (iv) each bettor has identical beliefs that each horse i will win the race with probability p_i, where $\sum_{k=1}^{n} p_k = 1$, and these beliefs are unbiased, i.e., each horse i does win with probability p_i.

The above conditions give rise to an equilibrium condition: bettors must be indifferent between all horses in the race, i.e., denoting the desirability of horse i by D_i, we must have that $D_i = D_j$ for all i, j (Ali, 1977). Otherwise, there would be at least one pair of horses such that $D_i > D_j$ and no one would bet on horse j (contravening the assumption that each horse is bet on by at least one bettor). With two mild assumptions (continuity of the D_i and first-order stochastic dominance), the equilibrium condition is unique (Jullien and Salanié, 2000).

The intuition in the following models is that, when equilibrium prices are biased due to bettors being subject to EU theory or PT, the level of bias in prices changes depending on the distribution of prices across all of the horses in the race. In our PT model, each bettor has value function $v(t)$, which is concave for gains and convex for losses, and their decisions are based on a weighted probability $w(p_i)$ (for simplicity, we make the assumption that, as in the original PT, the weighting function is the same for gains and losses). Each horse i has pari-mutuel 'win' odds R_i, i.e., a bet of size x on horse i returns a profit of xR_i if horse i wins and a loss of $-x$ otherwise. Thus, the value to a bettor of a winning bet of size x on horse i is given by $v(xR_i)$ and the value of a losing bet is given by $v(-x)$. Then the desirability of a bet of size x on horse i is given by the expected value of profits, where the associated probabilities are weighted by $w(p_i)$, i.e.,

(4.1) $$D_i = w(p_i)v(xR_i) + \sum_{j=1, j \neq i}^{n} w(p_j)v(-x).$$

Empirical estimates have found that the weighting functions typically sum to an amount less than one (e.g., Kahneman and Tversky, 1979), so we suppose that

(4.2) $$\sum_{k=1}^{n} w(p_k) = c < 1,$$

although the exact amount c is not important in our model. Hence (4.1) becomes

(4.3) $$D_i = w(p_i)v(xR_i) + (c - w(p_i))v(-x).$$

Solving the equlibrium condition $D_i = A$, where A is a constant, for $w(p_i)$ gives

(4.4) $$w(p_i) = \frac{A - cv(-x)}{v(xR_i) - v(-x)}.$$

Then, condition (4.2) gives

(4.5) $$A = cv(-x) + \frac{c}{\sum_{k=1}^{n} \frac{1}{v(xR_i) - v(-x)}}.$$

Substituting this back into (4.4) gives

(4.6) $$w(p_i) = \frac{1}{v(xR_i) - v(-x)} \frac{c}{\sum_{k=1}^{n} \frac{1}{v(xR_k) - v(-x)}}.$$

In our EU model, each bettor has an increasing utility of wealth function $u(t)$, has rational expectations, and current wealth *a*. So, the utility to a bettor of a winning bet of size *x* on horse *i* with odds R_i is given by $u(a + xR_i)$ and the utility of a losing bet is given by $u(a - x)$. Then the desirability of such a bet is given by the expected utility of future wealth, i.e.,

(4.7) $$D_i = p_i u(a + xR_i) + (1 - p_i)u(a - x).$$

This model leads to an equivalent expression to that in (4.6), which is

(4.8) $$p_i = \frac{1}{u(a + xR_i) - u(a - x)} \frac{1}{\sum_{k=1}^{n} \frac{1}{u(a + xR_k) - u(a - x)}}.$$

This equation is also Jullien and Salanié's (2000) explicit formula for the probabilities in terms of the odds for an EU model. Equation (4.6) is our PT equivalent. We extend their rationale and show that the PT model not only explains the FLB but also predicts a reduced or reverse FLB in races where the variance of odds for non-favourite is low, i.e., races with strong favourites and handicaps. On the other hand, for the EU model to explain the FLB, we require that $u(t)$ is convex. In this case, the model predicts an increased FLB in strong favourite races or handicaps.

In order to achieve the above, we index the favourite by 1, and note that we can rearrange (4.6) and (4.8) to give

(4.9)
$$w(p_1) = \frac{c}{1 + [v(xR_1) - v(-x)]X_v},$$

$$p_1 = \frac{1}{1 + [u(a + xR_1) - u(a - x)]X_u},$$

where

(4.10)
$$X_v = \sum_{k=2}^{n} \frac{1}{v(xR_k) - v(-x)},$$

$$X_u = \sum_{k=2}^{n} \frac{1}{u(a + xR_k) - u(a - x)}.$$

Now, X_v is proportional to the sum of the prices on each non-favourite (indexed by k = 2, 3, … , n), after each price has been modified by the value function for gains, where the prices are given by $r_k = 1/R_k$. To see this, note that, since $v(-x)$ is independent of k, we can set $v(-x) = 0$. Then, representing the value function for gains by a concave power function $v(t) = t^{\alpha}$, $0 < \alpha < 1$, we have that

(4.11)
$$X_v = x^{-\alpha} \sum_{k=2}^{n} r_k^{\alpha}.$$

So X_v varies in a predictable way depending on the variance of the odds for non-favourites, and hence so does the equilibrium condition in (4.6); X_u has a similar interpretation.

Recall two conditions on the value function $v(t)$ in PT: first, it is an increasing function, and second, it is concave for gains. Hence, the $v(xR_k)$ are concave and increasing. These conditions ensure that X_v is decreasing in the variance of the odds for non-favourites (see Appendix 3), i.e., X_v is greater in the 'strong favourite' condition (associated with low variance of odds for non-favourites). Thus, since (4.6) is an equilibrium condition, if we increase X_v, *ceteris paribus*, we must also decrease $v(xR_1)$. Consequently, since $v(t)$ is increasing, we must decrease R_1, which is the odds for the favourite. Hence, this model based on PT predicts that the FLB will be reduced in races with lower variance of odds for non-favourites. Note that this result is unique to the PT model; the EU model would only predict the FLB if the representative bettor's utility function is convex, i.e., if the representative bettor is risk-loving. However, in this case, the EU model predicts the opposite effect to the PT model: since the $u(a + xR_k)$ are convex and increasing, X_u is lower in the strong favourite condition. Thus, if we decrease X_u, *ceteris paribus*, we must also increase R_1, i.e., there is an increased FLB. Crucially, these differing predictions are independent of whether probabilities are weighted or not. So, while the EU model predicts an increased FLB in the strong favourite condition because the utility function must be convex, under PT, bettors can have a concave value function for gains, because it is the probability weighting function that explains the FLB.

The model we have developed has a further testable implication. Around half of races in the UK are 'handicap' races, where horses are allocated differential weights to carry (based on their previous performances), in an effort to equalize the winning chances of all horses in the race. The result of this equalizing procedure is that handicap races have lower variance of odds over all runners than non-handicap races. Hence, we might also expect the variance of odds for non-favourites to be lower. Testing this prediction using our data, we find that this is the case; the standard deviations of odds for non-favourites being 15.6 and 22.8 in handicaps and non-handicaps, respectively. Hence, in handicap races (at least for our dataset), we should expect the FLB to be reduced. We now develop hypotheses related to the implications of the model.

4.4. Hypotheses, data, and method

4.4.1. Hypotheses

Our models, based on PT and EU theory, respectively, predict that the FLB will be reduced/increased in races with low variance of odds for non-favourites (which we call the 'strong favourite' effect). Although previous studies have made assumptions that we do not make in this paper, the weight of evidence from these studies is in favour of PT as an explanation for the FLB. Consequently, we test the following as a 'strong favourite' hypothesis:

1. The level of the FLB will be reduced in races where the variance of odds for non-favourites is lower.

Our PT and EU models predict that the level of FLB will be reduced/increased, respectively, when the variance of the odds for non-favourites is relatively low. In handicap races, the variance of odds over all runners is generally lower because weights are distributed to horses in an effort to equalize horses' winning chances. We therefore might also expect the variance of odds for non-favourites to be lower, and this is empirically the case in our dataset. The weight of evidence from previous studies is in favour of the PT explanation; consequently, we test the following hypothesis:

2. The level of the FLB will be reduced in handicap (cf. non-handicap) races.

4.4.2. Data

The data employed in this study are final pari-mutuel odds and finishing positions for 25,644 horses running in 2447 races in the UK. Pari-mutuel odds are profits from a winning £1 bet on each horse, before transaction costs (track take and breakage) are deducted from the winnings. So, in order to adjust for transaction costs, odds-implied probabilities (hereafter, odds-probabilities), which are the probabilities of each horse winning the race as implied by the odds available, are given by

$$(4.12) \qquad q_{ij} = \frac{\frac{1}{1+R_{ij}}}{\sum_{k=1}^{n_j} \frac{1}{1+R_{kj}}},$$

where R_{ij} is the pari-mutuel odds for horse i in race j and n_j is the number of horses running in race j. The number of runners in each race in the database ranges from 2 to 29, with a mode of 8.

4.4.3. Method

In order to quantify the level of the FLB in the data, we use a conditional logit (CL) modeling approach (McFadden, 1974), which has been employed in many studies of the efficiency of betting markets (e.g., Figlewski, 1979; Asch et al, 1984; Bolton and Chapman, 1986; Benter, 1994; Sung and Johnson, 2010). In the context of horseracing, the CL model estimates the probability of each horse winning that race, from variables related to the horses, while taking into account the competitive nature of the race. With log of odds-probability as the only independent variable, the CL model is an effective method for estimating the level of FLB, and has the advantage of accounting for the intensity of competition between runners in each race. It is formulated as follows.

First, define an estimate of the horse *i*'s ability to win race *j*, W_{ij}, as

(4.13) $$W_{ij} = \beta \ln q_{ij} + \varepsilon_{ij} ,$$

where β is the parameter that determines the importance of the log of odds-probability q_{ij} for horse *i* in race *j*, and ε_{ij} is an independent error term. McFadden (1974) shows that if the independent errors are identically distributed according to the double exponential distribution, then the winning probability for horse *i* in race *j* is given by

(4.14) $$p_{ij} = \Pr(W_{ij} > W_{kj}, k = 1,2,\ldots,n_j, k \neq i) = \frac{\exp(\beta \ln q_{ij})}{\sum_{k=1}^{n_j} \exp(\beta \ln q_{kj})} = \frac{q_{ij}{}^{\beta}}{\sum_{k=1}^{n_j} q_{kj}{}^{\beta}} .$$

The parameter β is estimated by maximizing the joint probability of observing the results of all the races in the sample, i.e., by maximizing the log-likelihood

(4.15) $$\ln L = \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij} ,$$

where y_{ij} = 1 if horse *i* won race *j* and y_{ij} = 0 otherwise, and *N* is the total number of races. On estimating a CL model with log of odds-probabilities as the only independent variable, we refer to its parameter as the FLB β. An estimated value of the FLB β of one implies that the odds-probabilities are, on average, equal to the true winning probabilities. A value of the FLB β greater than one indicates a standard FLB, where longshots are relatively overbet. The greater the value of β, the greater is the degree of the FLB (Bacon-Shone et al., 1992). On the other hand, a value of β less than one indicates a reverse FLB, where favourites are relatively overbet.

Before addressing our hypotheses, it is instructive to investigate to what extent the FLB is present in all the races in our dataset. To test whether the FLB is present we employ the following standard normal test statistic to test whether the FLB β value significantly exceeds or is less than one:

$$(4.16) \qquad z = \frac{\beta - 1}{S.E.(\beta)}.$$

Subsequently, in order to address the issue central to each hypothesis, we divide the dataset into races that (i) do or do not satisfy the 'strong favourite' condition, and (ii) are handicap or non-handicap races. Here, a race satisfies the 'strong favourite' condition when the standard deviation of odds for non-favourites is lower than *x*, where we choose *x* to be 5 or 10. Then a race that does not satisfy this condition has this same variable greater than *x*. To test the strong favourite hypothesis, we separately estimate β in races that satisfy the strong favourite condition and races that do not. We then compare the two β values using the standard normal test statistic

$$(4.17) \qquad z = \frac{\beta_1 - \beta_2}{S.E.(\beta_1, \beta_2)}.$$

A similar test is carried out to test the handicaps hypothesis. We note that the choices for *x* of 5 and 10 are fairly arbitrary, so we also investigate the effect of altering these choices on the level of the FLB.

4.5. Results

4.5.1. Strong favourite hypothesis

The first row of Table 4.1 shows that there is a FLB over all the races in our data, with favourites underbet and longshots overbet (β = 1.091, p = 0.0014). To test whether the decisions of horserace bettors are consistent with a reduced FLB in races with a strong favourite (races with low variance of odds for non-favourites), we estimate the levels of FLB in races with the strong favourite condition. The results of these estimations are displayed in Table 4.1. We split races into those where the standard deviation of odds for non-favourites is lower and greater than some cut-off value. When we set the cut-off value for a strong favourite race at 5, we find that the level of FLB is significantly lower in races with a strong favourite compared to races with no strong favourite (NSF: β = 1.109, SF: β = 0.820, p = 0.0061).

Table 4.1. Results of estimating conditional logit models with log of odds-implied probability as the single explanatory variable for UK races from 2004, assessing the prevalence of a strong favourite effect.

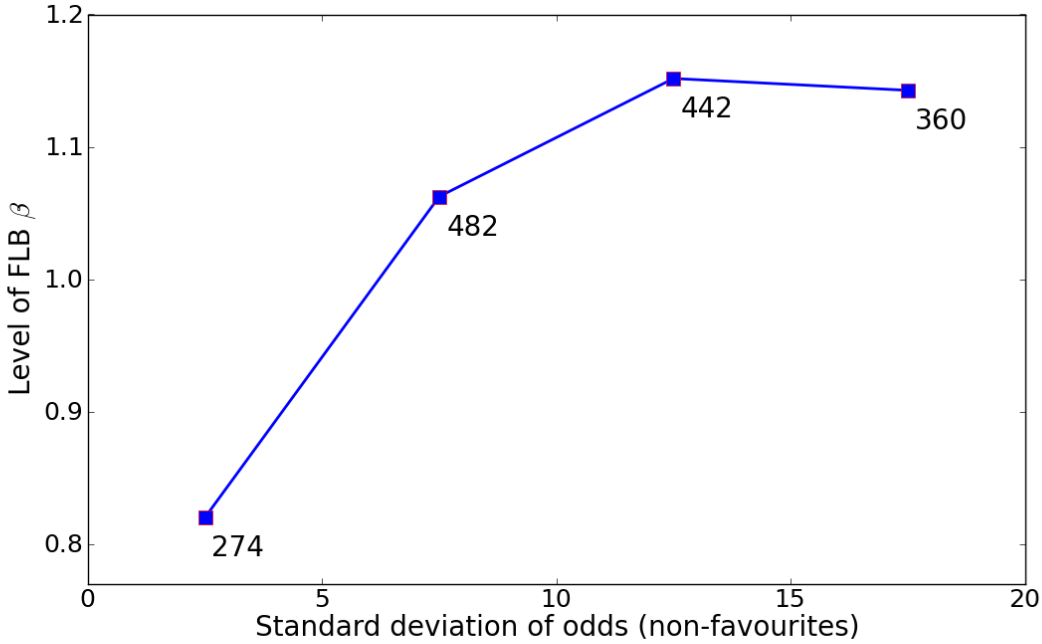
	Number of races	Mean number of runners	Level of FLB β	z (S.E.)	p
All races	2447	10.5	1.091	3.19*** (0.041)	0.0014
NSF(5)	2173	11.0	1.109	2.51*** (0.115)	0.0061
SF(5)	274	6.2	0.820		
NSF(10)	1691	11.7	1.118	1.83** (0.068)	0.0337
SF(10)	756	7.6	0.994		
Non-handicaps	1161	9.5	1.128	1.47* (0.057)	0.0703
Handicaps	1286	11.4	1.043		

$z = (\beta - 1)$/S.E. (all races), $z = [\beta(1) - \beta(2)]$/S.E. (comparisons).
'SF(x)' indicates a race where the standard deviation of odds for non-favourites (where the favourite is the horse believed by the bettors to be most likely to win the race) is lower than x.
'NSF(x)' indicates a race where the the standard deviation of odds for non-favourites is higher than x.
***: significant at the 1% level; *: significant at the 5% level; *: significant at the 10% level (1 tailed-test).

In fact, the strong favourite effect is so strong here that the level of FLB in races with a strong favourite is such that there is a reverse FLB for these races. When we set the cutoff value at 10, we also find lower FLB in races with a strong favourite than races without this condition (NSF: β = 1.118, NSF: β = 0.994, p = 0.0337). Hence, there is strong evidence to suggest that the FLB is reduced in races with a strong favourite. In order to see whether different values of the standard deviation of odds for non-favourites would give a different result, we plot the value of the FLB β for other values in Figure 4.1.

We see that when examining the races in the dataset there is general trend of increasing FLB as the standard deviation increases (i.e., for lower variance, the bias is lower). However, the effect appears to be marginally lower as the standard deviation of odds for non-favourites increases, suggesting that the strong favourite effect is most prominent for races with a relatively strong favourite (i.e., relatively low variance of odds for non-favourites).

Figure 4.1. The level of FLB with differing standard deviation of odds for non-favourites (where the favourite is the horse believed by the bettors to be most likely to win the race); number of races for each data-point are shown.



In sum, these results support the strong favourite hypothesis, i.e., they suggest that the level of the FLB is reduced in races with strong favourites. Hence, we find strong evidence to support our PT (cf. EU) model in explaining the FLB in these markets.

4.5.2. Handicaps hypothesis

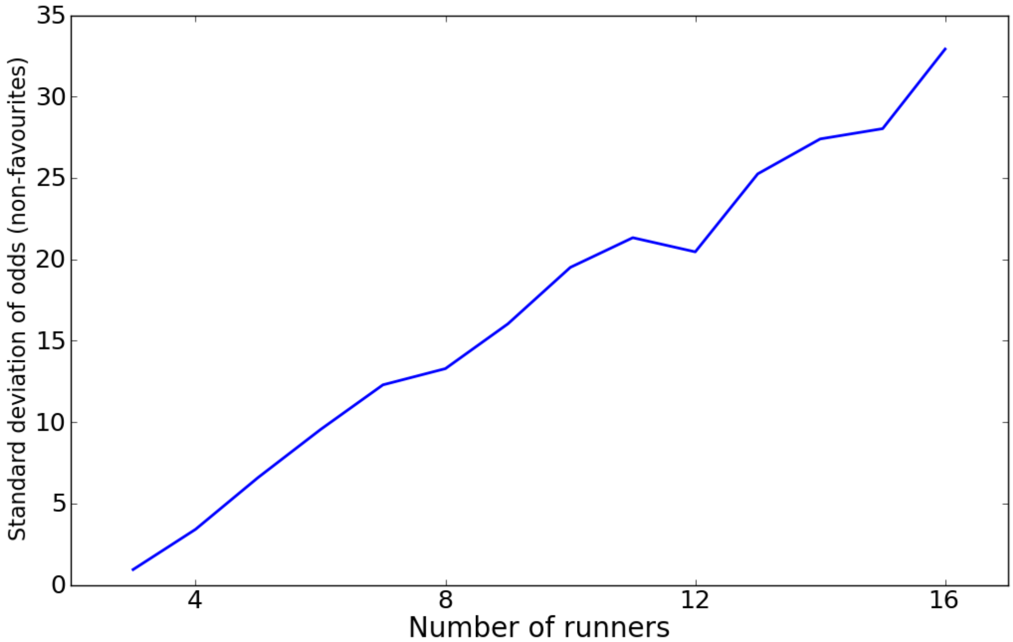
We now compare the level of FLB in handicap and non-handicap races. The results are presented in the last row of Table 4.1. We find that there is a modest reduction in of the level of FLB in handicap (cf. non-handicap) races (non-handicaps: β = 1.128, handicaps: β = 1.043, p = 0.0703). This provides some evidence to support the handicaps hypothesis, i.e., as a result of the strong favourite effect, the level of FLB is reduced in handicap races. This provides further evidence to support PT as an explanation for decision making bias under uncertainty.

4.6. Discussion

The main prediction of our model is that when the variance of odds for non-favourites is lower, favourites will be relatively under-/overbet, depending on the risk preferences of the representative bettor; specifically, there will be an increased/decreased FLB if the bettor is risk-loving/averse. In addition, our model predicts that the strong favourite effect should be replicated in handicap races, since the rules of entry for these races aim to equalize horses' chances; this in turn is likely to have the effect of reducing the variance of the odds for non-favourites. Hence, the model offers two related methods for empirically distinguishing between two competing theories of decision making under uncertainty. If decisions are made under PT, we should expect a reduced FLB in strong favourite races. Conversely, if decisions are consistent with EU theory, we should expect an increased FLB in strong favourite races. Our empirical findings support the former alternative, with stronger evidence provided by the strong favourite effect tests than the handicaps test. The slightly weaker evidence from the handicaps test could be explained by noting that, while it is to be expected that handicap races have reduced variance of odds over all competitors, this does not directly mean that variance of odds for non-favourites will be low. Instead, there are unlikely to be many strong favourites in handicap races. Thus, while we empirically see that there is a negative correlation between a race being a handicap and its variance of odds for non-favourites, we might expect the strong favourites' effect to be slightly less in evidence.

We also found that, while the strong favourite effect is prominent for races with low variance of odds for non-favourites, the effect is less significant for higher variances (Figure 4.1). This could be due to a confounding effect related to the number of runners in each race. As we show in Figure 4.2, there is a strong positive correlation between the number of runners in a race and the average standard deviation of odds for non-favourites.

Figure 4.2. The relationship between the number of runners in each race and the average standard deviation of odds for non-favourites (where the favourite is the horse believed by the bettors to be most likely to win the race); races with 3 or fewer runners, or 16 or more runners, are grouped.



Consequently, higher values of the variance are associated with greater numbers of runners, and this suggests that we may be observing a complexity effect which reduces the level of FLB. In particular, it has been shown that, as the complexity of a decision task increases (i.e., the number of alternatives in the choice set increases), decision makers opt for simplistic, compensatory strategies, including the 'take-the-best' heuristic (Gigerenzer and Goldstein, 1999), in which the decision maker prepares an order of cues based on their relative prediction validity, before choosing according to the first cue that discriminates between the alternatives. In a betting decision associated with a horserace, odds have the highest cue validity of all sources of information (they are the best predictor of race results), and we can, therefore, expect that under increased task complexity bettors are likely to adopt such heuristic strategies, leading to the relative overbetting of strong favourites in races with a higher number of runners.

Primarily, our results provide support to the hypothesis that cognitive errors (i.e., prospect theory), rather than a preference for risk, explain the FLB. Our empirically verified predictions are *only* consistent with a representative bettor model based on PT, rather than EU theory. Specifically, our results demonstrate that the representative bettor must be risk-averse for gains, and expected utility models where the average bettor is risk-averse are unable to explain the FLB. Hence, EU models are not consistent with our data. This conclusion, unlike that arrived at in other studies attempting to determine the origin of the FLB, is independent of parametric assumptions (Jullien and Salanié, 2000; Bradley, 2003) and makes no assumptions about the choice set of the decision maker (Golec and Tamarkin, 1995; Snowberg and Wolfers, 2010).

There have been a wide variety of previous explanations for the FLB: risk-love (Weitzman, 1965), the pricing policies of bookmakers (Shin, 1993), the limited information of bettors (Sobel and Raines, 2003), and misperceptions of probabilities (Henery, 1985). Our study sheds new light on this issue by showing definitively that, in markets that do not involve a market maker, prospect theory (specifically, risk-aversion with biased probabilities) is necessary to explain an observed reduction in the level of the FLB in races with strong favourites or in handicaps.

4.7. Conclusion

In sum, we find evidence that decisions in a real-world environment appear consistent with a 'strong favourite' effect, whereby bettors who generally underbet favourites give undue preference to favourites in races where the variance of the odds of their rivals is low. This effect is predicted by a representative agent model employing prospect theory as the driving force behind bettors' decisions. A similar effect is also observed when comparing the FLB in handicap and non-handicap races. The fact that our empirical results are consistent with prospect theory rather than expected utility, without parametric or choice set assumptions, contributes powerful new evidence that prospect theory (cf., expected utility theory) is better able to explain decision making biases under uncertainty in naturalistic environments. Demonstrating such biases in the real world is crucial to establishing the generalizability of laboratory research, and while the original experimental evidence supporting the conclusions of prospect theory was compelling, future work in this area could further investigate the consistency of empirical evidence of other large populations of decision makers with the predictions of theoretical models.

Finally, there have been a wide range of proposed explanations for the FLB, from risk preferences and cognitive errors to limited information of traders. However, we have demonstrated in this paper that the decisions made by a large population of traders in a naturalistic environment are only consistent with risk-averse traders with biased subjective probability estimates.

5. Properties of pseudo-R^2 as an estimate of forecast accuracy for discrete choice models

Abstract

While R^2 in ordinary least squares linear regression is a widely-used and well-justified measure, the same is typically not true of pseudo-R^2s in logistic regression. This has important implications for the evaluation of pseudo-R^2s as estimates of forecast accuracy for discrete choice models. We show both theoretically and empirically that at least one of the definitions of pseudo-R^2 is not robust to variations in the number of alternatives, and suggest an adjustment to correct for this bias. We describe and evaluate the relative merits of two methods (bootstrap and asymptotic) for estimating the variance of pseudo-R^2s so that their values can be compared across non-nested models or across models fitted on different datasets. Finally, we derive relationships that describe the usefulness of pseudo-R^2 measures in terms of their economic value in the context of competitive event prediction. As a result of the above, we arrive at a far more rigorous understanding of the value of pseudo-R^2s in evaluating the predictions of discrete choice models.

5.1. Introduction

Discrete choice models are a widely used class of statistical models, and include the multinomial logit, conditional logit, multinomial probit, mixed logit, and other models (Maddala, 1983). Their primary use is in the modelling of choice, specifically prediction of individuals' choices from a range of alternatives. Hence, they have wide applications in marketing (McFadden, 2001) and econometrics (Maddala, 1983), although they have also been adopted in such diverse fields as epidemiology (Breslow and Day, 1994), operations research (Cheng and Stough, 2006), and the forecasting of competitive events (Smith et al., 2009). While a great degree of attention has been given to the development of effective discrete choice models (Edelman, 2007), particularly for the forecasting of competitive events (e.g., Lessman et al., 2007), there has been little consideration of the best method for evaluating the predictions of these models. A key property of any means of evaluation of a forecast is its comparability across empirical models (Kvålseth 1985). Otherwise, the researcher can not be certain whether differences in the evaluator arise because of changes in the predictive power of the model or because of alternative confounding factors, such as properties of the datasets on which the models were fitted. Moreover, it is desirable in any forecasting context to assign degrees of uncertainty to any point estimates reported, in order to ensure that conclusions drawn from evaluating such statistics are statistically significant.

In linear regression, the coefficient of determination R^2 is widely used as a measure of a model's ability to explain variation in the data, and thus the accuracy of the model's predictions, and its properties and correct usage are now well understood (e.g., Kvålseth 1985, Draper and Smith, 1998). However, a similar consensus has not been reached for its analog for logistic regressions, the pseudo-R^2, because of significant differences between the two types of measure. So, while pseudo-R^2s are commonly reported, their usage is seldom justified (Veall and Zimmerman, 1996). There are in fact a number of issues associated with pseudo-R^2s that reamin unresolved: first, unlike R^2, there is no single definition of pseudo-R^2 that is universally employed. Instead, a variety of measures have been proposed, which are not necessarily mathematically (as in, the same formula) or conceptually (the same interpretation) equivalent to R^2 (Menard, 2000). Second, care must be taken when comparing values of pseudo-R^2 between datasets with different characteristics. For instance, one of the advantages of the conditional logit model in its application to discrete choice modelling is that each observation (event) need not consist of an equal number of competitors (e.g., in predicting a consumer's choice of healthcare products, the number of products available could be different for each consumer). Finally, the distributions of R^2s are complex and depend on unknown parameters (Ohtani, 2000), so while R^2 values are often reported, they are seldom accompanied by standard errors (Press and Zellner, 1978). Hence, the comparability of these measures between models is difficult. For pseudo-R^2s, this issue is exacerbated because not only are the distributional properties of pseudo-R^2s different to those of R^2, they also depend on the particular defintion of pseudo-R^2 employed *and* the choice of model. Little attention has previously been given to the consequences that these considerations have for the evaluation of discrete choice models using pseudo-R^2s.

One of the many applications for discrete choice models is in the forecasting of competitive event outcomes. A competitive event is a contest between at least two rival participants where one or more winners are declared and the outcome is uncertain: political elections or sporting events, for example. Often, these events are asssociated with markets for betting or trading on their outcome, e.g., betting markets in the case of sporting events, or prediction markets for political contests or the outcomes of business policies (Wolfers and Zitzewitz, 2006a). Since the outcomes of competitive events are of particular interest for economic reasons (in the case of sporting events) or policy implications (elections), the forecasting of competitive events is a prominent subject in the literature (e.g., Schnytzer et al., 2010; Smith and Vaughan Williams, 2010). Particular attention has been given to the properties of competitive events that mean that modelling techniques that would normally be effective in forecasting are not suitable for the forecasting of competitive events. For instance, the modeller must take into account the intensity of competition between the participants: hence, the standard modelling approach is to view competitors as alternatives in a choice set with the winner being the participant whose attributes lead it to being preferred (Lessman et al, 2011). If there is some uncertainty about the precise values of the attributes of each competitor, this is reflected in error terms, and the outputs of the model are the probability that each competitor will emerge as the preferred alternative, given the distribution of the error terms; thus, models typically involve some form of the logistic function. The use of pseudo-R^2s in competitive event forecasting is similar to any other use of discrete choice models. However, while the typical motivation for pseudo-R^2s is as a measure of improvement from null model (where each alternative is considered equally likely) to fitted model, a more useful measure would be improvement to fitted model from a model based on the predictions of prices from the associated market, a measure that we call *relative* pseudo-R^2. Then, there is a direct link between relative pseudo-R^2s and the economic value of the estimated model probabilities.

In this paper we address the various unresolved issues related to pseudo-R^2s and illustrate these points with specific reference to the forecasting of competitive events. Consequently, throughout the paper we refer to the conditional logit (CL) model, which is the most widely-used model in this context. We define two alternative pseudo-R^2 measures, and show that at least one of these is not robust to changes in the number of alternatives in each choice problem. Consequently, in order for discrete choice models to be comparable using these measures across non-nested models or across data fitted on different datasets, we suggest an adjustment to account for the bias. Then, we describe and compare two methods for obtaining the variance of pseudo-R^2 measures, the bootstrap and asymptotic methods. We find that each method results in variances that are reasonably close; hence, either method could be used to conduct significance tests for comparing pseudo-R^2 values. Finally, we define relative pseudo-R^2s as improvement to fitted model from a model based on the predictions of prices from the associated market in the context of the forecasting of competitive events. The purpose of this is to show that there is a relationship between relative pseudo-R^2s and the economic value of estimated model probabilities, a finding that has implications for assessing the efficiency of financial markets.

The remainder of this chapter is structured as follows. In section 5.2, we outline the CL model and define two alternative pseudo-R^2 measures. In section 5.3, we explore properties of these measures in relation to the number of alternatives. In section 5.4, we discuss the bootstrap and asymptotic methods. In section 5.5, we discuss relative pseudo-R^2s. We conclude in section 5.6.

5.2. The conditional logit model and pseudo-R^2s

5.2.1. The conditional logit model

The conditional logit (CL) model (McFadden, 1974) is a widely used model of discrete choice, and is particularly useful in the forecasting of the outcomes of competitive events, such as horseraces (e.g., Figlewski, 1979; Asch et al., 1984; Bolton and Chapman, 1986; Benter, 1994; Sung and Johnson, 2010). For a discrete choice problem, it results in estimates of the probabilities of each alternative being chosen based on variables relating to the alternatives. So, in the context of competitive events, it provides, for each event, estimates of the probabilities of each competitor winning the event based on variables related to the competitors, while taking into account the competition between the participants in the event. Its primary advantage over other discrete choice models in the forecasting of competitive events is that the probabilities can be expressed in an analytic form, thus estimation of the parameters is straightforward.

The formulation of the CL model begins with an estimate of the ability of competitor *i* to win event *j*, given by

(5.1) $$W_{ij} = \beta^T x_{ij} + \varepsilon_{ij},$$

where $\beta = (\beta_1, \beta_2, \ldots, \beta_m)^T$ are the coefficients that determine the importance of the variables $x_{ij} = (x_{ij1}, x_{ij2}, \ldots, x_{ijm})^T$ to the winning chances of the competitor, and ε_{ij} is an independent error term. The key assumption for the CL model, which makes the estimated probabilities analytically tractable, is that the errors are identically distributed according to the double exponential distribution

(5.2) $$f(x) = \exp[-x - \exp(-x)],$$

which is plotted in Figure 5.1 with a normal curve for comparison.

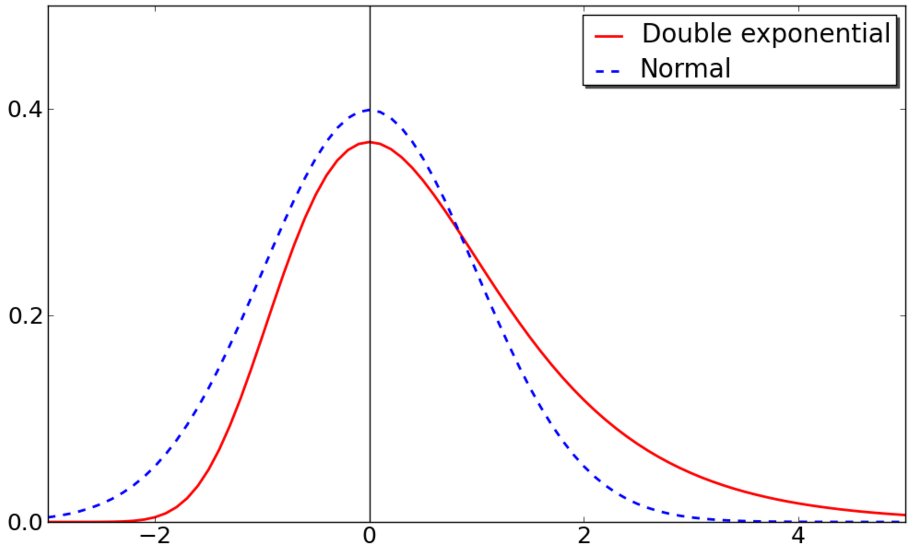
Then the estimated winning probability for competitor *i* in event *j* is given by

(5.3) $$p_{ij} = \Pr(W_{ij} > W_{kj}, k = 1, \ldots, n_j, k \neq i) = \frac{\exp(\beta^T x_{ij})}{\sum_{k=1}^{n_j} \exp(\beta^T x_{kj})}.$$

Alternatively, the probabilities can be written in the form of a logistic function as

(5.4) $$p_{ij} = \frac{1}{1 + \sum_{k=1, k \neq i}^{n_j} \exp[\beta^T (x_{kj} - x_{ij})]}.$$

Figure 5.1. The double exponential distribution.



The coefficients β are estimated by maximizing the joint probability of observing the results of all the events in the dataset, i.e., by maximizing the log-likelihood (ln$L$) of the full model (the model that includes all the independent variables in which we are interested),

(5.5) $$\ln L(\beta) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij} \, ,$$

where y_{ij} = 1 if competitor i won event j and y_{ij} = 0 otherwise, and N is the total number of events in the dataset. We denote the maximized likelihood function for the full model by $\ln L(\hat{\beta})$.

Since the estimated probabilities in the CL model are analytically tractable, it is straightforward to estimate the parameters using the Newton-Raphson method. In this case, if we let

(5.6) $$\begin{aligned} F_{ij} &= \exp(\beta^T x_{ij}) \\ S_j^0 &= \sum_{i=1}^{n_j} F_{ij} \\ S_{ju}^1 &= \sum_{i=1}^{n_j} F_{ij} x_{iju} \\ S_{juv}^2 &= \sum_{i=1}^{n_j} F_{ij} x_{iju} x_{ijv} \end{aligned} \, ,$$

then

(5.7) $$\frac{d\ln L(\beta)}{d\beta_u} = \sum_{j=1}^{N}\left[\sum_{i=1}^{n_j} y_{ij}x_{iju} - \frac{S_{ju}^1}{S_j^0}\right],$$

$$\frac{d^2\ln L(\beta)}{d\beta_u\beta_v} = \sum_{j=1}^{N}\frac{S_{ju}^1 S_{jv}^1 - S_j^0 S_{juv}^2}{(S_j^0)^2}.$$

5.2.2. Pseudo-R^2s

In linear regression, a popular measure of 'goodness-of-fit' is the coefficient of determination R^2 (Draper and Smith, 1998); indeed, this is perhaps the most widely used statistic in ordinary least squares regression (Kvålseth 1985). It varies between 0 and 1, and can be interpreted in a number of ways. First, it is the variation in the data that is explained by the model, a value of 1 implying that the model fully explains variability in the data. Second, it is the square of the correlation between the model's predicted values and the actual values. Third, it is the improvement from null model (a model that includes no independent variables) to fitted model, with 1 being a model that perfectly predicts any new datapoint.

However, the standard definition of R^2 is not applicable in logistic regression models, such as CL. Rather, several alternative pseudo-R^2 measures have been proposed. The motivating criterion for pseudo-R^2s is primarily the same as the third interpration of R^2 above, i.e., as improvement from null model to fitted model. The CL model is an example of a model that is estimated by maximum likelihood, and in fact, for *any* model that is estimated by maximum likelihood, pseudo-R^2s that satisfy this criterion can be defined.

The most popular measure (e.g., Benter, 1994) is McFadden's (1974) pseudo-R^2, which is given by

(5.8) $$R_M^2 = 1 - \frac{\ln L(\hat{\beta})}{\ln L(0)},$$

where $\ln L(\hat{\beta})$ is the maximized log-likelihood (ln$L$) of the full model, including all independent variables,

(5.9) $$\ln L(\beta) = \sum_{j=1}^{N}\sum_{i=1}^{n_j} y_{ij}\ln p_{ij}$$

and $\ln L(0)$ is the $\ln L$ of the naive model, where each competitor in the event is assigned the same probability of winning:

(5.10) $$\ln L(0) = \sum_{j=1}^{N} \ln(1/n_j).$$

An alternative is the Maddala (1983) pseudo-R^2, given by

(5.11) $$R_D^2 = 1 - \exp\{(2/N)[\ln L(0) - \ln L(\hat{\beta})]\}.$$

Note that, although the McFadden pseudo-R^2 possesses the desirable property that it takes a maximum value of 1, this is not true of the Maddala pseudo-R^2, which has a maximum of $R_D^2 = 1 - \exp\{(2/N)\ln L(0)\}$. While there are many other definitions of pseudo-R^2, for the remainder of this chapter, we consider only these two definitions, but our results are easily extended to other definitions. We begin with a consideration of their comparability across datasets.

5.3. Pseudo-R^2s and dependence on the number of alternatives

In this section, our primary motivation is to ensure that pseudo-R^2s are comparable between similarly specified models on alternative datasets. The concern here is that, in any discrete choice problem, alternative datasets may differ in their characteristics. Specifically, depending on how the data are sampled, the average number of alternatives available to each subject may vary. For example, in predicting a consumer's preferred medical care, the number of alternatives available to them might depend on their geographical location. To give an extreme example from competitive events, suppose a researcher seeks to analyze variations in the predictability of horseraces depending on the number of horses in each race. In this instance, the researcher will sample alternative datasets depending on the number of runners in each race, and so necessarily the average number of competitors in each subset of the data will differ. In this section, we investigate whether each of the McFadden and Maddala pseudo-R^2s are robust to variations in this particular characteristic of the data. We find that the value of the McFadden pseudo-R^2 varies predictably depending on the average number of alternatives, even while the predictive power of model estimates remains constant. Consequently, we define an adjusted version that does not have this undesirable property, while still satisfying the original criterion of improvement from null model to fitted model.

Recall that the motivation behind the formulation of pseudo-R^2s is that they represent the degree of improvement from null model to fitted model. Suppose, therefore, that, in each event *j*, the model assigns a winning probability to the eventual winner of $f(n_j)/n_j$, i.e., if y_{ij} = 1, $p_{ij} = f(n_j)/n_j$ (note that since the p_{ij} are probabilities, $0 \le f(n_j) \le n_j$ for all *j*). Then the dependence of the two definitions of pseudo-R^2 can be evaluated for their dependence on (or independence from) the number of competitors in each event. The formulae are given by the following proposition (for a proof, see Appendix 4).

Proposition 1. *If the model probabilities are* $p_{ij} = f(n_j)/n_j$, *then the McFadden and Maddala pseudo-R^2s are given by*

(5.12) $$R_M^2 = \frac{\ln \tilde{f}}{\ln \tilde{n}}$$

and

(5.13) $$R_D^2 = 1 - \frac{1}{\tilde{f}^2},$$

respectively, where $\tilde{n} = (\prod_{j=1}^{N} n_j)^{1/N}$ *and* $\tilde{f} = [\prod_{j=1}^{N} f(n_j)]^{1/N}$ *are the geometric means of the number of competitors and of* $f(n_j)$*, respectively. Here, N is the number of events.*

Here, $\tilde{f}$ can be thought of as the part of the pseudo-R^2 that measures the predictive power of the model probabilities. In each case, as $\tilde{f}$ increases, so does the pseudo-R^2. However, the McFadden pseudo-R^2 has a predictable dependence on the the number of alternatives: if the number of alternatives in each choice problem increases, R_M^2 decreases. Hence, in order to define an adjusted McFadden pseudo-R^2 that is independent of the number of alternatives, we multiply R_M^2 by $\ln \tilde{n}$, i.e.,

(5.14) $$\tilde{R}_M^2 = (\ln \tilde{n})\left[1 - \frac{\ln L(\hat{\beta})}{\ln L(0)}\right].$$

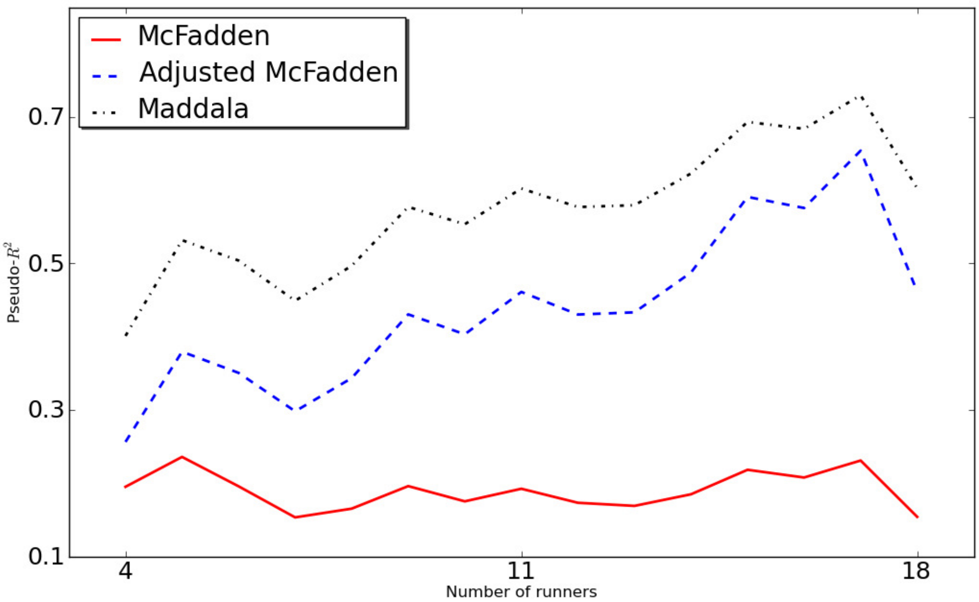
Note that this definition no longer has a maximum value of 1; it now has a maximum value of $\ln\tilde{n}$. For the Maddala pseudo-R^2, we find that it is, as required, already independent of the number of alternatives in each choice problem. However, it has a maximum value of $1-1/\tilde{n}^2$.

For an empirical demonstration that the adjusted McFadden and Maddala pseudo-R^2s are consistent even when the number of alternatives is varied, we fit CL models on subsets of horserace betting data categorized by the number of competitors in each event. The data employed are final bookmaker odds from 6064 UK races from 2009 and 2010. The CL models have just one independent variable, which is the log of odds-implied probability; the coefficient of this variable is given by β. The results are presented in Table 5.1 and Figure 5.2.

Table 5.1. Conditional logit models with log of odds-implied probability as the single variable fitted to different subsets of the data depending on the number of competitors in each event.

Number of competitors	Number of events	$\tilde{n}$	$\hat{\beta}$	$\ln L(0)$	$\ln L(\hat{\beta})$	R_M^2	$\tilde{R}_M^2$	R_D^2
4 or fewer	201	3.7	1.10	-264.2	-212.7	0.195	0.256	0.401
5	335	5	1.24	-539.2	-412.0	0.236	0.379	0.532
6	448	6	1.19	-802.7	-645.7	0.196	0.351	0.504
7	578	7	1.05	-1124.7	-952.4	0.153	0.298	0.449
8	611	8	1.11	-1270.5	-1060.7	0.165	0.343	0.497
9	646	9	1.24	-1419.4	-1141.3	0.196	0.430	0.577
10	585	10	1.17	-1347.0	-1111.1	0.175	0.403	0.554
11	572	11	1.23	-1371.6	-1107.8	0.192	0.461	0.602
12	533	12	1.17	-1324.5	-1095.1	0.173	0.430	0.577
13	450	13	1.21	-1154.2	-959.3	0.169	0.433	0.580
14	345	14	1.24	-910.5	-742.3	0.185	0.488	0.623
15	222	15	1.46	-601.2	-470.0	0.218	0.591	0.693
16	189	16	1.38	-524.0	-415.2	0.208	0.576	0.684
17	83	17	1.72	-235.2	-180.9	0.231	0.654	0.730
18 or more	266	20.0	1.29	-797.2	-674.5	0.154	0.461	0.603
All events	6064	9.6	1.21	-13686.1	-11194.8	0.182	0.411	0.560

Figure 5.2. A comparison of the pseudo-R2s across subsets of the data categorized by the number of competitors in each event.



Clearly, the adjusted McFadden and Maddala pseudo-R^2s vary consistently with each other when the number of competitors is changed, while the standard definition of the McFadden pseudo-R^2 does not vary in the same way. Hence, while there appears to be an increasing trend in model goodness-of-fit as the number of runners is increased, this trend is not captured by the standard definition of the McFadden pseudo-R^2, because it has the tendency to decrease as the number of runners is increased; see equation (5.12). We confirm this by fitting linear regressions of $R_D^2 - R_M^2$ and $R_D^2 - \widetilde{R}_M^2$ on the number of runners; gradients are given by 0.0174 (t = 10.67, p = 0.0000) and -0.0035 (t = 0.20, p = 0.4211), respectively, i.e., the difference between the Maddala and McFadden pseudo-R^2s increases with the number of runners while the difference between the Maddala and adjusted McFadden pseudo-R^2s do not. In the next section, we continue to address the comparability of pseudo-R^2s, by describing and comparing two methods for estimating their distributions.

5.4. Distributional properties of pseudo-R^2s

5.4.1. Bootstrapping pseudo-R^2s

It is straightforward to compare nested CL models (i.e., models where one of the models includes all the independent variables from the other model) fitted on the same data (e.g., using the likelihood ratio statistic). However, as researchers we run into difficulty when attempting to compare either non-nested models fitted on the same data or models that are fitted on mutually exclusive data. One approach would be to make comparisons between the pseudo-R^2s for each model. However, it is not a simple task to assign a measure of precision to values of pseudo-R^2 because their distributions are complex and depend on unknown parameters. An alternative method for estimating the distribution of pseudo-R^2s is to adopt an M-bootstrap approach (Efron, 1979), as recommended by Ohtani (2000) for ordinary R^2s. The bootstrap is commonly used when the theoretical distribution of a statistic is complicated, which is the case for the CL model. Suppose we have fitted CL models to two datasets, D_1 and D_2, consisting of N_1 and N_2 events, respectively. The M-bootstrap method proceeds as follows:

1. Sample N_1 events, with replacement, from D_1, to form a new dataset BD_1. Similarly, sample N_2 events, with replacement, from D_2, to form a new dataset BD_2.
2. Fit CL models on BD_1 and BD_2 and record the resulting values of pseudo-R^2.
3. Perform M iterations of steps 1 and 2.
4. The sample means, $\mu(R_1^2)$ and $\mu(R_2^2)$, and sample variances, $s^2(R_1^2)$ and $s^2(R_2^2)$, of the sets of M pseudo-R^2s are used to derive a standard normal test statistic, $z[\mu(R^2)] = [\mu(R_1^2) - \mu(R_2^2)] / \sqrt{s^2(R_1^2) + s^2(R_2^2)}$, which can be used to test the alternative hypothesis that the estimated probabilities from one model are more accurate than the other vs. the null hypothesis of no difference.

5.4.2. The asymptotic distribution of pseudo-R^2s

An alternative method is to estimate the *asymptotic distribution* of the pseudo-R^2, i.e., the expected distribution as the number of events tends to infinity. Hu, Shao, and Palta (2006) derive analytically the asymptotic distribution of the Maddala pseudo-R^2 in the multinomial logit model (like the conditional logit model, the multinomial logit model is a model of discrete of choice, but with different underlying assumptions). Here, we adapt their analysis to derive the asymptotic distribution of the McFadden and Maddala pseudo-R^2s for the conditional logit model.

Proposition 2. *Assume that the independent variables x_{ij}, j = 1, 2, ... , N, i = 1, 2, ... , n_j, are independent and identically distributed random m-vectors with finite second moment (i.e., $E(x_{ij}{}^2)$ finite). Let*

(5.15) $$H_1 = E(\ln n_j),$$

$$H_2 = -E(\textstyle\sum_{i=1}^{n_j} p_{ij} \ln p_{ij}).$$

and let

(5.16) $$\Sigma = \begin{pmatrix} Var(n_j) & \eta \\ \eta & \varepsilon \end{pmatrix},$$

$$g_1 = \frac{1}{\ln \lambda}\left(\frac{H_2}{\lambda \ln \lambda}, \quad 1\right)^T,$$

$$g_2 = \frac{2e^{2H_2}}{\lambda^2}\left(\frac{1}{\lambda}, \quad 1\right)^T,$$

where

(5.17) $$\lambda = E(n_j),$$

$$\eta = E[n_j \textstyle\sum_{i=1}^{n_j} p_{ij} \ln p_{ij}] + \lambda H_2,$$

$$\varepsilon = E[\textstyle\sum_{i=1}^{n_j} p_{ij} (\ln p_{ij})^2] - H_2{}^2.$$

Then, as $N \to \infty$,

(5.18) $$\sqrt{N}[R_M^2 - (1 - H_2 / H_1)] \rightarrow_d N(0, \sigma_1^{\ 2}),$$

$$\sqrt{N}[R_D^2 - (1 - e^{2(H_2 - H_1)})] \rightarrow_d N(0, \sigma_2^{\ 2}),$$

where $\sigma_1^{\ 2} = g_1^{\ T} \Sigma g_1$ *and* $\sigma_2^{\ 2} = g_2^{\ T} \Sigma g_2$.

Proof. For a proof, see Appendix 4.

The above proposition gives the asymptotic distributions of the McFadden and Maddala pseudo-R^2s. Hence, to obtain estimates of the variance of point estimates of these pseudo-R^2s, we can replace the unknown quantities with consistent estimators. So, denote by $\overline{n}$, $\widetilde{n}$, and $s^2(n)$ the arithmetic mean, geometric mean, and sample variance of the number of competitors, respectively, i.e.,

(5.19) $$\overline{n} = (1/N)\sum_{j=1}^{N} n_j ,$$

$$\widetilde{n} = (\prod_{j=1}^{N} n_j)^{1/N} ,$$

$$s^2(n) = \frac{1}{N-1}\sum_{j=1}^{N}(n_j - \overline{n})^2 .$$

Then, let

(5.20) $$\hat{\Sigma} = \begin{pmatrix} s^2(n) & \hat{\eta} \\ \hat{\eta} & \hat{\varepsilon} \end{pmatrix},$$

$$\hat{g}_1 = \frac{1}{\ln \overline{n}} \begin{pmatrix} \frac{\hat{H}}{\overline{n} \ln \overline{n}}, & 1 \end{pmatrix}^T ,$$

$$\hat{g}_2 = \frac{2}{\widetilde{n}^2} e^{2\hat{H}} \begin{pmatrix} \frac{1}{\overline{n}}, & 1 \end{pmatrix}^T ,$$

where

(5.21) $$\hat{\eta} = (1/N)\sum_{j=1}^{N} n_j \sum_{i=1}^{n_j} p_{ij} \ln p_{ij} + \overline{n}\hat{H}_2 ,$$

$$\hat{\varepsilon} = (1/N)\sum_{j=1}^{N}\sum_{i=1}^{n_j} p_{ij} (\ln p_{ij})^2 - \hat{H}_2^{\ 2} .$$

Here

(5.22) $$\hat{H}_2 = -(1/N)\sum_{j=1}^{N}\sum_{i=1}^{n_j} p_{ij} \ln p_{ij} \ .$$

Then estimates of the variance of the McFadden and Maddala pseudo-R^{2}s are given by

(5.23) $$s^2(R_M^2) = \frac{1}{N}(\hat{g}_1{}^T \hat{\Sigma} \hat{g}_1) \, ,$$

$$s^2(R_D^2) = \frac{1}{N}(\hat{g}_2{}^T \hat{\Sigma} \hat{g}_2) \, ,$$

respectively.

5.4.3. An empirical comparison of the asymptotic method with the bootstrap method

Now we discuss the differences between the two alternative methods for estimating variances of pseudo-R^2s described above. Teebagy and Chatterjee (1989) show that the bootstrap method overestimates standard errors in large samples for standard logistic regression (relative to the asymptotic distribution). Here we briefly compare values of the standard deviations of the pseudo-R^2s from CL models estimated using each method on some real data. The data employed are final exchange odds from 6058 UK races from 2009 and 2010. There are two subsets of the data, consisting, respectively, of those races with above-/below-median Shin z (with 3029 races in each set), where Shin z measures the extent of informed trading in the market (see Appendix 1 for more details). We again fit CL models with just the log of odds-implied probability as the single explanatory variable; the coefficient of this variable is given by β. The results are presented in Table 5.2.

Table 5.2. A comparison of the asymptotic and bootstrap methods for estimating the distributions of the McFadden and Maddala pseudo-R^2s.

Dataset	High Shin z	Low Shin z
Number of events	3029	3029
Mean number of competitors	7.5	13.1
β	1.03	1.01
$\ln L(0)$	-5967.3	-7702.9
$\ln L(\hat{\beta})$	-4900.1	-6224.6
Asymptotic		
R_M^2	0.1788	0.1919
$S.E.(R_M^2)$	0.00757	0.00690
R_D^2	0.5057	0.6232
$S.E.(R_D^2)$	0.01282	0.01218
Bootstrap		
$\mu(R_M^2)$	0.1793	0.1925
$S.E.(R_M^2)$	0.00671	0.00649
$\mu(R_D^2)$	0.5064	0.6240
$S.E.(R_D^2)$	0.01307	0.01235
$F_{3028,3028}(R_M^2)$	1.27**	1.13**
$F_{3028,3028}(R_D^2)$	1.04	1.03

**: significant at the 1% level (1-tailed F test).

It is clear that, while the asymptotic and bootstrap methods produce estimates of the variances of the pseudo-R^2 statistics that are reasonably close, there is some difference: the asymptotic method appears to overestimate the variance of the McFadden pseudo-R^2 relative to the bootstrap method, while relatively underestimating the variance of the Maddala pseudo-R^2. However, the difference is only statistically significant for the McFadden pseudo-R^2, with F-tests of differences of variances for the high and low Shin z datasets given by 1.27 (p = 0.0000) and 1.13 (p = 0.0004), respectively. The F values for the Maddala pseudo-R^2s for the high and low Shin z datasets are given by 1.04 (p = 0.1383) and 1.03 (p = 0.2269), respectively.

5.5. Pseudo-R^2 as a predictor of the economic value of a discrete choice model

Competitive events are often associated with a market for trading on the outcome. For example, horseraces have an associated betting market in which traders can wager money on their predicted outcome. This results in it being possible to derive from market prices 'public' predictions of the probabilities of each outcome occurring. This is important because discrete choice models are often employed to test for market efficiency, i.e., the degree to which the market appropriately discounts the value of information. Hence, in this context, the standard motivation for pseudo-R^2s, as improvement from null model to fitted model, is less relevant. Instead, it is more useful to the modeller to have some understanding of the improvement of their model over the public model. We denote this measure *relative* pseudo-R^2 and show theoretically that there is a direct link between this measure, the transaction costs for betting on an event, and the expected profit to a bettor employing this measure.

Recall that our dataset consists of N events, where each event j is between an integer number $n_j \geq 2$ competitors i; for each event, there is just one winner, given by y_{ij}, where $y_{ij} = 1$ if competitor i wins, and $y_{ij} = 0$ otherwise. Suppose further that the decimal odds are denoted by $R_{ij} > 1$, with corresponding prices given by $r_{ij} = 1/R_{ij}$. The over-round is given by $B_j = \sum_{i=1}^{n_j} r_{ij} - 1$, so odds-implied probabilities are given by $q_{ij} = r_{ij}/(1+B_j)$ for all i. Suppose that the bettor assigns winning probabilities to each competitor of p_{ij}. Then the expected profit from a £1 bet on competitor i in event j is $p_{ij}R_{ij} - 1$. Denote the winning probability that the bettor assigns to the eventual winning competitor by p_j, and the odds-implied probability of the same competitor by q_j. Then the bettors' expected profit from a bet on the eventual winner, or 'edge', is given by

(5.24) $$W_j = \frac{p_j}{q_j(1+B_j)} - 1.$$

Now, define the relative McFadden pseudo-R^2 by

(5.25) $$\overline{R}_M^2 = 1 - \frac{\ln L(p)}{\ln L(q)},$$

and the relative Maddala pseudo-R^2 by

(5.26) $$\overline{R}_D^2 = 1 - \exp\{(2/N)[\ln L(q) - \ln L(p)]\},$$

where the log likelihood of the bettor's model is given by

(5.27) $$\ln L(p) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij}$$

and the log likelihood of the public's model is given by

(5.28) $$\ln L(q) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln q_{ij} .$$

Defined in this way, the relative pseudo-R^2s measure the degree of improvement of the model over the public odds *for the winning competitor only*. Thus we can derive a link between the bettor's realised edge on the winning competitor and the relative pseudo-R^2 of the model.

Since the relative McFadden pseudo-R^2 can be written as

(5.29) $$\overline{R}_M^2 = 1 - \frac{\sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij}}{\sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln q_{ij}},$$

and (5.24) can be rewritten as

(5.30) $$p_j = q_j (1 + W_j)(1 + B_j),$$

substituting this into (5.29) gives

(5.31) $$\overline{R}_M^2 = 1 - \frac{\sum_{j=1}^{N} \ln[q_j (1 + W_j)(1 + B_j)]}{\sum_{j=1}^{N} \ln q_j},$$

i.e.,

(5.32) $$\overline{R}_M^2 = \frac{\sum_{j=1}^{N} \ln[(1 + W_j)(1 + B_j)]}{\sum_{j=1}^{N} \ln[R_j (1 + B_j)]}.$$

Rearranging this gives

(5.33) $$\overline{R}_M^2 = \frac{\ln \prod_{j=1}^{N} (1 + W_j)(1 + B_j)}{\ln \prod_{j=1}^{N} R_j (1 + B_j)},$$

i.e.,

(5.34) $$\overline{R}_M^2 = \frac{\ln GM(1 + W_j) + \ln GM(1 + B_j)}{\ln GM(R_j) + \ln GM(1 + B_j)},$$

where $GM(x_j) = (\prod_{j=1}^{N} x_j)^{1/N}$ denotes geometric mean. Hence, the McFadden pseudo-R^2 is, *ceteris paribus*, increasing in average edge, increasing in average odds-implied probability of the winner, and usually increasing in average over-round (since, typically, $R_j > 1 + W_j$). Rearranging (5.34) gives

(5.35) $$GM(p_j / q_j) = (1 / q_j)^{\overline{R}_M^2},$$

i.e, the bettor's edge over the public is this function of their relative McFadden pseudo-R^2 and the odds-implied probability of the winning competitor.

Similarly, since the relative Maddala pseudo-R^2 can be written as

(5.36) $$\overline{R}_D^2 = 1 - \exp[(2/N)\sum_{j=1}^{N}\sum_{i=1}^{n_j} y_{ij} \ln(q_{ij} / p_{ij})],$$

and (5.24) can be rewritten as

(5.37) $$\frac{q_j}{p_j} = \frac{1}{(1+W_j)(1+B_j)},$$

substituting this into (5.36) gives

(5.38) $$\overline{R}_D^2 = 1 - \exp\{-(2/N)[\sum_{j=1}^{N} \ln(1+W_j) + \ln(1+B_j)]\}.$$

Rearranging this gives

(5.39) $$\overline{R}_D^2 = 1 - \prod_{j=1}^{N}[(1+W_j)(1+B_j)]^{-2/N},$$

i.e.,

(5.40) $$\overline{R}_D^2 = 1 - \frac{1}{GM(1+W_j)^2 \, GM(1+B_j)^2}.$$

So the Maddala pseudo-R^2 is, *ceteris paribus*, increasing in average edge and increasing in average over-round. Rearranging (5.40) gives

(5.41) $$GM(p_j / q_j) = \frac{1}{\sqrt{1 - \overline{R}_D^2}},$$

i.e., the bettor's edge is positively related to the relative Maddala pseudo-R^2 but independent of the odds-implied probability of the winner. Clearly, relative pseudo-R^2s are useful in competitive event prediction. However, we have also shown in equations (5.35) and (5.41) that direct relationships can be derived between predicted model probabilities, 'public' model probabilities, and relative pseudo-R^2 measures. These relationships are important because they show us first that the economic value of estimated model probabilities is increasing in pseudo-R^2s, and also the functional form of this relationship. These formulae are useful for understanding the context in which pseudo-R^2s should be reported, and could be employed in wider contexts, e.g., when there is a prior probability of an alternative being chosen that is more appropriate than a null probability.

5.6. Conclusion

In this paper, we have set out to describe and evaluate properties of pseudo-R^2s as a measure of forecast accuracy in discrete choice models, a class of models that have a wide range of applications, from predicting consumer demand to epidemiology and operations research. While R^2 in ordinary least squares linear regression is a widely-used and well-justified measure, the same is typically not true of pseudo-R^2s. For instance, we have shown both theoretically and empirically that at least one of the definitions of pseudo-R^2 (McFadden's definition) is not robust to variations in the number of alternatives in each choice problem. We have therefore suggested an adjustment to correct for this bias. This has important implications for the comparability of pseudo-R^2 measures across models, particularly non-nested models or models fitted on different datasets, which is a key desirable property of any forecast evaluator. Further work could investigate the comparability of other definitions of pseudo-R^2 in discrete choice models or other models that involve the logistic function, such as generalized additive models.

We have also described two methods for estimating the variance of pseudo-R^2s so that their values can be statistically compared: the bootstrap and asymptotic methods. A comparison of the two methods on actual data demonstrates that they are reasonably close in the estimates that they produce, so we would recommend that either method is useful in obtaining standard errors for pseudo-R^2s. Our findings here contribute to an understanding of the use of pseudo-R^2s in general, which are often simply reported without a justification or without standard errors when comparing them across models. Moreover, these methods are particularly useful when the pseudo-R^2 itself is the value of interest in hypothesis testing; for instance, in comparing the predictive power of discrete choice models, or evaluating the efficiency of speculative financial markets. Finally, we have derived simple relationships between relative pseudo-R^2 measures and the expected profit to a bettor from betting on competitive events, which is an important relationship because choice modelling is often employed in the context of competitive events in order to assess market efficiency, and efficient markets are a desirable goal in the regulation of financial markets to minimize risks of financial shocks. As a result of all of the above, we have arrived at a much more rigorous understanding of the value of pseudo-R^2s in evaluating the predictions of discrete choice models.

Conclusion

This section briefly summarizes the main findings of each of the five papers in this thesis, states the contributions of each, discusses the contributions of the thesis as a whole to knowledge and understanding of biased decision making, and finally discusses the implications of the research for policy-making and future research.

The theme of Chapter 1 was that there are a variety of reasons why evidence of biased decision making from laboratory-based studies need not necessarily translate to the population as a whole, particularly to individuals who have expertise in the task they are carrying out, are offered meaningful incentives, and dedicate enough time to the task so as to receive regular, unambiguous feedback on the outcomes of their decisions. In discussing evidence (or lack of evidence) of biased decision making by participants in betting markets, this paper shows numerous examples that bettors are rational and well-calibrated decision makers, who generally aren't subject to heuristics and biases that are well-documented from laboratory-based research (the favourite-longshot bias is an important exception). The main conclusion was that future research into biased decision making should always account for the observed differences in behaviour between laboratory and naturalistic settings. While there are advantages and disadvantages of naturalistic and laboratory studies, a tandem approach, with each informing the other, seems to be the way forward in order to assess the true nature of biased decision making.

Chapter 2 was the first study to investigate noise trading and herding in a betting exchange market. The main conclusions were as follows: first, noise trading is associated with increased market efficiency, a result that is attributed to the improved liquidity attracting an increased level of participation by informed traders. Second, herding is prevalent, particularly so in the later, more active stages of the market and in different levels depending on whether price changes follow a 'buy' or a 'sell' signal. The findings contribute to the understanding of the role of noise in the efficiency of financial markets in general, as this paper is able to overcome a significant methodological problem with using regular financial market data, in that regular financial markets always represent current expectation of future prices, and so are inherently uncertain. In contrast, betting market prices reach a defined termination point at which all uncertainty is resolved. The main finding of this paper, that noise trading, volatility, and efficiency of final market prices all move in tandem, has important policy implications for all financial markets. For example, the results support arguments that regulatory measures to protect investors from the destabilizing effects of noise are self-defeating, and suggest that focusing on innovative means of reducing the risks to arbitrageurs, rather than discouraging speculators, may be the best approach to achieving efficient markets. It is also demonstrated that the inefficiency resulting from herding is of such a magnitude that it is possible to make positive returns from strategies in counter to those of herding traders. The contribution of these findings is that regulations in financial markets should be devised to minimize the impact of herding while avoiding restrictions to noise traders: particular attention should be given to situations where uninformed traders may incorrectly believe that there are traders with privileged information operating. Furthermore, markets that involve contingent returns at a fixed point in time (such as the markets examined here) should always be allowed sufficient time to reach efficient levels.

Chapters 3 and 4 investigated the favourite-longshot bias (FLB) using recent data from all three major types of betting market in the UK and Ireland: bookmakers, exchanges, and pari-mutuel pools. Previous research has shown that the FLB has existed or is absent in a variety of jurisdictions over many decades. These papers provide new reasons for the presence or absence of the bias in each type of market. In the case of bookmakers and exchanges, Chapter 3 makes a number of important contributions. First, it contains evidence of (i) FLB in bookmaker markets, and (ii) no FLB in exchange markets. Second, it confirms that predictive accuracy of exchange prices is largely superior to that of bookmaker prices. However, in the case of exchange markets, it also uncovers significant relationships between the FLB and betting volume, and between the FLB and over-round. It also shows that there are no such relationships in bookmaker markets. Most significantly, a model is developed that suggests that the optimal pricing policy for a bookmaker, who competes with other operators for betting on favourites, is to set prices which include the FLB. The empirical results are largely supportive of the predictions of this model.

In pari-mutuel markets, Chapter 4 has developed a model that shows that, in order to explain the FLB without reference to market makers, one must account for a 'strong favourite' effect, whereby bettors who generally underbet favourites give undue preference to favourites in races where the variance of the odds of their rivals is low. This effect is predicted by a representative agent model employing prospect theory as the driving force behind bettors' decisions, but is not predicted by a model employing expected utility theory. Most importantly, these predictions do not depend on parametric assumptions or assumptions about the bettor's choice set as in previous research. Hence, the results provide definitive evidence that cognitive errors of traders, rather than their risk preferences, explain the FLB. Together, these two papers contribute more robust evidence, using new and innovative methodologies, that explain the FLB in each type of market.

The final chapter explores and solves a number of methodological issues relating to pseudo-R^2s as a measure of evaluating forecast accuracy of discrete choice models, particularly the conditional logit (CL) model. Having employed the CL model throughout this thesis, we were able to use results derived in this chapter in order to test some of the hypotheses in earlier chapters. Moreover, this paper shows how these concerns relate to wider issues in the field of discrete choice modelling as a whole. Future work could explore a much more general understanding of the role that pseudo-R^2s have to play in statistical modelling.

Overall, this thesis makes a significant contribution towards understanding the extent and nature of biased decision making in naturalistic environments. As it was argued in Chapter 1, there are many reasons to doubt the generalizability of laboratory research, so demonstrating such biases in the real world is crucial to establishing solid foundations to the knowledge and understanding of the manner in which individuals make decisions. Indeed, in surveying previous literature in Chapter 1, the conclusion that was arrived upon is that decision makers operating in their natural environment generally do not make biased decisions, or at least, the extent and generality of such biases is significantly lower. This puts the main findings in Chapters 2 and 3 in context: each of these papers demonstrated evidence of biased decision making in the naturalistic environment of betting on horseracing (herding and the FLB, respectively), and thus make a significant contribution to the literature. Moreover, in Chapter 2, by categorizing market activity by the time until the end of the market and whether it results in large 'buy' or 'sell' movements, this paper was able to identify and discuss reasons why herding might occur. This is in itself is a major contribution over previous studies of herding in betting markets (discussed in Chapter 1) and in regular financial markets (discussed briefly in Chapter 2). Similarly, in Chapter 3, it was shown that the extent of the FLB is dependent on a range of factors, particularly the type of market (exchange or bookmaker), since transaction costs and the manner in which prices are set is different in each type of market. The different costs ensure that the type of decision maker, in terms of the information they hold, operating in each market is distinct: traders in exchange markets are more informed, and so the extent of systematic bias is reduced, and conversely, systematic bias is not eliminated in bookmaker markets because costs are restrictive. Furthermore, the bookmaker's effective monopoly in setting prices as well as the general demand preferences of bettors participating in these markets ensures that bookmaker markets display significant bias throughout their duration. All of the above considerations contribute to the developing understanding of how systematic decision making biases can be allowed to persist (monopoly pricing and high costs), but also how they can be eliminated (competitive pricing and low costs). Chapter 4 advanced a new method (free of certain restrictive assumptions) that allows researchers to distinguish between different models of decision making that can account for biases such as the FLB. Hence, it makes a significant contribution in guiding future research to assessing the relative merits of alternative models of decision making. Similarly, Chapter 5 discussed a range of limitations in the current methodology for evaluating discrete choice models. In a sense, this paper makes a contribution to what is hoped can be an informative and highly useful avenue of future research.

Research into decision making biases in general, this thesis included, has major policy implications in a number of fields. First, a detailed understanding of heuristics and biases in a management context is key to guiding operational decision making and avoiding costly errors at an organizational level. Hence, a greater understanding of biases such as herd behaviour and overweighting of small probabilities can guide policy-makers in organizations at a high level. Second, the efficient operation of financial markets depends on the appropriate regulation of such markets. A greater understanding of the extent and causes of herding, the effects of noise trading, and other factors, therefore contributes to better deployment of regulations. Finally, a large part of psychological research into behaviour is driven by real-world events, and so naturalistic research aids the future direction and topicality of laboratory research.

Finally, it should be offered that the main strength of this thesis is also its greatest limitation: in studying decision making in the naturalistic environment of betting markets for horseracing, the scope of the empirical research is necessarily narrow. Future work in the area of naturalistic biased decision making could further investigate the consistency of empirical evidence of other large populations of decision makers with the predictions of theoretical models. The FLB is a decision making bias that has received a great deal of attention in the literature not only because of its prevalence but because of what it tells us about the manner in which people make decisions in a general sense. This thesis has described new and more satisfactory explanations for the FLB in the three major types of betting market in the UK (and indeed, globally), and in doing so has devised new and innovative theoretical models and empirical methodologies and resolved unanswered questions related to the topic. It has also shown and explained the prevalence of herd behaviour in these markets and the effect that noise trading has on market volatility and efficiency, and outlined a range of implications of these findings for policy-making in wider financial markets, an important subject because regular financial market data are unsuited to robust tests of market efficiency. As a result of the above, this thesis has made significant contributions to the existing literature on biased decision making in speculative financial markets.

Appendix 1. The Shin model

In a series of three papers, Shin (1991, 1992, and 1993) demonstrated that, in a bookmaker market, the FLB can be explained by supply-side factors: specifically, price-setting by the bookmakers themselves. Shin modelled bookmaker markets as a game between a profit-maximizing bookmaker and a randomly chosen bettor. The model assumes that the bettor is likely to be a noise trader, but could be a perfectly informed insider, who knows precisely the identity of the winning horse. The model predicts that, since the bookmaker is not perfectly informed, they will depress odds on longshots (the horses with the least chances of winning the race) relative to those on favourites in order to protect themselves from the possibility of large losses to the insider, who is in possession of superior information.

Formally, the model is of an *n*-horse race that involves a monopoly bookmaker, a perfectly informed insider trader, and a set of uninformed outsiders. The bookmaker sets prices r_i (corresponding to decimal odds of $R_i = 1/r_i$) on all horses, subject to 0 < r_i < 1 for all *i*, and $\sum_{i=1}^{n} r_i \leq 1 + B$, where B > 0 is as small as is required for the bookmaker to obtain monopoly rights through competition with other potential market makers. The bookmaker knows the true winning probabilities p_i, but only the insider knows the identity of the winning horse in advance. A bettor is randomly selected to face the bookmaker; the bettor is the insider with probability z ($0 \leq z < 1$), or an outsider who attaches probability 1 to the *i*-th horse with probability $(1-z)p_i$. The bettor is then permitted to bet £1 with the bookmaker on their preferred horse, which is the winning horse if the bettor is the insider, or the *i*-th horse if the bettor is the *i*-th outsider. Hence, the problem for the expected profit-maximizing bookmaker is to set the r_i to maximize

(A1.1) $$1 - \sum_{i=1}^{n} \frac{zp_i + (1-z)p_i^2}{r_i}$$

subject to $\sum_{i=1}^{n} r_i \leq 1 + B$ and 0 < r_i < 1 for all *i*. The solution of this problem is given by

(A1.2) $$1 + B = \sum_{s=1}^{n} r_s = [\sum_{s=1}^{n} F(p_s)]^2 ,$$

$$r_i = F(p_i) \sum_{s=1}^{n} F(p_s) ,$$

where

(A1.3) $$F(p_i) = [zp_i + (1-z)p_i^{\,2}]^{1/2} ,$$

which gives rise to a FLB, i.e., $r_i / r_j < p_i / p_j \Leftrightarrow p_i > p_j$.

The value of *z* in Shin's (1993) model gives rise to a direct means (known as Shin *z*) of measuring the proportion of market participation that can be attributed to traders with privileged information. Although in Shin's original model, informed traders are perfectly informed, Fingleton and Waldron (1999) relaxed this assumption, showing that it is equivalent to suppose that the precision of the informed trader's information can vary, and that the Shin *z* value is equal to the level of informed trading times the degree of precision. Hence, it is reasonable to assume a more general situation in which a range of different types of informed traders operate, but that the level of influence they have in the market is likely to vary in tandem. The Shin *z* value itself is directly derived from final bookmaker prices and has been used extensively in betting market studies in order to investigate claims relating to the level of informed trading (e.g., Vaughan Williams and Paton, 1997; Smith et al., 2006).

There are several accepted methods for estimating the Shin *z* value for a given event. Shin's (1993) own method was based around the Taylor series expansion of $F(p_i)$ around 1/*n*. Jullien and Salanié (1994) noticed that the equations in (A1.2) can be rearranged to give

(A1.4) $$F(p_i)^2 = r_i^{\,2} / (1+B) ,$$

which in turn can be solved for p_i to yield

(A1.5) $$p_i = \left[\frac{z^2}{4(1-z)^2} + \frac{r^2}{(1+B)(1-z)} \right]^{1/2} - \frac{z}{2(1-z)} .$$

Thus, the condition $\sum_{i=1}^{n} p_i = 1$ can be used to estimate the value of *z* for a given event. In Chapter 2 we adopt the iterative method of Law and Peel (2002), in which we square and sum (A1.5) before rearranging the resulting expression to give

(A1.6) $$z = \frac{[1/(1+B)]\sum_{i=1}^{n} r_i^2 - \sum_{i=1}^{n} p_i^2}{1 - \sum_{i=1}^{n} p_i^2}.$$

The iterative procedure is then to start with an initial estimate of z, calculate the p_i using (A1.5), calculate a new value of z using (A1.6), and repeat these two steps until convergence.

Shin's model can explain the FLB in bookmaker markets and gives a useful measure of insider (or just informed) trading in these markets. However, one must accept the assumptions of Shin's model. The key assumption is that knowledgeable insiders are more likely to bet on longshots than favourites, which is a reasonable assumption to make given anecdotal evidence. However, it has been pointed out (e.g., Schnytzer and Shilony, 2005) that some of the assumptions in Shin's model are unrealistic. However, similar conclusions can be reached with these assumptions relaxed. For example, Schnytzer and Shilony (2005) found that bookmakers should raise prices on longshots more than favourites in order to defend themselves against insider knowledge, without assuming that insiders know which horse will win the race, or that insiders are more likely to bet on longshots. Peirson and Smith (2010) extend the Shin model while relaxing the assumptions that insiders know which horse will win the race, and that the amount bet by insiders is fixed and not related to the odds on offer. Their model demonstrates that bookmakers should increase prices on those horses where there is a higher probability of inside information being employed.

Appendix 2. A model of competing markets

Here we prove the main propositions from the model in section 3.3. We consider two markets, a bookmaker and an exchange, which offer prices q_i and r_i, respectively, on a single race with n runners, with over-rounds given by $B_b = \sum_{i=1}^{n} q_i - 1$ and $B_e = \sum_{i=1}^{n} r_i - 1$, respectively, with $B_b > B_e$. We assume that the bookmaker's best estimates for true probabilities are the exchange odds-implied probabilities, i.e., $p_i = r_i/(1+B_e)$. We also make the approximation that the q_i are continuous on the interval (0, 1). Considering a small time interval, over which B_e, B_b, and the r_i are constant the bookmaker's goal is to maximize their expected returns $G(q_1,\ldots,q_n)$ over this time interval. Denote the demand curve for horse i, which is the amount bet on horse i when the bookmaker and exchange prices are q_i and r_i, respectively, by $f(q_i, r_i)$, normalized so that it satisfies $f(r, r) = r/(1+B_e)$. Therefore, $G(q_1,\ldots,q_n)$ is given by

(A2.1) $$G(q_1,\ldots,q_n) = \sum_{i=1}^{n} f(q_i, r_i)\left(1 - \frac{r_i}{(1+B_e)q_i}\right)$$

subject to the over-round condition $\sum_{i=1}^{n} q_i = 1 + B_b$.

For fixed r_i, B_b and B_e, this is a constrained optimization problem to maximize

(A2.2) $$H(q_1,\ldots,q_n) = G(q_1,\ldots,q_n) - \lambda\left(\sum_{i=1}^{n} q_i - 1 - B_b\right),$$

where λ is a constant. The solutions are given by the system of equations

(A2.3) $$\partial H / \partial q_i = f(q_i, r_i)\frac{r_i}{(1+B_e)q_i^{\,2}} + f'(q_i, r_i)\left(1 - \frac{r_i}{(1+B_e)q_i}\right) - \lambda = 0,\ i = 1, \ldots, n,$$

$$\sum_{i=1}^{n} q_i = 1 + B_b,$$

We do not seek to solve this system of equations, but note a sufficient condition for the FLB in the bookmaker prices: for two horses j (a longshot) and k (a favourite: $r_j < r_k$), with odds-implied probabilities equal across the exchange and bookmaker markets ($q_j = (1+B_b)r_j/(1+B_e)$, $q_k = (1+B_b)r_k/(1+B_e)$), the marginal increase in expected returns for an increase in price is greater for the longshot ($\partial H / \partial q_j > \partial H / \partial q_k$). Furthermore, the greater the difference, the greater the level of FLB. So, denoting $\hat{q}_i = (1+B_b)r_i/(1+B_e)$ and with some manipulation, this condition becomes

(A2.4) $$\frac{1}{r_j} f(\hat{q}_j, r_j) - \frac{1}{r_k} f(\hat{q}_k, r_k) + \frac{1+B_b}{1+B_e} B_b [f'(\hat{q}_j, r_j) - f'(\hat{q}_k, r_k)] > 0.$$

We have now the following as a sufficient condition for the FLB:

(A2.5) $$f'(xr_j, r_j) > f'(xr_k, r_k),$$

for all x such that $1 \le x \le (1+B_b)/(1+B_e)$. That (A2.4) follows from (A2.5) is proved in the following proposition.

Proposition 1. *If condition* (A2.5) *is satisfied, then so is* (A2.4).

Proof. If $r_j < r_k$, and $1 \le x \le (1+B_b)/(1+B_e)$, then by (A2.5), $f'(xr_j, r_j) > f'(xr_k, r_k)$. So, by an integration inequality,

(A2.6) $$\int_{r_j}^{\hat{q}_j} f'(s, r_j) ds > \int_{r_k}^{\hat{q}_k} f'(s, r_k) ds.$$

Hence, by the Fundamental Theorem of Calculus,

(A2.7) $$f(\hat{q}_j, r_j) - f(r_j, r_j) > f(\hat{q}_k, r_k) - f(r_k, r_k).$$

Furthermore, $1/r_j > 1/r_k$, so

(A2.8) $$\frac{1}{r_j}[f(\hat{q}_j, r_j) - f(r_j, r_j)] > \frac{1}{r_k}[f(\hat{q}_k, r_k) - f(r_k, r_k)].$$

Now, $f(r, r) = r/(1+B_e)$, so

(A2.9) $$\frac{1}{r_j} f(\hat{q}_j, r_j) - \frac{1}{1+B_e} > \frac{1}{r_k} f(\hat{q}_k, r_k) - \frac{1}{1+B_e},$$

i.e.,

(A2.10) $$\frac{1}{r_j} f(\hat{q}_j, r_j) > \frac{1}{r_k} f(\hat{q}_k, r_k).$$

So inequality (A2.4) holds.

Moreover, the difference $\partial H/\partial q_j - \partial H/\partial q_k$ is increasing in both B_b (holding $(1+B_b)/(1+B_e)$ fixed) and $(1+B_b)/(1+B_e)$, i.e., the level of FLB should increase both with bookmaker over-round and the level of competition between the bookmaker and the exchange.

Proposition 2. *If there is no competition* (B_b = B_e), *then there is still FLB* (*provided* (A2.5) *holds*).

Proof. If B_b = B_e, then $\hat{q}_i = r_i$, so equation (A2.4) becomes

(A2.11) $$\frac{1}{r_j} f(r_j, r_j) - \frac{1}{r_k} f(r_k, r_k) + B_e[f'(r_j, r_j) - f'(r_k, r_k)] > 0 .$$

Now, $f(r,r) = r/(1+B_e)$, so the first two terms in (A2.11) cancel, leaving the condition

(A2.12) $$f'(r_j, r_j) - f'(r_k, r_k) > 0 .$$

This is satisfied by (A2.5) when x = 1, so (A2.4) is a sufficient condition for the FLB even when B_b = B_e.

Proposition 3. *Furthermore, if there is no over-round* (B_b = 0) *then there is no FLB.*

Proof. We show that $\partial H / \partial q_i = 0$ is satisfied by q_i = r_i. If B_b = B_e = 0, then (A2.3) simplifies to

(A2.13) $$\partial H / \partial q_i = f(r_i, r_i) \frac{r_i}{r_i^2} + f'(r_i, r_i) \left(1 - \frac{r_i}{r_i}\right) - \lambda .$$

Now, $f(r,r) = r/(1+B_e)$, so

(A2.14) $$\partial H / \partial q_i = 1 - \lambda .$$

This is true for all i so we can choose λ = 1, i.e., $\partial H / \partial q_i = 0$.

Appendix 3. Proof from Chapter 4

We show here that, for $v(t)$ an increasing function that is concave for $t > 0$,

(A3.1) $$X = \sum_{k=2}^{n} \frac{1}{v(xR_k) - v(-x)}$$

is decreasing in the variance of the R_k. To see this, first suppose there are just 3 horses in the race. Denote the odds on the non-favourites by R_2 and R_3, with respective 'prices' given by $r = 1/R_2$ and $K - r = 1/R_3$. Since $v(-x)$ is constant for all k, we set $v(-x) = 0$. Representing the value function for gains by a concave power function $v(t) = t^{\alpha}$, $0 < \alpha < 1$, we have that

(A3.2) $$X = x^{-\alpha}[r^{\alpha} + (K - r)^{\alpha}].$$

Hence

(A3.3) $$\frac{dX}{dr} = \alpha x^{-\alpha}[r^{\alpha-1} - (K - r)^{\alpha-1}],$$

which is negative if and only if $r > K/2$, i.e., increasing the variance of the prices (and hence odds) for non-favourites has the effect of decreasing X. The converse result holds for the expected utility model if the utility of wealth function is convex.

Appendix 4. Proofs from Chapter 5

Proposition 1. *If the model probabilities are* $p_{ij} = f(n_j)/n_j$, *then the McFadden and Maddala pseudo-*R^2*s are given by*

(A4.1) $$R_M^2 = \frac{\ln \tilde{f}}{\ln \tilde{n}}$$

and

(A4.2) $$R_D^2 = 1 - \frac{1}{\tilde{f}^2},$$

respectively, where $\tilde{n} = (\prod_{j=1}^{N} n_j)^{1/N}$ *and* $\tilde{f} = [\prod_{j=1}^{N} f(n_j)]^{1/N}$ *are the geometric means of the number of competitors and of* $f(n_j)$, *respectively. Here, N is the number of events.*

Proof. First, recall that the definitions of the McFadden and Maddala pseudo-R^2s are given by

(A4.3) $$R_M^2 = 1 - \frac{\ln L(\hat{\beta})}{\ln L(0)}$$

and

(A4.4) $$R_D^2 = 1 - \exp\{(2/N)[\ln L(0) - \ln L(\hat{\beta})]\},$$

respectively, where $\ln L(\hat{\beta})$ is the maximum log-likelihood of

(A4.5) $$\ln L(\beta) = \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij}$$

and

(A4.6) $$\ln L(0) = \sum_{j=1}^{N} \ln(1/n_j).$$

Hence, for the McFadden pseudo-R^2,

(A4.7) $$\begin{aligned} R_M^2 &= 1 - \frac{\sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij}}{\sum_{j=1}^{N} \ln(1/n_j)} \\ &= 1 - \frac{\sum_{j=1}^{N} \ln[f(n_j)/n_j]}{\sum_{j=1}^{N} \ln(1/n_j)} \\ &= 1 - \frac{\sum_{j=1}^{N} [\ln f(n_j) + \ln(1/n_j)]}{\sum_{j=1}^{N} \ln(1/n_j)} \end{aligned}$$

$$= \frac{\sum_{j=1}^{N} \ln f(n_j)}{\sum_{j=1}^{N} \ln n_j}$$

$$= \frac{\ln \prod_{j=1}^{N} f(n_j)}{\ln \prod_{j=1}^{N} n_j}$$

$$= \frac{\ln(\tilde{f}^N)}{\ln(\tilde{n}^N)}$$

$$= \frac{\ln \tilde{f}}{\ln \tilde{n}}.$$

For the Maddala pseudo-R^2,

(A4.8) $$R_D^2 = 1 - \exp\{(2/N)[\sum_{j=1}^{N} \ln(1/n_j) - \sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ij} \ln p_{ij}]\}$$

$$= 1 - \exp\{(2/N)[\sum_{j=1}^{N} \ln(1/n_j) - \sum_{j=1}^{N} \ln\{f(n_j)/n_j\}]\}$$

$$= 1 - \exp\{-(2/N)\sum_{j=1}^{N} \ln f(n_j)\}$$

$$= 1 - \exp\{-(2/N) \ln \prod_{j=1}^{N} f(n_j)\}$$

$$= 1 - [\prod_{j=1}^{N} f(n_j)]^{-2/N}$$

$$= 1 - \frac{1}{\tilde{f}^2}.$$

The following lemmas are adapted from Hu, Shao, and Palta (2006) for the conditional logit model (their proofs are for the multinomial logit model).

Lemma 1. *Suppose the independent variables x_{ij}, j = 1, 2, ... , N, i = 1, 2, ... , n_j, are independent and identically distributed random m-vectors with finite second moment (i.e., $E(x_{ij}{}^2)$ finite). Then $(1/\sqrt{N})[\ln L(\hat{\beta}) - \ln L(\beta)] \to_p 0$ as $N \to \infty$, where $\to_p$ denotes convergence in probability.*

Proof. We first prove that

(A4.9) $$\frac{\partial^2 \ln L(\beta)}{\partial \beta_u \partial \beta_v} = O_p(N).$$

Let

(A4.10)
$$\begin{aligned} F_{ij} &= \exp(\beta^T x_{ij}) \\ S_j^0 &= \sum_{i=1}^{n_j} F_{ij} \\ S_{ju}^1 &= \sum_{i=1}^{n_j} F_{ij} x_{iju} \\ S_{juv}^2 &= \sum_{i=1}^{n_j} F_{ij} x_{iju} x_{ijv} \end{aligned},$$

so that $p_{ij} = F_{ij} / S_j^0$. Then

(A4.11)
$$\frac{\partial F_{ij}}{\partial \beta_u} = x_{iju} F_{ij}$$

and

(A4.12)
$$\frac{\partial p_{ij}}{\partial \beta_u} = p_{ij}\left[x_{iju} - \frac{S_{ju}^1}{S_j^0} \right].$$

Hence

(A4.13)
$$\frac{\partial \ln L(\beta)}{\partial \beta_u} = \sum_{j=1}^{N}\left[\sum_{i=1}^{n_j} y_{ij} x_{iju} - \frac{S_{ju}^1}{S_j^0} \right],$$

so

(A4.14)
$$\frac{\partial^2 \ln L(\beta)}{\partial \beta_u \partial \beta_v} = \sum_{j=1}^{N}\left[\frac{S_{ju}^1 S_{jv}^1 - S_j^0 S_{juv}^2}{(S_j^0)^2} \right],$$

which is finite by assumption of finite second moment. Hence, each element of the second order derivative matrix is $O_p(N)$. The remainder of the proof is from Hu, Shao, and Palta (2006). Let $S_N(\beta) = \partial \ln L(\beta) / \partial \beta$ and $I_N(\beta) = E(J_N(\beta))$ be the score function and information matrix, respectively, where $J_N(\beta) = -\partial^2 \ln L(\beta) / \partial \beta \partial \beta^T$. Then it follows from the above that $I_N(\beta) = O_p(N)$. Hence, by a second order Taylor expansion

(A4.15)
$$\begin{aligned} \frac{1}{\sqrt{N}}[\ln L(\hat{\beta}) - \ln L(\beta)] &= \frac{1}{\sqrt{N}} S_N(\beta)^T (\hat{\beta} - \beta) - \frac{1}{2\sqrt{N}} (\hat{\beta} - \beta)^T J_N(\beta^*)(\hat{\beta} - \beta) \\ &= (\hat{\beta} - \beta)^T I_N(\beta)^{1/2} I_N(\beta)^{-1/2} \sqrt{N} \frac{J_N(\beta^*)}{2N^{3/2}} \sqrt{N} I_N(\beta)^{-1/2} I_N(\beta)^{1/2} (\hat{\beta} - \beta), \end{aligned}$$

where β^* is a vector between $\hat{\beta}$ and β. The asymptotic normality results of the maximum likelihood estimator gives $I_N(\beta)^{1/2}(\hat{\beta} - \beta) \to N(0,1)$. The lemma then follows from $I_N(\beta)^{-1/2}\sqrt{N} = O_p(1)$ and $J_N(\beta^*) / N = O_p(1)$.

Lemma 2. *Assume that the independent variables x_{ij}, j = 1, 2, … , N, i = 1, 2, … , n_j, are independent and identically distributed random m-vectors with finite second moment (i.e., $E(x_{ij}{}^2)$ finite). Let*

(A4.16) $$H_1 = E(\ln n_j),$$

$$H_2 = -E(\textstyle\sum_{i=1}^{n_j} p_{ij} \ln p_{ij}).$$

Then, as $N \to \infty$, $R_M^2 \to_p 1 - H_2 / H_1$, *and* $R_D^2 \to_p 1 - \exp[2(H_2 - H_1)]$.

Proof. Let $f_1(x) = 1 - x$ and $f_2(x) = (1/2)\ln(1-x)$, so $f_1^{-1}(x) = 1 - x$ and $f_2^{-1}(x) = 1 - \exp(2x)$. Then

(A4.17) $$\begin{aligned} f_1(R_M^2) &= \frac{(1/N)\ln L(\hat{\beta})}{(1/N)\ln L(0)} \\ &= \frac{(1/N)\ln L(\beta) - (1/N)[\ln L(\beta) - \ln L(\hat{\beta})]}{(1/N)\sum_{j=1}^{N} \ln(1/n_j)} \\ &= \frac{-(1/N)\sum_{j=1}^{N}\sum_{i=1}^{n_j} y_{ij} \ln p_{ij} + (1/N)[\ln L(\beta) - \ln L(\hat{\beta})]}{(1/N)\sum_{j=1}^{N} \ln n_j}. \end{aligned}$$

Similarly,

(A4.18) $$\begin{aligned} f_2(R_D^2) &= (1/N)\ln L(0) - (1/N)\ln L(\hat{\beta}) \\ &= (1/N)\sum_{j=1}^{N} \ln(1/n_j) - (1/N)\ln L(\beta) + (1/N)[\ln L(\beta) - \ln L(\hat{\beta})] \\ &= -(1/N)\sum_{j=1}^{N} \ln n_j - (1/N)\sum_{j=1}^{N}\sum_{i=1}^{n_j} y_{ij} \ln p_{ij} + (1/N)[\ln L(\beta) - \ln L(\hat{\beta})]. \end{aligned}$$

By the Law of Large Numbers, as $N \to \infty$,

(A4.19) $$(1/N)\sum_{j=1}^{N} \ln n_j \to_p H_1,$$

$$-(1/N)\sum_{j=1}^{N}\sum_{i=1}^{n_j} y_{ij} \ln p_{ij} \to_p H_2.$$

Moreover, from Lemma 1, as $N \to \infty$,

(A4.20) $$(1/N)[\ln L(\beta) - \ln L(\hat{\beta}] \to_p 0.$$

Hence, as $N \to \infty$, $f(R_M^2) \to_p H_2 / H_1$ and $f(R_D^2) \to_p H_2 - H_1$. So, by the Continuous Mapping Theorem, as $N \to \infty$, $R_M^2 \to_p f_1^{-1}(H_2 / H_1)$, i.e., $R_M^2 \to_p 1 - H_2 / H_1$, and $R_D^2 \to_p f_2^{-1}(H_2 - H_1)$, i.e., $R_D^2 \to_p 1 - \exp[2(H_2 - H_1)]$.

Proposition 2. *Assume that the independent variables* x_{ij}, $j = 1, 2, \ldots, N$, $i = 1, 2, \ldots, n_j$, *are independent and identically distributed random m-vectors with finite second moment* (*i.e.,* $E(x_{ij}^{\ 2})$ *finite*). *Let* H_1, H_2 *be given by* (A4.16) *and let*

(A4.21) $$\Sigma = \begin{pmatrix} Var(n_j) & \eta \\ \eta & \varepsilon \end{pmatrix},$$

$$g_1 = \frac{1}{\ln\lambda}\left(\frac{H_2}{\lambda\ln\lambda}, \quad 1\right)^T,$$

$$g_2 = \frac{2e^{2H_2}}{\lambda^2}\left(\frac{1}{\lambda}, \quad 1\right)^T,$$

where

(A4.22) $$\lambda = E(n_j),$$

$$\eta = E[n_j \sum_{i=1}^{n_j} p_{ij} \ln p_{ij}] + \lambda H_2,$$

$$\varepsilon = E[\sum_{i=1}^{n_j} p_{ij} (\ln p_{ij})^2] - H_2^{\ 2}.$$

Then, as $N \to \infty$,

(A4.23) $$\sqrt{N}[R_M^2 - (1 - H_2 / H_1)] \to_d N(0, \sigma_1^{\ 2}),$$

$$\sqrt{N}[R_D^2 - (1 - e^{2(H_2 - H_1)})] \to_d N(0, \sigma_2^{\ 2}),$$

where $\sigma_1^{\ 2} = g_1^{\ T}\Sigma g_1$ *and* $\sigma_2^{\ 2} = g_2^{\ T}\Sigma g_2$.

Proof. Define $Z_j = (n_j, W_j)$, where $W_j = \sum_{i=1}^{n_j} y_{ij} \ln p_{ij}$. Then the Z_j form an independent and identically distributed random sequence with $\mu = E(Z_j) = (\bar{n}, -H_2)$ and $Cov(Z_j) = \Sigma$. To see this, note that $Cov(n_j, n_j) = Var(n_j)$ and $Cov(n_j, W_j) = \eta$ follow immediately, and

(A4.24) $$\begin{aligned} Cov(W_j, W_j) &= E(W_j^{\ 2}) - E(W_j)^2 \\ &= E[(\sum_{i=1}^{n_j} p_{ij} \ln p_{ij})^2] - H_2^{\ 2} \\ &= E[(\sum_{i=1}^{n_j} p_{ij}) \sum_{i=1}^{n_j} p_{ij} (\ln p_{ij})^2] - H_2^{\ 2} \end{aligned}$$

$$= E[\sum_{i=1}^{n_j} p_{ij} (\ln p_{ij})^2] - H_2{}^2 .$$

By the Central Limit Theorem (in two dimensions),

(A4.25) $$\sqrt{N}(\bar{Z} - \mu) \to N(0, \Sigma) .$$

Let

(A4.26) $$\phi_1(x_1, x_2) = 1 + \frac{x_2}{\ln x_1} ,$$

$$\phi_2(x_1, x_2) = 1 - \exp[2(-\ln x_1 - x_2)] .$$

Applying the delta method with ϕ_1 and ϕ_2 to (A4.25) gives

(A4.26) $$\sqrt{N}[\phi_1(\bar{Z}) - \phi_1(\mu)] \to_d N(0, \nabla \phi_1(\mu)^T \Sigma \nabla \phi_1(\mu)) ,$$

$$\sqrt{N}[\phi_2(\bar{Z}) - \phi_2(\mu)] \to_d N(0, \nabla \phi_2(\mu)^T \Sigma \nabla \phi_2(\mu)) .$$

From Lemma 2, and since

(A4.27) $$\nabla \phi_1(x_1, x_2) = \left(-\frac{x_2}{x_1 (\ln x_1)^2} , \quad \frac{1}{\ln x_1} \right) ,$$

$$\nabla \phi_2(x_1, x_2) = \left(\frac{2}{x_1} e^{2(-\ln x_1 - x_2)} , \quad 2e^{2(-\ln x_1 - x_2)} \right) ,$$

this leads to the asymptotic normality results in (A4.23).

List of references

Ali, M. M. (1977). Probability and utility estimates for racetrack bettors. *The Journal of Political Economy*, 85(4), 803-815.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'Ecole Americaine. *Econometrica*, 21(4), 503-546.

Anderson, G., and Brown, R. I., (1984). Real and laboratory gambling, sensation-seeking and arousal. *British Journal of Psychology*, 75(3), 401-410.

Asch, P., Malkiel, B. G., and Quandt, R. E. (1982). Racetrack betting and informed behavior. *Journal of Financial Economics*, 10(2), 187-194.

Asch, P., Malkiel, B. G., and Quandt, R. E. (1984). Market efficiency in racetrack betting. *The Journal of Business*, 57(2), 165-175.

Avery, C., and Chevalier, J. (1999). Identifying investor sentiment from price paths: the case of football betting. *The Journal of Business*, 72(4), 493-521.

Avery, C., and Zemsky, P. (1998). Multidimensional uncertainty and herd behavior in financial markets. *The American Economic Review*, 88(4), 724-748.

Ayton, P., and Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness? *Memory & Cognition*, 32(8), 1369-1378.

Ayton, P., and Wright, G. (1994). Subjective probability: what should we believe? In G. Wright, and P. Ayton (Eds.), *Subjective Probability* (pp. 163-183). Chichester: Wiley.

Bacon-Shone, J., Lo, V. S. Y., and Busche, K. (1992). Modelling winning probability. *Research report 10*, Department of Statistics, University of Hong Kong.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3), 797-817.

Bar-Eli, M., Avugos, S., and Raab, M. (2006). Twenty years of "hot hand" research: review and critique. *Psychology of Sport and Exercise*, 7(6), 525-553.

Benter, W. (1994). Computer based horse race handicapping and wagering systems: a report. In D. B. Hausch, V. S. Y. Lo, and W. T. Ziemba (Eds.), *Efficiency of Racetrack Betting Markets* (pp. 183-198). London: Academic Press.

Betfair. (2011, June 29). *2011 Annual Report*. Retreived August 25, 2011 from http://corporate.betfair.com/investorrelations/~/media/Files/B/Betfair/pdf/annual-Report-2011.pdf

Bhattacharjee, S., and Moreno, K. (2002). The impact of affective information on the professional judgments of more experienced and less experienced auditors. *Journal of Behavioral Decision Making*, 15(4), 361-377.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *The Journal of Political Economy*, 100(5), 992-1026.

Bird, R., and McCrae, M. (1987). Tests of the efficiency of racetrack betting using bookmaker odds. *Management Science*, 33(12), 1552-1562.

Black, F. (1986). Noise. *The Journal of Finance*, 41(3), 529-543.

Bloomfield, R., O'Hara, M., and Saar, G. (2009). How noise trading affects markets: an experimental analysis. *Review of Financial Studies*, 22(6), 2275-2302.

Bolton, R. N., and Chapman, R. G. (1986). Searching for positive returns at the track: a multinomial logit model for handicapping. *Management Science*, 32(8), 1040-1060.

Bradley, I. (2003). The representative bettor, bet size, and prospect theory. *Economics Letters*, 78(3), 409-413.

Breiman, L. (1961). Optimal gambling systems for favorable games. *Proceedings 4th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 65-78.

Breslow, N. E., and Day, N. E. (1994). *Statistical methods in cancer research. Volume I – the analysis of case-control studies*. Lyon: International Agency for Research on Cancer.

Brown, W. O., and Sauer, R. D. (1993a). Does the basketball market believe in the hot hand? Comment. *The American Economic Review*, 83(5), 1377-1386.

Brown, W. O., and Sauer, R. D. (1993b). Fundamentals or noise? Evidence from the professional basketball betting market. *The Journal of Finance*, 48(4), 1193-1209.

Brown, L. D., D'Amato, R., and Gertner, R. (1994). Racetrack betting: do bettors understand the odds? *Chance*, 7(3), 17-23.

Bruce, A. C., and Johnson, J. E. V. (1992). Toward an explanation of betting as a leisure pursuit. *Leisure Studies*, 11(3), 201-218.

Bruce, A. C., and Johnson, J. E. V. (1995). Costing excitement in leisure betting. *Leisure Studies*, 14(1), 48-63.

Bruce, A. C., and Johnson, J. E. V. (2000). Investigating the roots of the favourite-longshot bias: an examination of decision-making by supply- and demand-side agents in parallel betting markets. *Journal of Behavioral Decision Making*, 13(4), 413-430.

Bruce, A. C., and Johnson, J. E. V. (2003). Investigating betting behaviour: a critical discussion of alternative methodological approaches. In L. Vaughan Williams (Ed.), *The Economics of Gambling* (pp. 224-246). London: Routledge.

Burns, P. (1985). Experience and decision making: a comparison of students and businessmen in a simulated progressive auction. In V. L. Smith (Ed.), *Research in Experimental Economics* (pp. 139-153). Greenwich: JAI Press.

Burns, B. D., and Corpus, B. (2004). Randomness and inductions from streaks: "gambler's fallacy" versus "hot hand". *Psychonomic Bulletin & Review*, 11(1), 179-184.

Busche, K. (1994). Efficient market results in an Asian setting. In D. B. Hausch, V. S. Y. Lo, and W. T. Ziemba (Eds.), *Efficiency of Racetrack Betting Markets* (pp. 615-616). London: Academic Press.

Busche, K., and Hall, C. D. (1988). An exception to the risk preference anomaly. *The Journal of Business*, 61(3), 337-346.

Cain, M., and Peel, D. (2004). The utility of gambling and the favourite-longshot bias. *The European Journal of Finance*, 10(5), 379-390.

Camerer, C. F. (1987). Do biases in probability judgment matter in markets? Experimental evidence. *The American Economic Review*, 77(5), 981-997.

Camerer, C. F. (1989). Does the basketball market believe in the 'hot hand'? *The American Economic Review*, 79(5), 1257-1261.

Camerer, C. F. (1998). Can asset markets be manipulated? A field experiment with racetrack betting. *The Journal of Political Economy*, 106(3), 457-482.

Camerer, C. F. (2001). Prospect theory in the wild: evidence from the field. In D. Kahneman, and A. Tversky (Eds.), *Choices, Values, and Frames* (pp. 283-300). New York: Cambridge University Press.

Campbell, J. Y., and Kyle, A. S. (1993). Smart money, noise trading and stock price behaviour. *The Review of Economic Studies*, 60(1), 1-34.

Ceci, S. J., and Liker, J. K. (1986). A day at the races: a study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General*, 115(3), 255-266.

Chadha, S., and Quandt, R. E. (1996). Betting bias and market equilibrium in racetrack betting. *Applied Finanical Economics*, 6(3), 287-292.

Cheng, S., and Stough, R. S. (2006). Location decisions of Japanese new manufacturing plants in China: a discrete-choice analysis. *The Annals of Regional Science*, 40(2), 369-387.

Cipriani, M., and Guarino, A. (2005). Herd behavior in a laboratory financial market. *The American Economic Review*, 95(5), 1427-1443.

Cipriani, M., and Guarino, A. (2008). Herd behavior and contagion in financial markets. *The B.E. Journal of Theoretical Economics*, 8(1), Art. 24.

Clotfelter, C. T., and Cook, P. J. (1993). The "gambler's fallacy" in lottery play. *Management Science*, 39(12), 1521-1525.

Cox, J. R., and Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. *Memory & Cognition*, 10(5), 496-502.

Crafts, N. F. R. (1985). Some evidence of insider knowledge in horse race betting in Britain. *Economica*, 52(207), 295-304.

Croson, R., and Sundali, J. (2005). The gambler's fallacy and the hot hand: empirical data from casinos. *Journal of Risk and Uncertainty*, 30(3), 195-209.

Dale, R. S., Johnson, J. E. V., and Tang, L. (2005). Financial markets can go mad: evidence of irrational behaviour during the South Sea Bubble. *The Economic History Review*, 58(2), 233-271.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *The Journal of Political Economy*, 98(4), 703-738.

Devenow, A., and Welch, I. (1996). Rational herding in financial economics. *European Economic Review*, 40(5), 603-615.

Dipboye, R. L., and Flanagan, M. F. (1979). Research settings in industrial and organizational psychology: are findings in the field more generalizable than in the laboratory? *American Psychologist*, 34(2), 141-150.

Dowie, J. (1976). On the efficiency and equity of betting markets. *Economica*, 43(170), 139-150.

Draper, N. R., and Smith, H. (1998). *Applied Regression Analysis*. New York: Wiley.

Durham, G. R., Hertzel, M. G., and Martin, J. S. (2005). The market impact of trends and sequences in performance: new evidence. *The Journal of Finance*, 60(5), 2551-2569.

Ebbeson, E. B., and Konecni, V. J. (1980). On the external validity of decision making research: what do we know about decisions in the real world? In S. Wallsten (Ed.), *Cognitive Processes in Choice and Decision Behavior* (pp. 21-45). Hillsdale, NJ: Erlbaum.

Edelman, D. (2007). Adapting support vector machine methods for horserace odds prediction. *Annals of Operations Research*, 151(1), 325-336.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1), 1-26.

Ellison, G., and Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica*, 77(2), 427-452.

Epstein, L. G., and Schneider, M. (2008). Ambiguity, information quality, and asset pricing. *The Journal of Finance*, 63(1), 197-228.

Falk, R., and Konold, C. (1997). Making sense of randomness: implicit encoding as a basis for judgment. *Psychological Review*, 104(2), 301-318.

Ferrell, W. R. (1994). Calibration of sensory and cognitive judgments: a single model for both. *Scandinavian Journal of Psychology*, 35(4), 297-314.

Festinger, L. (1953). Laboratory experiments. In L. Festinger, and D. Katz (Eds.), *Research Methods in the Behavioral Sciences* (pp. 136-172). New York: Holt, Rinehart and Winston.

Figlewski, S. (1979). Subjective information and market efficiency in a betting market. *The Journal of Political Economy*, 87(1), 75-88.

Fingleton, J., and Waldron, P. (1999). Optimal determination of bookmaker betting odds: theory and tests. *Trinity Economic Paper Series*, Technical Paper No. 96/9.

Franck, E., Verbeek, E., and Nüesch, S. (2010). Prediction accuracy of different market structures - bookmakers versus a betting exchange. *International Journal of Forecasting*, 26(3), 448-459.

Friedman, M., and Savage. L. J. (1948). The utility analysis of choices involving risk. *The Journal of Political Economy*, 56(4), 279-304.

Funder, D. C. (1987). Errors and mistakes: evaluating the accuracy of social judgment. *Psychological Bulletin*, 101(1), 75-90.

Gambling Commission (2010). Industry Statistics 2009/10. http://www.gamblingcommission.gov.uk/pdf/Gambling Industry Statistics 2009 2010 WEB - January 2011.pdf (last accessed: 15 February 2011).

Gandar, J. M., Zuber, R. A., and Johnson, R. S. (2001). Searching for the favourite-longshot bias down under: an examination of the New Zealand pari-mutuel betting market. *Applied Economics*, 33(13), 1621-1629.

Gao, X., and Lin, T. (2011). Do individual investors trade stocks as gambling? Evidence from repeated natural experiments. *Mimeo,* University of Hong Kong.

Gigerenzer, G. (2000). *Adaptive Thinking: Rationality in the Real World*. New York: Oxford University Press.

Gigerenzer, G., and Goldstein, D. G. (1999). Betting on one good reason. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 75-95). New York/Oxford: Oxford University Press.

Gilovich, T., Vallone, R., and Tversky, A. (1985). The hot hand in basketball: on the misperception of random sequences. *Cognitive Psychology*, 17(3), 295-314.

Gjerstad, S. (2005). Risk aversion, beliefs, and prediction market equilibrium. *Mimeo*, University of Arizona.

Golec, J., and Tamarkin, M. (1995). Do bettors prefer longshots because they are risk-lovers, or are they just overconfident? *Journal of Risk and Uncertainty*, 11(1), 51-64.

Golec, J., and Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *The Journal of Political Economy*, 106(1), 205-225.

Goodman, J. S. (1998). The interactive effects of task and external feedback on practice performance and learning. *Organizational Behavior and Human Decision Processes*, 76(3), 223-252.

Grant, A., and Johnstone, D. (2010). Finding profitable forecast combinations using probability scoring rules. *International Journal of Forecasting*, 26(3), 498-510.

Griffith, R. M. (1949). Odds adjustments by American horse-race bettors. *The American Journal of Psychology*, 62(2), 290-294.

Grossman, S. J., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3), 393-408.

Harrison, G. W., and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(4), 1009-1055.

Hausch, D. B., Ziemba, W. T., and Rubinstein, M. (1981). Efficiency of the market for racetrack betting. *Management Science*, 27(12), 1435-1452.

Henery, R. J. (1985). On the average probability of losing bets on horses with given starting price odds. *Journal of the Royal Statistical Society. Series A (General)*, 148(4), 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. Series A (General)*, 137(2), 227-230.

Hogarth, R. M. (1981). Beyond discrete biases: functional and dysfunctional aspects of judgmental heuristics. *Psychological Review*, 90(2), 197-217.

Hong, H., and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54(6), 2143-2184.

Hu, B., Shao, J., and Palta, M. (2006). Pseudo-R^2 in logistic regression model. *Statistica Sinica*, 16, 847-860.

Hurley, W., and McDonough, L. (1995). A note on the Hayek hypothesis and the favorite-longshot bias in parimutuel betting. *The American Economic Review*, 85(4), 949-955.

Hurley, W., and McDonough, L. (1996). The favourite-longshot bias in parimutuel betting: a clarification of the explanation that bettors like to bet longshots. *Economics Letters*, 52(3), 275-278.

Johnson, J. E. V., and Bruce, A. C. (1993). Gluck's second law: an empirical investigation of horserace betting in early and late races. *Psychological Reports*, 72(2), 1251-1258.

Johnson, J. E. V., and Bruce, A. C. (1998). Risk strategy under task complexity: a multivariate analysis of behaviour in a naturalistic setting. *Journal of Behavioral Decision Making*, 11(1), 1-17.

Johnson, J. E. V., and Bruce, A. C. (2001). Calibration of subjective probability judgments in a naturalistic setting. *Organizational Behavior and Human Decision Processes*, 85(2), 265-290.

Johnson, J. E. V., Jones, O., and Tang, L. (2006). Exploring decision makers' use of price information in a speculative market. *Management Science*, 52(6), 897-908.

Johnson, J. E. V., O'Brien, R., and Sung, M. (2010). Assessing bettors' ability to process dynamic information: policy implications. *Southern Economic Journal*, 76(4), 906-931.

Johnson, J. E. V., Schnytzer, A., and Liu, S. (2009). To what extent do investors in a financial market anchor their judgments excessively? Evidence from the Hong Kong horserace betting market. *Journal of Behavioral Decision Making*, 22(4), 410-434.

Jullien, B., and Salanié, B. (1994). Measuring the incidence of insider trading: a comment on Shin. *The Economic Journal*, 104(427), 1418-1419.

Jullien, B., and Salanié, B. (2000). Estimating preferences under risk: the case of racetrack bettors. *The Journal of Political Economy*, 108(3), 503-530.

Jullien, B., and Salanié, B. (2008). Empirical evidence on the preferences of racetrack bettors. In D. B. Hausch, and W. T. Ziemba (Eds.), *Handbook of Sports and Lottery Markets* (pp. 27-49). Amsterdam: North Holland.

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2), 263-291.

Kahneman, D., Slovic, P., and Tversky, A. (Eds.) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kelly, J. L. (1956). A new interpretation of information rate. *Bell System Technical Journal*, 35(4), 917-926.

Keren, G. B., and Wagenaar, W. A. (1985). On the psychology of playing blackjack: normative and descriptive considerations with implications for decision theory. *Journal of Experimental Psychology: General*, 114(2), 133-158.

Keren, G. B., and Wagenaar, W. A. (1987). Violation of utility theory in unique and repeated gambles. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3), 387-391.

Koehler, J. J. (1996). The base rate fallacy reconsidered: descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, 19(1), 1-53.

Kopelman, R. E., and Minkin, B. L. (1991). Toward a psychology of parimutuel behavior: test of Gluck's Laws. *Psychological Reports*, 68(2), 701-702.

Kvålseth, T. O. (1985). Cautionary note about R^2. *The American Statistician*, 39(4), 279-285.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6), 1315-1335.

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics*, 32(1), 23-43.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2), 311-328.

Laplace, P. (1825). *Philosophical essay on probabilities* (translated from the fifth French edition of 1825 by A. I. Dale). New York: Springer-Verlag.

Larkin, J., McDermott, J., Simon, D. P., and Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208(4450), 1335-1342.

Law, D., and Peel, D. A. (2002). Insider trading, herding behaviour and market plungers in the British horse-race betting market. *Economica*, 69(274), 327-338.

Lehrer, J. (2009). *The Decisive Moment*. Edinburgh: Canongate.

Lessman, S., Sung, M., and Johnson, J. E. V. (2007). Adapting least-square support vector regression models to forecast the outcome of horseraces. *The Journal of Prediction Markets*, 1(3), 169-187.

Lessman, S., Sung, M., Johnson, J. E. V., and Ma, T. (2011). A new methodology for generating and combining statistical forecasting models to enhance competitive event prediction. *European Journal of Operational Research*, forthcoming.

Levitt, S. D. (2004). Why are gambling markets organized so differently from financial markets? *The Economic Journal*, 114(495), 223-246.

Levitt, S. D., and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2), 153-174.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1974). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 551-578.

Liu, S., and Johnson, J. E. V. (2007). Searching for anchoring effects in a naturalistic environment: evidence from the Hong Kong horserace betting market. *The Journal of Gambling Business and Economics*, 1(1), 69-84.

MacLean, L. C., Ziemba, W. T., and Blazenko, G. (1992). Growth versus security in dynamic investment analysis. *Management Science*, 38(11), 1562-1585.

MacLean, L. C., Thorp, E. O., Zhao, Y., and Ziemba, W. T. (2010). Medium term simulations of the full Kelly and fractional Kelly strategies investment. In L. C. MacLean, E. O. Thorp, and W. T. Ziemba (Eds.), *The Kelly Capital Growth Investment Criterion: Theory and Practice* (pp. 543-561). Singapore: World Scientific.

Maddala, G. S. (1983). *Limited dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.

Marginson, D. (2010). Market *in*efficiency in person-to-person betting: examining 'known loser' insider trading on the exchanges. *Southern Economic Journal*, 76(4), 932-952.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (105-142). New York: Academic Press.

McFadden, D. (2001). Economic choices. *The American Economic Review*, 91(3), 351-378.

McGlothlin, W. H. (1956). Stability of choices among uncertain alternatives. *The American Journal of Psychology*, 69(4), 604-615.

Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, 54(1), 17-24.

Meulbroek, L. K. (1992). An empirical analysis of illegal insider trading. *The Journal of Finance*, 47(5), 1661-1699.

Metzger, M. A. (1985). Biases in betting: an application of laboratory findings. *Psychological Reports*, 56(3), 883-888.

Murphy, A. H., and Brown, B. G. (1984). A comparative evaluation of objective and subjective weather forecasts in the United States. *Journal of Forecasting*, 3(4), 369-393.

Mwabu, G., Ainsworth, M., and Nyamete, A. (1993). Quality of medical care and choice of medical treatment in Kenya: an empirical analysis. *The Journal of Human Resources*, 28(4), 838-862.

Northcraft, G. B., and Neale, M. A. (1987). Experts, amateurs, and real estate: an anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, 39(1), 84-97.

Ohtani, K. (2000). Bootstrapping R^2 and adjusted R^2 in regression analysis. *Economic Modelling*, 17(4), 473-483.

Oorlog, D. R. (1995). Serial correlation in the wagering market for professional basketball. *Quarterly Journal of Business and Economics*, 34(2), 96-109.

Orasanu, J., and Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, and C. E. Zsambok (Eds.), *Decision Making in Action: Models and Methods* (pp. 3-20). Chichester: Wiley.

Ottaviani, M., and Sørensen, P. N. (2005). The timing of bets and the favorite-longshot bias. *Mimeo*, London Business School and University of Copenhagen.

Ottaviani, M., and Sørensen, P. N. (2008). The favorite-longshot bias: an overview of the main explanations. In D. B. Hausch, and W. T. Ziemba (Eds.), *Handbook of Sports and Lottery Markets* (pp. 83-101). Amsterdam: North Holland.

Papachristou, G. (2004). The British gambler's fallacy. *Applied Economics*, 36(18), 2073-2077.

Paul, R. J., and Weinbach, A. P. (2005). Bettor misperceptions in the NBA: the overbetting of large favorites and the "hot hand". *Journal of Sports Economics*, 6(4), 390-400.

Peirson, J., and Blackburn, P. (2003). Betting at British racecourses: a comparison of the efficiency of betting with bookmakers and at the Tote. In L. Vaughan Williams (Ed.), *The Economics of Gambling* (pp. 30-42). London: Routledge.

Peirson, J., and Smith, M. A. (2010). Expert analysis and insider information in horse race betting: regulating informed market behavior. *Southern Economic Journal*, 76(4), 976-992.

Piron, R., and Smith, L. R. (1994). Testing risklove in an experimental racetrack. *Journal of Economic Behavior & Organization*, 27(3), 465-474.

Press, S. J., and Zellner, A. (1978). Posterior distribution for the multiple correlation coefficient with fixed regressors. *Journal of Econometrics*, 8(3), 307-321.

Preston, M. G., and Baratta, P. (1948). An experimental study of the auction-value of an uncertain outcome. *The American Journal of Psychology*, 61(2), 183-193.

Quandt, R. E. (1986). Betting and equilibrium. *The Quarterly Journal of Economics*, 101(1), 201-208.

Rabin, M. (2002). Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3), 775-816.

Rachlin, H. (1990). Why do people gamble and keep gambling despite heavy losses? *Psychological Science*, 1(5), 294-297.

Rosett, R. N. (1965). Gambling and rationality. *The Journal of Political Economy*, 73(6), 595-607.

Rosett, R. N. (1971). Weak experimental verification of the expected utility hypothesis. *The Review of Economic Studies*, 38(4), 481-492.

Sauer, R. D. (1998). The economics of wagering markets. *Journal of Economic Literature*, 36(4), 2021-2064.

Saunders, D. M., and Turner, D. E. (1987). Gambling and leisure: the case of racing. *Leisure Studies*, 6(3), 281-299.

Schnytzer, A., and Shilony, Y. (1995). Inside information in a betting market. *The Economic Journal*, 105(431), 963-971.

Schnytzer, A., and Shilony, Y. (2005). Insider trading and bias in a market for state-contingent claims. In L. Vaughan Williams (Ed.), *Information Efficiency in Financial and Betting Markets* (pp. 287-312). Cambridge: Cambridge University Press.

Schnytzer, A., and Snir, A. (2008). Herding in imperfect betting markets with insider traders. *The Journal of Gambling Business and Economics*, 2(2), 1-15.

Schnytzer, A., Lamers, M., and Makropoulou, V. (2010). The impact of insider trading on forecasting in a bookmakers' horse betting market. *International Journal of Forecasting*, 26(3), 537-542.

Shanteau, J. (1992). Competence in experts: the role of task characteristics. *Organizational Behavior and Human Decision Processes*, 53(2), 252-266.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, 71(3), 421-436.

Shiller, R. J. (1990). Market volatility and investor behavior. *The American Economic Review*, 80(2), 58-62.

Shin, H. S. (1991). Optimal betting odds against insider traders. *The Economic Journal*, 101(408), 1179-1185.

Shin, H. S. (1992). Prices of state contingent claims with insider traders, and the favourite-longshot bias. *The Economic Journal*, 102(411), 426-435.

Shin, H. S. (1993). Measuring the incidence of insider trading in a market for state-contingent claims. *The Economic Journal*, 103(420), 1141-1153.

Shleifer, A., and Summers, L. H. (1990). The noise trader approach to finance. *The Journal of Economic Perspectives*, 4(2), 19-33.

Shleifer, A., and Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 52(1), 35-55.

Sias, R. W. (2004). Institutional herding. *The Review of Financial Studies*, 17(1), 165-206.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99-118.

Smith, M. A. (2010). Exploring variations in odds bias across betting media. *The Journal of Gambling Business and Economics*, 4(1), 53-66.

Smith, M. A., Paton, D., and Vaughan Williams, L. (2006). Market efficiency in person-to-person betting. *Economica*, 73(292), 673-689.

Smith, M. A., Paton, D., and Vaughan Williams, L. (2009). Do bookmakers possess superior skills to bettors in predicting outcomes? *Journal of Economic Behavior & Organization*, 71(2), 539-549.

Smith, M., and Vaughan Williams, L. (2008). Betting exchanges: a technological revolution in sports betting. In D. B. Hausch, and W. T. Ziemba (Eds.), *Handbook of Sports and Lottery Markets* (pp. 403-418). Amsterdam: North Holland.

Smith, M. A., and Vaughan Williams, L. (2010). Forecasting horse race outcomes: new evidence on odds bias in UK betting markets. *International Journal of Forecasting*, 26(3), 543-550.

Snowberg, E., and Wolfers, J. (2010). Explaining the favorite-longshot bias: Is it risk-love or misperceptions? *The Journal of Political Economy*, 118(4), 723-746.

Snyder, W. W. (1978). Horse racing: testing the efficient markets model. *The Journal of Finance*, 33(4), 1109-1118.

Sobel, R. S., and Raines, S. T. (2003). An examination of the empirical derivatives of the favourite-longshot bias in racetrack betting. *Applied Economics*, 35(4), 371-385.

Spiwoks, M., Bizer, K., and Hein, O. (2008). Informational cascades: a mirage? *Journal of Economic Behavior & Organization*, 67(1), 193-199.

Sung, M., and Johnson, J. E. V. (2010). Revealing weak-form inefficiency in a market for state contingent claims: the importance of market ecology, modelling procedures and investment strategies. *Economica*, 77(305), 128-147.

Swidler, S., and Shaw, R. (1995). Racetrack wagering and the "uninformed" bettor: a study of market efficiency. *The Quarterly Review of Economics and Finance*, 35(3), 305-314.

Teebagy, N., and Chatterjee, S. (1989). Inference in a binary response model with applications to data analysis. *Decision Sciences*, 20(2), 393-403.

Terrell, D. (1994). A test of the gambler's fallacy: evidence from pari-mutuel games. *Journal of Risk and Uncertainty*, 8(3), 309-317.

Terrell, D. (1998). Biases in assessments of probabilities: new evidence from greyhound races. *Journal of Risk and Uncertainty*, 17(2), 151-167.

Terrell, D., and Farmer, A. (1996). Optimal betting and efficiency in parimutuel betting markets with information costs. *The Economic Journal*, 106(437), 846-868.

Thaler, R. H., and Johnson, E. J. (1990). Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Management Science*, 36(6), 643-660.

Thaler, R. H., and Ziemba, W. T. (1988). Anomalies: parimutuel betting markets: racetracks and lotteries. *The Journal of Economic Perspectives*, 2(2), 161-174.

Tuckwell, R. H. (1983). The thoroughbred gambling market: efficiency, equity and related issues. *Australian Economic Papers*, 22(40), 106-118.

Tversky, A., and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76(2), 105-110.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185(4157), 1124-1131.

Tversky, A., and Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297-323.

Vaughan Williams, L. (1999). Information efficiency in betting markets: a survey. *Bulletin of Economic Research*, 51(1), 1-39.

Vaughan Williams, L., and Paton, D. (1997). Why is there a favorite-longshot bias in British racetrack betting markets? *The Economic Journal*, 107(440), 150-158.

Vaughan Williams, L., and Paton, D. (1998). Why are some favourite-longshot biases positive and others negative? *Applied Economics*, 30(11), 1505-1510.

Veall, M. R., and Zimmerman, K. F. (1996). Pseudo-R^2 measures for some common limited dependent variable models. *Journal of Economic Surveys*, 10(3), 241-259.

Vlastakis, N., Dotsis, G., and Markellos, R. N. (2009). How efficient is the European football betting market? Evidence from arbitrage and trading strategies. *Journal of Forecasting*, 28(5), 426-444.

Wagenaar, W. A. (1988). *Paradoxes of Gambling Behaviour*. London: Lawrence Erlbaum Associates.

Wagenaar, W. A., and Keren, G. B. (1988). Chance and luck are not the same. *Journal of Behavioral Decision Making*, 1(2), 65-75.

Walls, W. D., and Busche, K. (2003). Breakage, turnover, and betting market efficiency: new evidence from Japanese horse tracks. In L. Vaughan Williams (Ed.), *The Economics of Gambling* (pp. 43-66). London: Routledge.

Wallsten, T. S., Budescu, D. V., Erev, I., and Diederich, A. (1997). Evaluating and combining subjective probability estimates. *Journal of Behavioral Decision Making*, 10(3), 243-268.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20(3), 273-281.

Weitzman, M. (1965). Utility analysis and group behavior: an empirical study. *The Journal of Political Economy*, 73(1), 18-26.

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *The Journal of Finance*, 54(2), 581-622.

Wilson, T. D., Houston, C. E., and Etling, K. M. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125(4), 387-402.

Wolfers, J., and Zitzewitz, E. (2006a). Five open questions about prediction markets. *Working Paper, National Bureau of Economic Research*.

Wolfers, J., and Zitzewitz, E. (2006b). Interpreting prediction market prices as probabilities. *Working Paper, National Bureau of Economic Research.*

Yaari, M. E. (1965). Convexity in the theory of choice under risk. *The Quarterly Journal of Economics*, 79(2), 278-290.

Yates, J. F. (Ed.) (1992). *Risk Taking Behaviour*. Chichester: John Wiley.