

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Price formation in parimutuel markets

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Management

in

Economics

at Massey University, Albany, New Zealand

Paul G Geertsema

2010

Copyright (C) 2010 Paul G Geertsema. The author asserts the moral right to be identified as the author of this work.

Abstract

Two types of betting are common in sports betting: *fixed odds* betting and *parimutuel* betting. In fixed odds betting, the payout conditional on winning is fixed once the bet is placed and is not affected by the placing of subsequent bets. By contrast, winning bettors in a parimutuel contest share pro-rata in the total betting pool. This means that the payout to winning bettors in a parimutuel contest depends not only on selecting the winning outcome, but also on the amounts bet by other bettors (which cannot be observed at the time a bet is placed). Therefore a parimutuel contest can be viewed as a game at the level of individual bettors. Existing models in the parimutuel literature explain the data by either assuming a single, representative bettor with certain risk preferences or by assuming that a number of risk neutral bettors compete strategically within a game theoretic framework. Our contribution is to construct a novel theoretical framework of parimutuel markets in which we model *both* strategic interaction *and* risk preferences at the level of individual insiders, in the presence of exogenous outsiders. We solve this model analytically for the optimal insider betting amount in a static symmetric Nash equilibrium. Using a new dataset of 1.6 million individual horse race bets in New Zealand from 2006 to 2009, we document a strong inverse linear relationship between our model-implied insider risk preferences and the strength of insider beliefs relative to outsiders. That is, as the strength of insiders' beliefs relative to that of outsiders decrease, implied risk sensitivity moves from risk averse to risk loving. At a level of insider beliefs congruent with actual performance in the data, average implied risk preferences are close to zero, that is, insiders are effectively risk neutral. While risk neutrality is a standard assumption in strategic interaction models of parimutuel betting, our study is the first to provide empirical support for this assumption. Finally, we document a strong relationship (not previously reported in the literature) between the average bet size and the average payout ratio, suggesting that bettors with inside information self-select by placing larger bets.

Contents

1	Summary	1
2	Literature review	6
2.1	Introduction	6
2.2	Why are betting markets important?	7
2.3	Parimutuel betting markets	8
2.4	Empirical regularities	9
2.4.1	Market efficiency	9
2.4.2	The favourite longshot bias	10
2.4.3	Factors correlated with the favourite longshot bias	12
2.4.4	Early betting	14
2.4.5	Late betting is smart betting	15
2.4.6	Rounding of odds	15
2.5	Theories of price formation	16
2.5.1	Overview	16
2.5.2	Representative agent models	16
2.5.3	Market structure	19
2.5.4	Strategic interaction	22
2.5.5	The current state of theory	28
2.6	Areas for further work	28
2.6.1	Extending strategic interaction models to encompass risk sensitive bettors	28
2.6.2	Endogenising the distinction between inside and outside bettors	28

2.6.3	Deriving additional testable implications from existing theoretical frameworks	29
2.6.4	Creating a richer set of empirical stylised facts	29
2.7	Next steps	29
3	Strategic interaction with risk preferences	31
3.1	Introduction	31
3.1.1	Motivation	31
3.1.2	Notation	33
3.1.3	Risk neutral model	33
3.2	Model	37
3.2.1	Payoff	37
3.2.2	Expectation	38
3.2.3	Variance	38
3.2.4	Utility	38
3.3	Equilibrium	39
3.3.1	Best response function (first order conditions)	39
3.3.2	Nash equilibrium	40
3.3.3	Admissibility and second order conditions	41
3.4	Discussion	43
3.4.1	Symmetry	43
3.4.2	No equilibrium zones	44
3.4.3	Insider beliefs	45
3.4.4	Equilibrium final market probabilities	46
3.4.5	Special case 1: Risk neutral insiders	48
3.4.6	Special case 2: Risk sensitive representative insider	49
3.4.7	Special case 3: A simplified model	49
3.4.8	Empirical statics	51

4	Empirical tests	52
4.1	Introduction	52
4.2	Data description and analysis	52
4.3	Insiders vs outsiders	53
4.3.1	Cut-off point	53
4.3.2	Payout ratio and average bet size	53
4.4	Race level data	55
4.5	Operational definition of model parameters	57
4.6	Empirical tests and results	58
4.6.1	Overview	58
4.6.2	Direct estimation of risk sensitivity	59
4.6.3	Estimation conditional on specified insider beliefs	60
4.6.4	Model vs empirical final market probabilities	64
5	Conclusion	68
6	Appendix	73
6.1	Maple code to verify derivation of equilibrium insider betting amounts	73
6.2	Average payout rate and average bet size by percentile	74
6.3	Statics based on an empirical approximation	76
6.4	Stata code for estimating reduced form final market probability (\hat{m}) .	78
6.5	Stata code for direct empirical testing of predicted final market probability	80
6.6	Stata code for indirect empirical testing of predicted final market probability	81

List of Figures

2.5.1 Final market probability m for $p = 0.6$ and $p = 0.8$ as a function of N	25
3.1.1 Amount bet by individual insiders for given q and N	35
3.1.2 Net payout of individual insiders for given q and N	36
3.4.1 Critical values of α as a function of q , assuming different values of b_0	44
3.4.2 Risk loving insider betting amounts on either horse as a function of q , with associated utility (utility scaled 20x)	45
3.4.3 Risk averse insider betting amounts on either horse as a function of q , with associated utility (utility scaled 20x)	46
3.4.4 Final market probabilities where $N \rightarrow \infty$ and $\tau = 0.2$	47
3.4.5 Insider betting with $N = 4$, $\tau = 0$ and $q = \frac{1}{2}$	50
4.3.1 Payout ratio (per percentile) against average bet (per percentile)	55
4.6.1 Histogram of implied α , at an insider/outsider cut-off point of $c = 100$ (Model compliant observations only)	59
4.6.2 Mean implied risk sensitivity α_c vs insider/outsider cut-off point c	60
4.6.3 Mean of model implied risk sensitivity α_q as a function of Δq	63
4.6.4 Estimated risk sensitivity α_q against Δq for different insider bet cut-off points	64
4.6.5 Regression coefficient γ against risk sensitivity α for different q	66

List of Tables

4.1	Data summary – horse racing bets from 1 Aug 2006 to 31 July 2009	52
4.2	Data summary – insiders vs outsiders	54
4.3	Race level data summary	56
4.4	Implied risk sensitivity $\alpha_{\Delta q}$ for a range of Δq	62
4.5	Best fit regression results for α and q	67
6.1	Bets and payouts by percentile	75
6.2	Empirical approximation of model final market probabilities	77

Chapter 1

Summary

Betting markets are of interest to economists because they are real-world markets which resemble simplified financial markets in many important ways. The similarities include uncertain future outcomes, observable trading and pricing, economically significant financial risk and the presence of potentially complex information and belief structures. At the same time, the outcomes in betting markets are directly observable and those outcomes are exogenous to the price formation process – simplifying features that are absent from most financial markets. As a consequence betting markets have become a popular venue for testing economic theories of price formation under conditions of uncertainty.

Sports betting markets can be classified as *fixed odds* markets, in which a fixed payout is made conditional on a specified outcome occurring; and *parimutuel* markets, in which the sum of all bets relating to an event is paid out pro-rata to successful bettors. It is common for bets on the outcome of team sports (such as football games) to be conducted in fixed odds betting markets. By contrast, bets on events in which the outcome is an ordinal ranking are often structured as parimutuel bets; horse racing being the prime example. While the payout on fixed odds betting is unaffected by other bets placed subsequently, this is not true for parimutuel bets – the payoff of a parimutuel bet depends also on the amounts bet on the event by other bettors (whether bet on the winning outcome or not). Crucially, at the time they place their own bets, bettors in a parimutuel contest are not able to observe the amounts simultaneously or subsequently bet by other bettors. Hence, at the level of individual bettors, parimutuel markets exhibit a strategic interaction element which is absent from fixed odds betting markets.

The early empirical literature dealing with parimutuel markets did not explicitly incorporate this strategic interaction element; instead, the focus was on the relationship between the probability distribution of outcomes implied by dollar betting

volumes and the actual empirical distribution of those outcomes. The genesis of this strand of the literature can be traced to Griffith (1949), in which it is shown that bettors in horse races exhibit a bias towards outcomes with low ex-ante probabilities and against outcomes with high ex-ante probabilities. The *favourite longshot bias*, as this has come to be known, is one of the more enduring anomalies in economics (as noted in Walls and Busche (2003)). The favourite longshot bias is by no means restricted to parimutuel markets, as it has also been widely documented in fixed odds betting markets. A large number of theories have been put forward to explain the favourite longshot bias; recent surveys can be found in Coleman (2004) and Ottaviani and Sørensen (2008).

Initial attempts to explain the favourite longshot bias theoretically made use of a representative agent framework in which a single representative agent is assumed to have certain risk preferences (Ali (1977) is a typical example of this approach). More recently, the focus has shifted towards modelling the strategic interaction in parimutuel betting markets explicitly as a game between many risk neutral bettors. A recent example of this approach is that of Ottaviani and Sørensen (2010). However, both approaches appear to be consistent with aggregate betting data.

Our contribution is to structure a theoretical model of parimutuel betting markets in which we incorporate both strategic interaction *and* bettor risk preferences at the individual bettor level. We believe this is a novel contribution to the literature¹. Our model postulates an arbitrary number of strategic insiders that compete in a simultaneous-move² Nash game in the presence of exogenously specified outsiders. Insiders have utility that allows for an arbitrary degree of risk sensitivity. We solve this model for the static symmetric Nash equilibrium in continuous strategies and obtain a closed-form solution to the optimal individual insider betting amount as well as the resulting implied “market prices” of event outcomes.

To test our model, we turn to a unique dataset of 1.6 million individual horse race bets in New Zealand from 2006 to 2009. In applying our model to the data, we

¹The only other work combining risk aversion and strategic interaction in a parimutuel setting (as far as we are aware) is that of Qiu (2007), in which the author demonstrates the existence of equilibrium in a sequential move framework with agents subject of prospect theory utility. In our model, by contrast, we consider a simultaneous move game between a number of insider bettors with a common arbitrary degree of individual risk aversion. We also explicitly solve for a closed form equilibrium solution, while Qiu is focused on demonstrating the existence and properties of equilibrium in his model.

²Betting normally opens a day or two before the start of the race. Bets can be placed at any time until closing time, and the total betting volume on each horse is publicly reported every 5 minutes or so throughout this time. Our decision to model insiders in a simultaneous game is motivated by a previous finding in the literature that around 40% of bets are placed in the last minute prior to closing and that last minute betting is a bettor predictor of final outcomes than earlier bets (Gramm and McKinney (2009)). These regularities are discussed more fully in our literature section.

need to operationalise our distinction between insiders and outsiders. We do so by classifying as insiders all bettors that bet NZD 100³ or more in a single race; the remaining bettors are classified as outsiders. Our classification rule is motivated by the observation that insiders as defined above outperform outsiders by 9.2% *ex-post* in our dataset, suggesting that bettors that bet more than NZD 100 do in fact bet on the basis of information not available to outsiders.

Based on this classification rule we are able to calculate proxy variables for all of our theoretical model parameters, save for inside information and insider risk preference. (Note that in the context of our model *inside information* denotes the difference in probabilities that insiders and outsiders assign to the winning outcome.) We then proceed to calculate the model implied risk sensitivity of insiders at different levels of inside information. The results reveal a very strong linear relationship between inside information and risk sensitivity – in effect, inside information and risk sensitivity could be considered close substitutes within our model, since the same insider betting behaviour can be induced by either assuming a sufficiently high level of risk sensitivity or a sufficiently high degree of inside information on the part of insider bettors. At a level of inside information consistent with the outperformance of insiders over outsiders observed in the data, the average implied risk sensitivity is very close to zero. This can be interpreted as evidence that insiders are risk neutral in deciding their bets on a race-by-race basis.

Previous literature applying game theory to parimutuel markets has *assumed* risk neutrality without any empirical justification. Given that in financial economics the assumption of individual risk aversion in decision making is pervasive, this is not a minor assumption. Our work suggests that, at a minimum, an assumption of risk neutrality on the part of insiders is not inconsistent with the data. This finding provides significant theoretical and empirical support to the growing body of literature seeking to explain parimutuel markets within a game theoretic risk neutral agent framework.

In addition, we find a strong positive linear relationship between the log of the average individual amount bet and the average payout ratio; a result not previously documented in the literature⁴. We find that the log of the average amount bet explains 74% of the variation in average payout rates when aggregated in percentiles by bet size. One plausible explanation for this regularity is that bettors who have inside information self-select by placing larger bets – an intriguing hypothesis that could form the basis for further research.

³As of 22 July 2010 1 NZD = 0.72 USD.

⁴This finding provides additional support for the classification rule we use to distinguish between insiders and outsiders.

The rest of the thesis is set out as follows: in chapter 2 we conduct a broad survey of the literature pertaining to parimutuel betting markets. Chapter 3 presents our theoretical model of parimutuel markets in which we combine individual risk preferences with strategic interaction. Chapter 4 introduces a new bet-level dataset and demonstrates that success in betting is systematically related to the size of the bet placed. We then proceed to test our hypothesis that insider bettors are risk neutral against the data using the theoretical results obtained in chapter 3. We find that individual bettors are effectively risk neutral conditional on a degree of inside information that is consistent with the observed data. Chapter 5 concludes. Relevant computer code and data tables are collected in the Appendix in chapter 6. A summary of commonly used variables follows; readers may find it useful to have this summary to hand while reading the remainder of the document.

Summary of commonly used variables

Variable	Explanation
N	Number of insider bettors
b_0	Fraction bet on horse 1 by exogenous outsiders
τ	Takeout rate
q	Belief of insiders regarding the probability of horse 1 winning
p	Objective probability of horse 1 winning
Δq	$= q - b_0$, or the amount of inside information available to insiders
α	Risk preference parameter (from insider utility function)
α_L	Lower bound on insider risk preference
α_H	Upper bound on insider risk preference
$\alpha_{\Delta q}$	Model implied risk preference conditional on $q = b_0 + \Delta q$
b_1	Amount bet by insider bettor 1
b_n	Amount bet by insider bettor n , for $n > 1$
b^*	The optimal amount for insider bettors to bet in equilibrium
π	Winning bet payoff for horse 1
w	Final period wealth of bettor
m	Final market probability of horse 1 winning
\hat{m}	OLS estimate of final market probability
\bar{m}	Empirical (observed) final market probability
$m_{(q,\alpha)}$	Model implied final market probability for given q and α
u	Utility of insider bettor
c	Cut-off point (in NZD) between insider and outsider bets
γ	Coefficient of OLS regression of \bar{m} on $m_{(q,\alpha)}$: $\gamma = 1 \Rightarrow \alpha = 0$

Chapter 2

Literature review

2.1 Introduction

Parimutuel betting markets, of which betting on horse races is the prime example, make up only a small fraction of total global economic activity. By way of example, the total amount bet on horse races in the US in 2008 amounted to c. \$14bn, of which around 80% was returned to bettors, resulting in gross revenue to the racing industry of slightly less than \$3bn¹. While this is not a small number, it pales into insignificance when considered in the context of US GDP of \$14.4 trillion² for the same year. This begs the question why these markets should be of interest to economists. It turns out that betting markets have a number of intrinsic features that make them particularly suited to testing economic theories of price formation, as have been noted by several contributors to the literature on betting markets (see for instance Sauer (2005) and Thaler and Ziemba (1988))³.

The purpose of this chapter is to conduct a broad review of the literature associated with price formation in parimutuel markets, with a view to identifying areas that might benefit from further theoretical development. To motivate this line of inquiry, in Section 2.2 we elaborate on the features of betting markets that make them particularly amenable to the study of price formation. Section 2.3 explains the operation of parimutuel markets before we turn to a review of empirical regularities that have been uncovered by researchers in Section 2.4. Attempts to explain these empirical regularities form the substance of Section 2.5, which is an introduction to

¹International Federation of horse racing Authorities, retrieved at 22 April 2010 from <http://www.horseracingintfed.com>. Assumes an exchange rate of 1 US dollar to 0.71 euros.

²US Bureau of Economic Accounts, Table 1.1.5, retrieved from <http://www.bea.gov> at 22 April 2010.

³Of course there is always the possibility that some economists may find horse racing intrinsically interesting. This author does not number among them. Instead, this author finds general theories of price formation to be intrinsically interesting.

the main theories of price formation in parimutuel markets. Section 2.6 presents an analysis of current gaps in the theoretical literature that could benefit from further work.

2.2 Why are betting markets important?

Understanding the way in which markets determine prices is of central importance in economic theory, since prices are inextricably linked to the production of goods and the allocation of resources. While not particularly important as a driver of general economic activity, betting markets do have significant potential to help economists form a clearer understanding of how markets work (a point that has been made by many authors – see for instance Thaler and Ziemba (1988), Sauer (1998), Vaughan Williams (1999), Coleman (2004) and Ottaviani and Sørensen (2008) for further elaborations). In particular, the following features of betting markets noted by the aforementioned authors are relevant:

1. Bets have a well-defined termination point which is fixed in advance and publicly known.
2. The outcomes can be ascertained with certainty and is publicly observable.
3. The payout conditional on winning is known in advance in the case of fixed-odds betting (but *not* in the case of parimutuel betting).
4. Outcomes are exogenous to the trading process and the prices that result from it. Therefore econometric problems with endogeneity can be avoided.
5. Parimutuel markets do not require the presence of a price-setting market maker, thereby eliminating a potential source of price distortion.
6. Though simpler, betting markets share important features with financial markets – potentially complex information structures, large numbers of (possibly heterogeneous) participants and the potential for differential information sets (aka insider trading). Therefore insights derived from examining betting markets may also have applicability to wider financial markets.

Also, in contrast to most experimental studies, betting market “experiments” are repeated on a daily basis, are large scale, involves significant amounts of money and are free from the Hawthorne effect since bettors do not generally consider themselves to be participants in an experiment.

However, betting markets do suffer from potential drawbacks as a test-bed for theories of price formation. First, as noted by Ottaviani and Sørensen (2008) and others, bettors may be motivated by recreational utility as opposed to monetary payoff, which makes it more difficult to model behaviour within the standard economic frameworks. Second, as discussed in the next section, parimutuel betting markets at the level of the individual bettor constitute a game between bettors in that payoffs depend not only on the decision of the individual bettor, but also on decisions by other bettors, who are in turn similarly affected. Hence the strategic interaction of bettors adds a layer of complexity to bettor's expectations that are not necessarily present in more conventional financial markets (or, indeed, in fixed-odds betting markets).

2.3 Parimutuel betting markets

Parimutuel (from the French *pari mutuel*, meaning mutual stake⁴) betting is a form of betting in which two or more *bettors* place bets on a number of mutually exclusive outcomes, such as the winning horse in a horse race. When the winning outcome is realised, the total amount originally bet - the *pool* - is returned to the winning bettors in proportion to their original bets, after deducting a certain percentage - the *takeout rate* - for taxes and expenses (typically in the range 15%-25%).

Mathematically, let $h \in \mathcal{H}$ be one of a set of horses in a race. Denoting by $b(h)$ the total amount bet on horse h and by B the total amount bet on all horses, the net payoff after deduction of the takeout rate τ to a unit bet on a winning horse is:

$$\pi(h) = (1 - \tau) \frac{B}{b(h)} - 1 = (1 - \tau) \frac{\sum_H b(h)}{b(h)} - 1$$

The invention of the parimutuel method of betting is credited to the French businessman Pierre Oller around 1870⁵. Since then it has become the dominant form of betting on the outcome of horse races in the US, Western Europe, the UK, Australasia and the Far East. In some jurisdictions, such as the UK and Australia, parimutuel betting on horse races coexists with fixed-odds betting, as noted in Sauer (1998). In other jurisdictions, parimutuel betting is mandated by law.

One notable feature of parimutuel betting is that the payoff to a bet depends not only on the outcome of the contest, but also on the bets placed by other players, either on the winning horse or on any of the other horses. Recall that winning bettors share the total pool of money bet (after subtracting the takeout), the size

⁴Merriam-Webster Online Dictionary (from <http://www.merriam-webster.com/dictionary/parimutuel> at 21 April 2010).

⁵Encyclopedia Britannica

of which depends on the total money bet on *all* of the horses. Furthermore, the proportion of the total pool that a winning bettor is entitled to depends not only on the size of her own bet, but also on the size of bets made by other winning bettors. This is because the total pool is distributed to winning bettors *in proportion* to their winning original bets.

This implies that payoffs to bettors in a parimutuel contest is strategically interdependent, being contingent not only on betting on the winning outcome, but also on the other bettor's bets. This is in contrast to fixed-odds betting where the outcome contingent on winning is known with certainty.

Note that, once the *final* distribution of bets is known, the payout on a winning parimutuel bet can be computed with ease. The key point, however, is that no bettor is privy to the final distribution of bets at the time she has to decide her bet - there is no practicable way to know with certainty how much other bettors will bet subsequently (or simultaneously, if our bettor is holding out until the last moment before betting closes).

2.4 Empirical regularities

2.4.1 Market efficiency

The concepts and methodologies of financial market efficiency (Fama, 1991) has also been applied to betting markets. Many researchers have found evidence inconsistent with market efficiency – Sauer (1998) and Vaughan Williams (1999) are surveys that discuss the topic in some depth. The favourite longshot bias, which we discuss in detail below, is often cited as evidence that sporting markets are not efficient. Despite this, there is a consensus in the literature that final market probabilities in parimutuel markets are generally good first-order predictors of actual outcomes.

For instance, Thaler and Ziemba (1988) writes:

The racetrack betting market is surprisingly efficient. Market odds are remarkably good estimates of winning probabilities. This implies that racetrack bettors have considerable expertise, and that the markets should be taken seriously.

but then goes on to admit the existence of apparent biases.

More recently, Sauer (1998) comments:

Although there are many published exceptions that reject efficient pricing, it is the benchmark result that prices are approximately efficient that yields insights to which economics should be proud to lay claim. For example, odds at the racetrack imbed subjective estimates of the probability of winning that are (a) quite close to their empirical counterparts, (b) similarly close to those obtained using sophisticated statistical methods, and (c) very difficult to exploit on a systematic basis.

Empirical testing of market efficiency in sports betting markets have been approached from a number of different angles. In discussing the literature, Sauer (1998) considers three approaches to testing for market efficiency:

1. Constant returns: Bets should all yield the same expected return.
2. Absence of profit opportunity: No way to make money in expectation.
3. Equilibrium pricing: Observed prices should match theoretical prices.

Thaler and Ziemba (1988) organises their discussion of market efficiency around the financial markets benchmarks of weak vs strong market efficiency. Weak market efficiency implies that bets should not have positive expected values while strong market efficiency implies that all bets should have expected values equal to the amount bet less the takeout.

Most of the literature on market efficiency in sports betting assumes that in an efficient market the market probability of an outcome (which can be inferred from the final market odds) should be an unbiased predictor of the observed outcome. This assumption will be true for some models of price formation and but not true for others. This leads us to an inescapable problem when testing for any form of market efficiency: any such test can only test the *joint* hypothesis that the market is inefficient *and* that the benchmark model used to define efficiency is correct. Therefore, any empirical rejection of efficiency may simply indicate that the benchmark model used is incorrectly specified, rather than proving that a particular market is inefficient (Cochrane, 2001 p128 citing Fama (1970)).

2.4.2 The favourite longshot bias

If one assumes that there is a single representative bettor which is risk neutral and a maximiser of expected utility over wealth, then it follows that market prices of

outcomes will be unbiased predictors of expected outcomes.⁶ To see why this must be the case, assume that there exists an outcome for which the market price is lower than the expected outcome. This implies that the representative bettor can increase his expected utility by betting a larger fraction of his money on this outcome. This process will continue until the market price of this outcome is equal to the expected outcome.

Therefore one might expect the fraction of money bet on each horse (its “price”) to be an unbiased predictor of the frequency of actual wins for that horse. However this is not the case. Griffith (1949) is credited with the discovery that horses with low odds (that is, favourites to win the race) are under-bet and horses with high odds (that is, longshots) are over-bet. In other words, favourites win more often than the final market odds suggest they should, while longshots win less often than the market odds suggest they should. This empirical regularity has been coined the *favourite longshot bias* (or sometimes the longshot bias or the FLB) and it holds a pride of place as one of the most robust anomalies in the literature on betting markets. It has even been described as “one of the most robust anomalous empirical regularities in economics” (Walls and Busche (2003) citing Thaler and Ziemba (1988)). The tendency to overestimate the likelihood of rare events also has support in the psychology literature where it is considered a subcategory of optimism bias.

In a survey of empirical papers Coleman (2004) found 19 studies addressing the longshot bias of which 15 found a statistically significant favourite longshot bias while a further 4 studies did not.⁷ So while the favourite longshot bias has been observed in several markets and over a time-frame of more than 50 years, it cannot be claimed to be universal. However, in a recent study employing a very large data set of 678,729 US races spanning 1992 to 2001 Snowberg and Wolfers (2010) continue to find strong evidence of the favourite longshot bias.⁸

More recently, empirical research has focused on identifying factors associated with the favourite longshot bias in the hope that it might shed more light on the underlying causes. (Sauer (1998) notes that work in this direction “would be particularly useful”.) Some of this work is examined below. It is perhaps also worth pointing out that the term favourite longshot *bias* presupposes that the discrepancy between

⁶The Law of Large Numbers ensures that the average empirical outcome will approach the mean of the distribution from which outcomes are drawn (in the limit). Therefore, over a sufficiently large number of trials the average observed outcome will be an unbiased estimate of the expected outcome.

⁷As will be discussed later, three of these studies considered races from Far East Asian countries, while the fourth used a small sample of 288 races.

⁸Just to confound matters further, some authors have documented a *reverse* favourite longshot bias in lotteries which are also commonly organised on a parimutuel basis.

market odds and empirical outcomes are the result of a bias. While early authors such as Griffith (1949) postulated bias on the part of bettors as an explanation for the discrepancy, there is no reason to suppose that the discrepancy must necessarily be the result of bettor bias – it could equally be the result of other factors such as loss aversion, strategic interaction or private information. We therefore use the term favourite longshot bias without prejudice as a shorthand for the discrepancy between market probabilities and empirical outcomes, as has become the practice in this literature.

2.4.3 Factors correlated with the favourite longshot bias

2.4.3.1 Pool size and number of horses

In a study of 10,000 races from North America, Japan and Hong Kong Busche and Walls (2000) find that the favourite longshot bias is inversely related to the size of the betting pool. For betting pools of the order of \$200 million to \$300 million, they do not find any favourite long shot bias - in other words, the market probabilities are unbiased predictors of empirical race outcomes. However, they do find a positive favourite longshot bias in races where the betting pool is in the range \$300,000 to \$1.8 million. They interpret this as evidence supporting decision cost theory, on the assumption that only large betting pools are able to generate sufficient profit opportunities to warrant the expense of obtaining superior information by sophisticated bettors. However, this evidence is also consistent with other explanations, such as inside information (it may be more difficult to exploit inside information in larger races) or the number of informed players (larger races may proxy for a greater number of sophisticated players, which forces the market price closer to the expected price).

In a similar paper that controls for additional race factors, Gramm and Owens (2005) also find support for the notion that larger betting pools are associated with a smaller favourite longshot bias. In addition, they find that the favourite longshot bias is decreasing in the number of horses in the race and for lower quality fields. These conclusions are robust to fixed track effects, for which they control.

On balance, the evidence does seem to indicate that “large” races (ie more dollars bet, more bettors, more horses) result in a smaller associated favourite longshot bias, even to the point where the bias becomes statistically insignificant.

2.4.3.2 Late races and capital constraints

In a paper examining 728 races held in 1978 in Atlantic City, Asch, Malkiel and Quandt (1982) finds that the favourite longshot bias is higher in races that are held later in the day.⁹¹⁰ They hypothesise that late races may proxy for capital constrained bettors – since bettors will, on average, lose the equivalent of the takeout rate on each race, there will be a large number of bettors who have lost a significant portion of their “race day money”. The idea is that these bettors would much prefer to end the day with a profit, but given the depleted state of their capital, the only way they can achieve this is by betting on a longshot horse with a high payout relative to the amount bet. Similarly, Ali (1977) found that a (representative agent) bettor at the last race of the day exhibited higher risk love than at the corresponding first two races of the day. However, Snowberg and Wolfers (2010) fail to find a last-race effect in their large scale study of US horse races from 1992 to 2001.

It appears that some bettors may be motivated by risk love or extreme loss aversion linked in some sense to the amount of money they bring to the races. This points to a potential Kahneman and Tversky (1979) prospect theory type loss aversion on the part of bettors, perhaps combined with mental accounting. (Also see Qiu (2007) for an attempt to incorporate prospect theory with mental accounting in a sequential parimutuel setting.)

2.4.3.3 Racing in East Asia

A number of studies have failed to find a favourite longshot bias. In his survey of previous empirical work, Coleman (2004) found four studies where a favourite longshot bias was not found, one of which (Swindler and Shaw, 1995) uses a small sample of 288 early season races in Texas. The remainder, which consists of c. 6000 races in Hong Kong ((Hausch and Ziemba, 1995) and (Busche and Hall, 1988)) and c. 13000 races in Japan (Walls and Busche, 2003), are both based on data from Far East Asian localities, which has prompted speculation that there may be cultural element at work.

⁹The effect is strong - for bets on favourites, the expected loss is 4%, while for bets on longshots, the expected loss is 69%. This can be compared to a takeout rate typically in the range 15%-25%, which is by definition equal to the dollar weighted average loss across all bets. The presence of such a high takeout rate means that it is usually not possible to profit from the observed favourite longshot bias despite the strength of the apparent bias.

¹⁰Coleman (2004) attributes the suggestion that the favourite longshot bias strengthens during the racing day to McGlothlin (1956).

2.4.4 Early betting

Most studies find that slightly over half of total betting dollars are wagered before closing time. In a study of 1644 races Gramm and McKinney (2009) find that 40% of all dollars are bet in the last minute before closing time, which implies around 60% must be wagered before then, often a day or more beforehand. In the case of fixed-odds betting, there may be a benefit to betting early if the odds offered by the bookmaker are sufficiently far away from the bettor's own estimate. If the odds offered by the bookmaker later converge to the bettor's own estimate, the bettor is still entitled to a payout at the fixed odds he originally contracted with the bookmaker in case the bet pays off. Therefore, it may be advantageous for an informed or sophisticated bettor to bet early if the fixed odds on offer from bookmakers are sufficiently different from his own estimate. The situation is slightly different in the case of parimutuel betting. A bettor in a parimutuel contest does not face a bookmaker with whom he can contract for a fixed payout – instead, the payout on a bet in the parimutuel system depends on the final amount bet on the winning horse as well as all the other horses. An informed or sophisticated bettor may therefore prefer to postpone betting to the last minute because

1. He cannot be sure of “locking in” what appear to be favourable odds based on current betting pools.
2. He gives up the opportunity of benefiting from future information that might become available.¹¹
3. He runs the risk of revealing his information to other bettors – and if other bettors also “get on the bandwagon” the effect of their additional bets would be to reduce the payout to the first bettor.

That is not to say that there are no advantages to betting early in a parimutuel race: Ottaviani and Sørensen (2006) shows in detail how sequential parimutuel betting corresponds to a Cournot competition in quantities, such that a Stackelberg first-mover advantage obtains, a correspondence also noted in passing by Adams, Rusco and Walls (2002).

There are other potential explanations of early betting, the most obvious being the possibility that bettors are motivated by non-monetary recreational utility, in which case they would be indifferent as to the exact timing of their bets.

¹¹This is mitigated by the fact that such future information is likely to alter the payoffs implicit in his bet so as to equalise returns across all bets.

2.4.5 Late betting is smart betting

If late betting were simply an implication of early betting as discussed above, then it would not warrant further discussion. But there is more to it: several studies have found that late betting moves market probabilities closer to the empirical probabilities. In other words, late betting eliminates some of the favourite longshot bias and results in final market probabilities that are closer to observed empirical outcomes – late betting is smart betting.

In one of the early studies, Asch, Malkiel and Quandt (1982) found that final odds were substantially better at predicting outcomes than the opening odds. In a study of 10,332 races in New Zealand, Gandar, Zuber and Dare (2000) finds that a favourite longshot bias is apparent in the early betting pools, but that the bias is mostly eliminated by late betting. Gramm and McKinney (2009) performs a similar analysis and finds that late betting volumes (in the last minute before closing) is the best predictor of empirical outcomes, even outperforming total betting volumes. The implication appears to be that late betting is largely free of bias and motivated to correctly assess winning probabilities, while early betting appears to be more susceptible to the favourite longshot bias and perhaps also more motivated by recreation than expected monetary reward.

2.4.6 Rounding of odds

A feature of most parimutuel markets is that the odds used to calculate payouts (which are based on the ratio between the total betting pool and the amount bet on the winning horse) are rounded downwards. The net effect of this is to reduce the amount of money paid out to winning bettors and to increase the amount of money retained by the organisers of the parimutuel market. Odds are typically rounded down to the nearest 5 or 10 cents. Walls and Busche (2003) point out that using rounded odds to calculate the favourite longshot bias results in a statistical bias towards rejecting market efficiency¹² – using exact betting volumes, they find the favourite longshot bias fails to materialise in their sample of c. 10,000 races. Since a number of the early studies used published (ie rounded) odds, it is possible that at least some of the reported favourite longshot bias might be attributable to the effect of rounding. However, as Coleman (2004) points out, the rounding effect has the greatest impact on favourites with short odds and the least impact on longshots with high odds. So while rounding biases published odds, it does so in a way that weakens the overall favourite longshot bias.

¹²This is because on favourites the payout odds will be less than suggested by the betting pools themselves. This will make it look like favourites were over-bet based on the payout odds.

2.5 Theories of price formation

2.5.1 Overview

The theoretical literature on parimutuel betting markets has developed in three strands. The first strand attempts to explain the favourite longshot bias as a result of certain attributes collectively inherent in the bettors themselves (within a representative agent framework). This includes bias or misperception, as well as risk-love or skew-love. We consider this literature in sub-section 2.5.2. The second strand, which we cover in sub-section 2.5.3, explains the favourite longshot bias by way of various market structure features. Examples include limited arbitrage and heterogeneous beliefs. Finally, the third strand explicitly models strategic interaction between individual bettors using the tools of game theory. In some of the models information is assumed to be public, while in other models, it is private to individual bettors. Sub-section 2.5.4 reviews this literature.

2.5.2 Representative agent models

Most of the early literature explained the favourite longshot bias by appealing to some aggregate characteristics of bettors, such as misperception or risk-love. As such we consider this segment of the literature under the heading of representative agent models. A weakness of this class of models is that a representative agent framework cannot incorporate strategic interaction within the context of parimutuel betting. Therefore, such models may confound the impact of strategic interaction (which is not modelled) with representative agent characteristics (which are).

2.5.2.1 Bettor bias

The early papers interpreted the discrepancy between market probabilities and empirical probabilities as evidence that bettors collectively suffer from bias (hence the choice of the term “favourite longshot *bias*” to describe the discrepancy). In the first paper to identify the favourite longshot bias, Griffith (1949) noted that, while bettors collectively managed to predict outcomes with a remarkable degree of accuracy, there remained a systematic under-estimation of the winning probabilities for favourite entries along with an over-estimation of the winning probabilities for longshot horses.

Remarkably, the debate whether the favourite longshot bias should be seen as evidence of collective bias or collective risk-love on the part of bettors rages on more

than 50 years later. Recently, Snowberg and Wolfers (2010) used a very large sample of North American races ($> 600,000$) to try and discriminate between these two competing explanations. Their approach centered around using straight win-bets to calibrate two simple models of decision making, the first corresponding to risk-love and the second corresponding to misestimation of probabilities. They then tested these calibrated models against data on compound bets¹³, and found that the compound bet data was more consistent with a misestimation of probabilities than risk-love.

2.5.2.2 Risk preference

In contrast with sub-section 2.5.2.1 (“Bettor bias”), Weitzman (1965) postulated a model of an *unbiased* representative bettor. While the representative bettor is not subject to bias, Weitzman postulates a utility function over betting outcomes for this bettor. Using data from 12,000 races during the period 1954 to 1963 at four New York racetracks, he estimates a utility curve which fits the observed aggregate betting data and concludes that the representative bettor is a risk-lover. Weitzman is careful to note that this does not imply that individual bettors are risk-loving (even in an average sense), only that bettors in a collective sense can be modelled closely by assuming they were represented by a hypothetical individual bettor with a risk-seeking utility over betting outcomes.

In a model with heterogeneous risk-loving bettors (in the mean-variance sense) Quandt (1986) constructs a model that exhibits the favourite longshot bias. His argument is based on the fact that, in contrast to standard mean-variance frameworks in the finance literature where risk aversion is assumed, risk-loving bettors *do not* achieve any benefit from diversification. Since they are risk-lovers, they would prefer to bet on a single horse, which would generally result in a higher variance than if they bet on a combination of horses. Quandt argues that, if market probabilities equalled empirical probabilities for each horse, this would result in a set of betting opportunities with different variances, but equal expected returns.¹⁴ However, such a set of betting opportunities cannot represent an equilibrium, since each bettor would prefer to bet on the single horse with the highest variance. Using a fixed-point argument, Quandt shows that in equilibrium, market probabilities will be above the empirical probabilities for favourites and below the empirical probabilities for longshots, hence giving rise to the favourite longshot bias.

¹³Compound (or exotic) bets are bets dependent on the performance of more than one horse. For instance, a *trifecta* bet pays off if the bettor correctly predicts the first, second and third placed horses in that order.

¹⁴This follows from the fact that market probabilities equal empirical probabilities

It is worth noting that while Quandt allows for an unspecified number of mean-variance optimising bettors, the way in which he solves for equilibrium effectively assumes a representative bettor and stops short of modelling the full strategic interaction between bettors in a parimutuel betting contest.

2.5.2.3 Skew preference

The presence of a positive track takeout rate implies that it is not possible for all bettors to be motivated by the average payout only. In sub-section 2.5.2.2 above, an attempt was made to explain bettor behaviour by linking their utility to the variance of payouts, the second moment. Golec and Tamarkin (1998) goes further by arguing that it is an affinity not for variance, but the third moment, skewness, that motivates bettors. They support their argument with a regression study showing that a specification that includes skewness provides a better description of the data than one based on variance only.

2.5.2.4 Prospect theory

A number of authors (see for instance Thaler and Ziemba (1988)) has noted that the behaviour of bettors appear to be consistent with the prospect theory of Kahneman and Tversky (1979). The key components of prospect theory is that people 1) overweight small probabilities and underweight large probabilities, 2) utility is defined separately over losses and gains, rather than total wealth and 3) the utility function is specified so that people exhibit risk-aversion for gains and risk-love for losses, with the added constraint that the utility function is flatter for gains than for losses, implying that people are loss-averse.

A recent attempt to incorporate prospect theory is due to Qiu (2007), which we discuss more fully in sub-section 2.5.4.4 since it also incorporates strategic interaction.

2.5.2.5 Recreational utility

Many papers implicitly acknowledge the existence of a class of bettors motivated by non-monetary or recreational utility. Typically such bettors are modelled as exogenous. For instance Isaacs (1953) considers a situation where a single informed bettor have to bet optimally given previous betting by exogenous bettors, while Ottaviani and Sørensen (2006) is a more recent example.

Hausch, Ziemba and Rubinstein (1981) considers a model in which some bettors derive recreational utility from making successful bets on longshots (ie “bragging

rights”), and show that the presence of such bettors implies a favourite longshot bias.

In contrast, Conlisk (1993) models recreational bettors endogenously by appending a “tiny utility of gambling” to an otherwise standard expected utility model, which allows him to explain the apparent risk seeking behaviour of gamblers and bettors within an otherwise risk averse framework. Ottaviani and Sørensen (2010) incorporates recreational utility within the framework of a simultaneous move game with private information. Participation of bettors is endogenous in this model, and bettors only obtain recreational utility if they choose to bet. The authors show that even bettors without strong private beliefs will bet on a horse if recreational utility is sufficiently high and that the favourite longshot bias becomes more pronounced as recreational utility decreases.

2.5.3 Market structure

A number of theoretical models of parimutuel betting has been inspired by assuming different market features. This stands in contrast to the earlier literature, which tried to explain observed betting probabilities by appealing to bettor misperception or attitudes towards risk within a representative agent setting. The key argument of the market structure literature is that the favourite longshot bias can arise as a result of specific market microstructure, without having to make assumptions about the characteristics of a representative bettor.

2.5.3.1 Single insider

The first explicit consideration of inside information is due to Isaacs (1953), who postulated a model with exogenous uninformed bettors (outsiders) and a single risk neutral informed bettor who bets optimally in response to the previous bets of the outsiders. In this model the insider knows the objective (true) win-probabilities of a set of H horses and also observes the amount bet by the outsiders. Assuming that no other bettors will bet subsequently, Isaacs solves for the insider’s optimal betting schedule. He also shows that if one ranks all horses by the ratio of their objective win probability to their outsider market probability (that is, by the ratio of $p(h)$ to $\frac{b_0(h)}{B_0}$) then if the insider bets on any horse j he will also bet on all the other horses ranked above j .

A consequence of equilibrium for a single informed bettor is that the final market probabilities (which include bets from both the outsiders and the insider) will fall short of the objective market probabilities for those horses that the insider bets

on, giving rise to the favourite longshot bias. The reason is that if the insider bets money until market probabilities equal objective probabilities, then the insider will make zero profits in expectation. Therefore, in order to profit from her inside information, the insider has to bet *less* than would be required to equate the final market probabilities with the objective probabilities. Since the final market probabilities are less than the objective probabilities in equilibrium, the favourite longshot bias obtains¹⁵.

2.5.3.2 Heterogeneous beliefs

It is reasonable to believe that bettors may have different subjective estimates regarding the winning probabilities of different horses. In a model with two horses and a continuum of risk neutral infinitesimally small bettors with heterogeneous beliefs, Ali (1977) shows that the mean of the distribution of bettors beliefs (which are equal the objective probability by assumption), is above the final market probability for the favourite horse and below the final market probability for the longshot, which is in accordance with the favourite longshot bias.

2.5.3.3 Endogenous bookmaker

A separate strand of literature investigates the impact of bookmakers on the price formation process. This line of inquiry is not directly applicable to parimutuel betting markets, since in a parimutuel market there is no bookmaker – bettors effectively bet against each other only and the parimutuel betting market operator (often referred to as the *track operator*) takes as compensation a fixed proportion of the total amount bet *without having to take any risk on the outcome of the race*. In contrast, bookmakers offering fixed-odds betting do take risk, at least to the degree that they are unable to balance betting interests evenly. We include a discussion of bookmakers because it shows clearly how the presence of a bookmaker can complicate the price setting process, and therefore further motivates the study of parimutuel markets in which this distortion is not present.

Shin (1991) and Shin (1992) provide a model in which a bookmaker posts fixed-odds prices against an uninformed public as well as an unknown insider. The uninformed public has heterogeneous beliefs distributed uniformly from zero to one, which are not correlated to the objective winning probabilities. The bookmaker has knowledge of the objective winning probabilities of all the horses. However, the insider has

¹⁵Interestingly, Isaacs (1953) fails to make any mention of the favourite longshot bias in his paper and appears to have been unaware of it. The fact that his equilibrium for a monopolist insider implies a favourite longshot bias nonetheless is therefore telling.

knowledge of the actual winning horse and bets her total wealth on this horse. The bookmaker always loses the bet contracted with the insider, but can generate profit by exploiting the mispricing of the uninformed public. Shin shows that in order to maximise expected profit in the presence of inside information, the bookmakers will post odds incorporating market probabilities that are different from the objective probabilities that the market maker is privy to. This relationship is given by a “square root rule”:

$$\frac{m(1)}{m(2)} = \sqrt{\frac{p(1)}{p(2)}}$$

The favourite longshot bias is inherent in this pricing rule as the favourite horse is quoted at a market probability below its objective probability, while the longshot horse is quoted at a market probability above its objective probability.

In a parimutuel race the track operator does not set prices, but it does have control over a choice variable that impacts its profitability - the track takeout rate¹⁶. Watanabe, Nonoyama and Mori (1994) puts forward a model in which the track takeout rate is endogenised. They consider a fixed number of bettors with heterogeneous *a-priori* beliefs which are public knowledge. The bettors play a simultaneous move game parameterised by the takeout rate in stage 2. In stage 1, the track operator chooses a takeout rate that will maximise his profit conditional on the beliefs of the bettors. By solving simultaneously for the outcome of the bettor’s game and the optimal choice of takeout rate for the track operator, Watanabe, Nonoyama and Mori are able to numerically characterise the dependence of the track takeout rate on bettor’s beliefs.

In a setting with a fixed-odds bookmaker, Levitt (2004) shows that an optimising bookmaker has an incentive to quote odds inconsistent with objective probabilities if faced with a betting public prone to predictable bias. The quoted odds results in betting volumes which are not balanced across outcomes, so that the bookmaker takes a net betting position against the public in a direction that exploits bettor biases. Levitt provides empirical evidence that bookmakers do in fact deviate from the “balanced books” model and that such a deviation earns them profits in excess of what they would have earned if they were seeking a balanced book.

2.5.3.4 Transaction costs and limits to arbitrage

In real-life parimutuel betting transaction costs (in the form of the takeout rate and downwards rounding of odds) is large, typically ranging from 15% to 25% (Thaler

¹⁶This is not always the case, as in some jurisdictions the takeout rate is fixed by law while other jurisdictions impose upper limits.

and Ziemba, 1988). Nor is it possible to directly “short” a horse if you think it is over-bet (ie overpriced)¹⁷. Taken together, this implies that arbitrage by sophisticated bettors may not be able to eliminate biases altogether, even if sophisticated bettors exhaust all profitable betting opportunities. It is notable, that despite the widespread occurrence of the favourite longshot bias, it has generally not been of a sufficient magnitude to present consistently profitable betting opportunities. Since the favourite longshot bias is economically significant, the lack of profitable arbitrage opportunities seems to be driven at least in part by high takeout rates.

Hurley and McDonough (1995) construct a model of insiders betting competitively to exploit the mispricing of a group of exogenous outsiders. While the insiders have access to inside information that should allow them to benefit at the expense of outsiders, they face high transaction costs in the form of a takeout rate and the generation of costly inside information. If transaction and information costs are equal to zero and the number of insiders approach infinity, final market probabilities approach objective probabilities. In the presence of positive transaction or information costs, this becomes an inequality with the objective probability forming the upper bound on the final market probability for the favourite horse, consistent with the favourite longshot bias. Hurley and McDonough then proceed to test their hypothesis that the bias is generated by the takeout rate in a laboratory experiment, in which the hypothesis is rejected. One is left with the conclusion that either the laboratory experiment does not generalise to the race track or the track takeout rate is not in itself an explanation for the existence of the longshot bias.

2.5.4 Strategic interaction

As noted in section 2.3, parimutuel betting comprises a *game* from the point of view of individual bettors in the sense that the optimal strategies of individual bettors depend on the optimal strategies of other bettors and *vice versa*. Despite the fact that the game theoretic nature of parimutuel betting has long been recognised¹⁸, the bulk of early theoretical models were constructed within a representative agent framework. This precluded any modelling of the game-theoretic interaction between bettors (see sub-section 2.5.2).

Another, more recent, literature models individual bettors but assume the number of bettors is so large and their individual bets so small that the impact of individual

¹⁷Although it is possible to create a position that might mimic a short interest in a particular horse by betting (probability weighted) amounts on all the other horses.

¹⁸(Ottaviani and Sørensen, 2008) cites a paper by Borel (1938) in which several key game-theoretic ideas relating to parimutuel betting are presented. This is even more remarkable considering that Borel’s paper pre-dates the development of modern game theory (formalised by von Neumann and Nash) by several decades.

decisions can be ignored. However, this approach is not without criticism, as Feeney and King (2001) point out:

Much of the theoretical work analysing parimutuel systems, however, makes a critical ‘small player’ assumption that one individual cannot influence the actions of others. In other words, there are always enough players in the system so that the effect of one player’s actions on the information and returns of other players can be ignored. For example, Potters and Wit (1995) examine a parimutuel system where players independently choose actions after receiving an individual signal that is drawn from a common distribution. But each player ignores the consequences of their action on the odds of the gamble. Similarly, Watanabe (1997) analyses a parimutuel system with a continuum of players. The small player assumption greatly simplifies the modelling of parimutuel systems, but it is very strong. For example, if there are few players or if some players wager relatively large sums of money then the interdependency of returns in a parimutuel system means that the ‘small player’ assumption is likely to be violated.

In this section we consider the strand of the literature that explicitly models the strategic interdependence between bettors inherent in the parimutuel method of betting. Our discussion follows the standard distinction in the game theoretical literature between simultaneous games and sequential games.¹⁹ We also make a distinction between games where bettors possess private information and those where information are common to bettors.

2.5.4.1 A note regarding rational risk neutral bettors

In a simultaneous parimutuel betting model that consists only of rational risk neutral bettors, no bets will take place if the takeout rate is non-zero. This follows from the fact that all bets lose money in expectation in such a market, and therefore rational risk neutral bettors will not place bets in the first place. Betting by rational risk neutral bettors can nonetheless be motivated if we assume some bettors are motivated by recreational utility, or if we assume some bettors have beliefs or inside information that differs from the implied market probabilities implicit in the existing betting odds they face. Most models proceed by postulating two groups of bettors -

¹⁹A simultaneous game is a game where the players have to decide their actions without having the benefit of observing the actions of other players. Therefore, it is not required that actions take place at the same time, only that those actions are not observable by other players. In a sequential move game, by contrast, players have the opportunity to observe the actions of previous players before having to decide their own actions.

outsiders that bet for recreational reasons and insiders that bet on the basis of inside information in order to take advantage of the mispricing of recreational bettors. In these models outsiders are for the most part modelled exogenously, so any strategic interaction takes place between insiders.

2.5.4.2 Simultaneous move – symmetric information

There are two ways to think about insiders' private information: 1) information that is private to the group of insiders but are common knowledge among those insiders (symmetric information) and 2) information that is strictly private to each individual in a group of insiders (asymmetric information). The first is addressed in this sub-section, while the second is addressed in the next sub-section.

In an early paper Eisenberg and Gale (1959) presents a model of m bettors and n horses, with each bettor's win-probability beliefs over horses modelled as a $m \times n$ matrix. They show that, if bettors are constrained by an arbitrary budget – which may differ between bettors – then there exists a unique Nash equilibrium of final market probabilities (but not necessarily betting amounts). In their own words “Thus, at equilibrium the bettors, as a group, maximize a weighted sum of logarithms of subjective expectations, the weights being the bettor's budgets. As noted previously, equilibrium probabilities turn out to be unique, although equilibrium bets need not be unique.” (Eisenberg and Gale, 1959, p168)

Hurley and McDonough (1995), previously discussed in some detail in sub-section 2.5.3.4 in the context of transaction costs, is an early paper explicitly seeking to explain the favourite longshot bias within a game theoretic framework using two horses and N insiders. They show that, as the number of insiders $N \rightarrow \infty$, the final market probability of the favourite horse becomes $m = (1 - \tau)p$. In other words, the market probability of the favourite will be less than the objective probability – the favourite longshot bias.

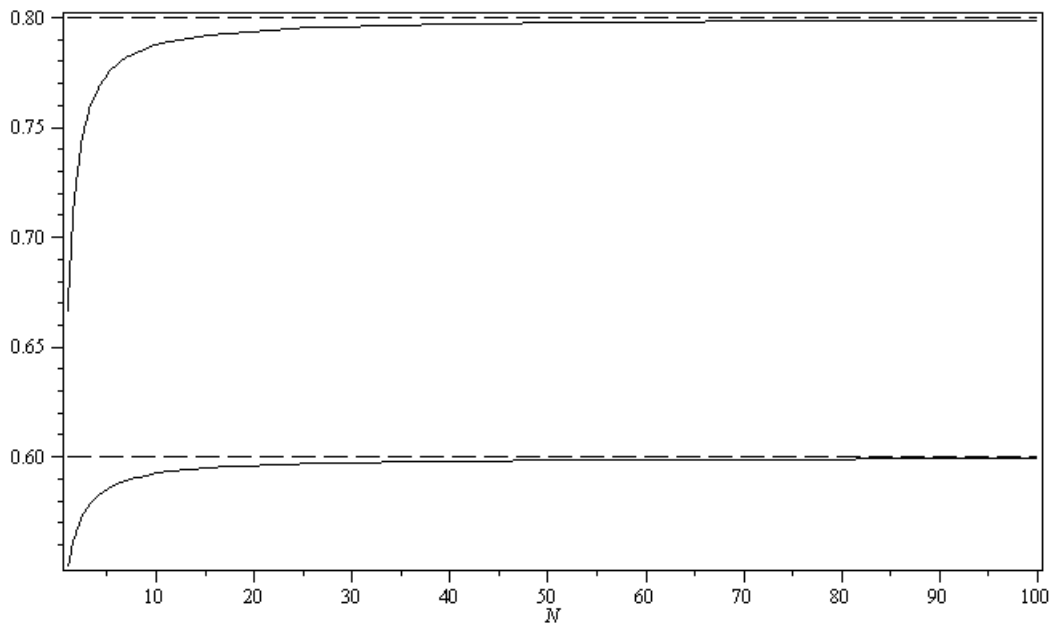
Although Hurley and McDonough neglects to mention this, their model is also able to explain the favourite longshot bias *without* an appeal to transaction costs. We take as a starting point the model set out in sub-section 3.1.3, which is equivalent to the Hurley and McDonough model in which we assume a zero takeout rate. Consistent with Hurley and McDonough we further assume $b_0 = 0.5$. Making use of the simplifying assumptions above and substituting the optimal insider betting amount (equation 3.1.4) into the formula for the final market probability of the winning horse (equation 3.1.6), we obtain a closed form solution to the final market probability

$$m = \frac{\sqrt{p(pN^2 - 6pN + 4N + p)} + pN - p}{\sqrt{p(pN^2 - 6pN + 4N + p)} - pN + 2N - p} \quad (2.5.1)$$

In the absence of a takeout rate, we have $p \geq b_0 = 0.5$. It can be shown that $m = p$ in equation 2.5.1 above is solved by only three values of p , that is, $[0, 0.5, 1]$. The first solution can be disregarded since it violates the inequality above, while the remaining two are corner solutions for the feasible range of p .

It can be shown that $m < p$ for all p in the range (0.5..1) for finite N , which is again consistent with the favourite longshot bias, since m and p refer to the final market implied probability and objective probability respectively of the favourite horse. For a small number of insiders N , the difference between the final market probability m and the objective probability p is substantial. As shown earlier, for large N , the final market probability converges to the objective probability. To give a sense of the relationship between m and p as N becomes larger, we plotted the curves corresponding to equation 2.5.1 for $p = 0.6$ and $p = 0.8$ in Figure 2.5.1 below.

Figure 2.5.1: Final market probability m for $p = 0.6$ and $p = 0.8$ as a function of N



It can therefore be argued that the favourite longshot bias may equally well be the result of a small number of insiders able to bet sizeable amounts of money.²⁰ The above exposition can be viewed as an extension of the Isaacs (1953) model of a single monopolistic insider to a model of finitely many insiders, and the same

²⁰We are assuming a zero takeout rate. More realistically, in the presence of a positive takeout rate, the difference between insiders assessment of winning probabilities and the probabilities implicit in existing betting odds needs to be sufficient to generate profits in expectation for insiders, otherwise insiders won't bet.

general results obtain. Chadha and Quandt (1996) constructs a more general model along the same lines in which they model a finite number of insiders betting on a finite number of horses in the presence of exogenously specified outsiders. The insiders have access to the objective win-probabilities of all the horses. The insiders attempts to bet optimally, given the expectation that other insiders will also bet optimally, gives rise to a Nash equilibrium. Chadha and Quandt are only able to solve for equilibrium final market probabilities numerically and arrive at similar results. Another paper taking a similar approach is that of Adams, Rusco and Walls (2002), again arriving at similar conclusions.

2.5.4.3 Simultaneous move – private information

Situations where insiders have asymmetric information are somewhat more complicated than situations where the inside information is symmetric to insiders. In the case of symmetric information, insiders can fairly easily form expectations regarding other insiders' actions in equilibrium, since all actions are conditioned on the same information set, which is known to all insiders. When insiders have asymmetric information, it is necessary for each insider to consider the *distribution* from which other insiders obtain their private information. This is to enable the formation of expectations regarding the private information of those other insiders. Koessler, Noussair and Ziegelmeyer (2008) is credited²¹ with being the first to integrate asymmetric information into a game theoretic formulation of parimutuel betting. The Koessler, Noussair and Ziegelmeyer model postulates N bettors that bet based on private binary signals, first in a simultaneous move setting and then in a predetermined sequential move setting. Among other results, Koessler, Noussair and Ziegelmeyer proves the existence of separating equilibria in which all bettors follow their private signals within a Bayesian-Nash framework, which yields final market probabilities consistent with the favourite longshot bias. They also show that such an equilibrium is unique if the number of bettors is large enough. However, separating equilibria are shown to disappear in a sequential move setup if the number of sequential betting periods grows large enough. In a simultaneous move model with a continuum of insiders and real valued private signals Ottaviani and Sørensen (2005) and Ottaviani and Sørensen (2009) also derives a favourite longshot bias, which is driven to zero as the number of insiders grows large.

In an ambitious undertaking, Ottaviani and Sørensen (2010) presents a model of parimutuel betting that incorporates k outcomes, N ex-ante identical bettors with a common prior belief distribution over all outcomes (the symmetric information component), independently drawn private signals (the asymmetric information compo-

²¹See Ottaviani and Sørensen, 2006, p5

ment) and fixed recreational utility which is dependent on participation. Each bettor can choose to bet an amount normalised to unity on a single outcome, or to abstain, thereby forfeiting their recreational utility. Unlike other contributions to the literature, the distinction between recreational bettors and insiders is endogenised – some bettors with poor private signal realisations will nonetheless bet in order to realise recreational utility, while other bettors with strong private signal realisations will bet in order to monetise the value of their private information. The model is able to simultaneously account for a number of known empirical regularities

- The favourite longshot bias in the presence of private information *and* the reverse favourite longshot bias as documented in lotteries (where private information is assumed lacking).
- Late betting being more informative
- The occurrence of the favourite longshot bias in both parimutuel and fixed-odds betting markets.

In addition the theory posits a number of empirically testable observations; a stronger favourite longshot bias is predicted when there are many bettors, bettors have more private information, there are fewer outcomes, recreational utility decreases or if the takeout rate increases.

2.5.4.4 Sequential move

In sequential move games, players commit to their actions in a predefined order. Previous actions are observable by players. This means that, unlike a simultaneous move game, players reveal information via their actions to subsequent players and need to incorporate this effect into their decision making process.

Feeney and King (2001) shows that there is an early mover advantage in a (symmetric information) sequential parimutuel game with two outcomes. Furthermore, early players may choose outcomes with low ex-ante probabilities in rational anticipation of later players opting for the more likely outcomes. Ottaviani and Sørensen (2006) considers a sequential parimutuel betting model and finds that large bettors relying on common (symmetric) information are incentivised to bet early, while small bettors with private (asymmetric) information are incentivised to bet late. They motivate their finding of early betting for large bettors with common information by drawing an explicit analogy between parimutuel betting and Cournot competition in quantities, such that there exists a Stackelberg first-mover advantage. Effectively, by moving early on the basis of common information, a large bettor moves the

odds so as to discourage subsequent bettors from placing similarly large bets, and thereby appropriates to himself a larger share of the total profit accruing to those in possession of the common information.

In a novel combination of strategic interaction and loss aversion, Qiu (2007) postulates a sequential move parimutuel betting model in which participants evaluate outcomes by reference to the prospect theory of Kahneman and Tversky (1979) and demonstrates the existence of a sub-game perfect Nash equilibrium in such a setting.

2.5.5 The current state of theory

In conclusion, the literature surrounding price formation in parimutuel markets in general, and the favourite longshot bias in particular, has matured to the point where there are now dozens of candidate theories of the favourite longshot bias. In order to progress our understanding in this area it is necessary to focus more effort on work that has the potential to falsify existing theories.

2.6 Areas for further work

2.6.1 Extending strategic interaction models to encompass risk sensitive bettors

Virtually all of the recent literature which models bettors within a game theoretic framework makes the assumption that bettors are risk neutral.²² While this assumption may hold for outsiders betting small amounts for recreational reasons, is it still plausible in the context of sophisticated insiders betting large amounts? In the financial literature, it is standard to assume some type of risk aversion on the part of agents that put a substantial part of their wealth at risk, and it seems natural to extend this approach also to the sophisticated insiders populating game theoretic models of parimutuel betting.

2.6.2 Endogenising the distinction between inside and outside bettors

There is a growing consensus in the literature that bettors tend to be of two types – either recreational outsiders or sophisticated insiders. However, this distinction

²²The only exception we are aware of is a paper by J. Qiu, ‘Loss aversion and mental accounting: the favorite longshot bias in parimutuel betting’, *Jena Economic Research Papers*, (2007), in which bettors have risk aversion in accordance with prospect theory within the context of a sequential parimutuel game.

is typically exogenously imposed (Ottaviani and Sørensen (2010) is an important exception that points the way). In many ways, it might be preferable to consider the distinction between insiders and outsiders in a more nuanced fashion. It is plausible to assume heterogeneity in the degree of utility different people derive from the recreational and monetary aspects of betting. Extending theories that endogenise this distinction would be a step towards a more realistic description of bettor behaviour.

2.6.3 Deriving additional testable implications from existing theoretical frameworks

The large number of theoretical contributions that lay claim to explaining the favourite longshot bias makes it all the more important to bring empirical evidence to bear on these theories. However, most of the empirical work to date has focused on aggregate prices determined via parimutuel betting. In order to improve our ability to test various theories against the data, it is necessary to derive more detailed empirically testable implications from existing theories. This could be achieved by elaborating in more detail the various relationships between elements such as the distribution of bet sizes, the variability of implied market prices and the composition of bettors implied by different theories.

2.6.4 Creating a richer set of empirical stylised facts

Being able to generate the favourite longshot bias is no longer a sufficient criterion for deeming a theory plausible. In order to guide the development of theory in the right direction it will be necessary to rely on a richer set of empirical “stylised facts”. Documenting additional empirical regularities would help both in falsifying existing theories and in guiding the development of new theories by providing hints as to the underlying mechanisms that are at work in parimutuel betting contests.

2.7 Next steps

Some areas were identified for further work in the previous section; in the following chapters we will address a number of these issues. In chapter 3 we formulate a theory of parimutuel markets in which we combine strategic interaction with individual risk preferences. The motivation for this endeavour is more fully explained in the introduction to the next chapter, but draws on the arguments set out above. Chapter 4 introduces a new dataset containing data at the individual bet level, which we use

to investigate the relationship between the amount bet and the level of betting “success”. We also use this data to test the theory expounded in chapter 3.

Chapter 3

Strategic interaction with risk preferences

3.1 Introduction

3.1.1 Motivation

Both risk preferences and strategic interaction has been employed to describe price formation in parimutuel markets (and empirical phenomena such as the favourite longshot bias). Generally, risk preferences have been framed within a representative agent framework, which obviously precludes strategic interaction since there can be no strategic interaction with a single agent. On the other hand, almost all strategic interaction models assume risk neutrality on the part of bettors. Our contribution to the literature is to formulate a theoretical model of parimutuel markets that combine both strategic interaction and risk preferences.

Combining risk aversion and strategic interaction in a single model comes at some cost in terms of complexity, therefore it is worth discussing the motivation behind this effort. We are motivated by both theoretical and empirical considerations. From a theoretical point of view, combining risk preferences and strategic interaction in a single model of parimutuel markets is a novel endeavour from which new insights into the interaction of these two effects may emerge.

From an empirical point of view, we note that both risk loving representative agent models and models of strategic interaction with inside information are capable of explaining empirically observed regularities such as the favourite longshot bias. However, these models are based on mutually inconsistent assumptions. To disentangle simultaneously the empirical implications of risk preferences and strategic interaction requires a theoretical framework that is capable of modelling both. For in-

stance, it may well be that both strategic interaction *and* risk preferences play a role in parimutuel price formation. Current theoretical frameworks offer no basis for testing such an hypothesis, since none combine both strategic interaction and risk preferences. However a model such as the one we propose holds out the prospect of testable implications which would allow empirical data to reject this hypothesis, or offer support for it if not rejected.

A strength of our model is that it not only incorporates risk preferences and strategic interaction, but that it parameterises those features. Risk preference is modelled by a real variable α such that negative values of α correspond to progressively higher levels of risk aversion, while positive values of α implies correspondingly higher levels of risk love. Setting α equal to zero is equivalent to assuming risk neutrality. This implies that this model encompass existing models of strategic interaction between risk neutral insiders as a special case.¹ Likewise, the degree of strategic interaction is modelled by the number of insiders N . Therefore the case of a representative bettor with risk preferences is a special case of our model with $N = 1$. Low values of $N > 1$ correspond to situations where game theoretical strategic interaction is an important driver of parimutuel price equilibrium. In the limit as N becomes large, the model reduces to a framework that assumes infinitesimally small inside bettors, such that the impact of each insider's bet on the parimutuel price becomes negligible. In addition to the above, we also model the common belief of insiders regarding the probability of horse 1 winning, q , which differs from that implied by the total amount bet by exogenous recreational outsiders, b_0 . Finally, we incorporate the track takeout rate τ . Solving for the symmetric static Nash equilibrium, we obtain an analytic closed form solution for the optimal insider betting amount. Since the takeout rate τ is directly observable and the number of insiders N and outsider total betting b_0 may be estimated by a classification rule, this holds out the promise of being able to jointly estimate the degree of insider risk preference α and the degree of inside information q directly from the individual bet level data.

The chapter is set out as follows: the remainder of this section introduces our notation and derives, for exposition, a simple strategic interaction model of risk neutral insiders corresponding to that of Hurley and McDonough (1995) with zero takeout rates. In section 3.2 we present our formal model combining strategic interaction with risk preferences, followed by section 3.3 in which we solve for the static Nash equilibrium of insider betting. Section 3.4 considers the results obtained in more detail using a number of simplified models.

¹At the cost of a simpler model in other respects. We assume a parimutuel contest with only two outcomes (ie horses), and we do not model information asymmetries between insiders.

3.1.2 Notation

Our notation closely follows that of Ottaviani and Sørensen (2008), while also borrowing from the earlier literature, notably Sauer (1998). We assume a parimutuel contest with two outcomes $h \in \mathcal{H} = \{1, 2\}$. The set of bettors are indexed by $n \in \mathcal{N} = \{0, 1, 2, \dots, N\}$. We use the convention that $n > 0$ indexes strategic *insiders* with common private belief $q = Pr(h = 1)$ not shared with a single representative outsider bettor denoted by $n = 0$. A representative exogenous *outsider* bets a total of $b_0(1) > 0$ in the first period on horse $h = 1$ and $b_0(2) > 0$ on horse $h = 2$, followed by $N \geq 1$ identical insiders betting $b_n(1)$ on horse $h = 1$ in period two. Without loss of generality, we label the outcomes such that $q \geq b_0$. The total amount bet in the race is $B = \sum_{h \in \mathcal{H}} \sum_{n \in \mathcal{N}} b_n(h)$. We normalise all amounts so that the total amount bet by outsiders equal unity, implying $b_0(1) + b_0(2) = 1$. For convenience, we drop the outcome index when referring to horse $h = 1$, so that $b_0 = b_0(1)$, $b_n = b_n(1)$ and so forth. Insiders are assumed to be financially unconstrained. The track takeout rate is denoted by $\tau \in [0, 1)$. The gross payoff of a winning bet is denoted by π . The final market probability of horse h winning is given by $m(h) = \frac{b(h)}{B}$.

When we explicitly take into account risk preferences, insiders are assumed to have utility over net final period wealth² $w_n = -b_n + \pi_n$ given by $u_n(w_n) = E[w_n] + \alpha(\sqrt{VAR[w_n]})$ with risk preference parameter $\alpha \in \mathbb{R}$.³ Insider expectations and variance are taken with regards to their horse $h = 1$ winning probability belief q .

3.1.3 Risk neutral model

By way of introducing above the notation, we construct a strategic interaction model of a horse race with risk neutral insiders which corresponds to that developed by Hurley and McDonough (1995), but with the track takeout rate $\tau = 0$. We assume insiders are risk neutral maximisers of expected payoffs. Since all insiders are *ex-ante* identical the game is symmetrical, so that in equilibrium $b_1 = b_2 = \dots = b_N$. We denote by b_n the amount individually bet by other insiders for $n \geq 2$. The expected net payoff π_1 of a bet b_1 on horse 1 by bettor 1 is then given by

$$\pi_1(b_1) = -b_1 + q \left(\frac{b_1}{b_0 + b_1 + b_n(N-1)} \right) (1 + b_1 + b_n(N-1)) \quad (3.1.1)$$

The first term represents the amount bet by bettor 1. The last term represents the expected winning payout on horse 1, which is equal to the insiders' belief regarding

²Since the time between placing a bet and receiving a payout is a matter of days at most, we ignore the time value of money in our analysis.

³However we will need to impose restrictions on the range of α in order to guarantee strictly positive betting amounts.

the probability of horse 1 winning ($q = Pr(h = 1)$) times the winning payout. The amount due to insider 1 in case horse 1 wins is equal to the product of insider 1's share of all bets laid on horse 1, being $b_1/(b_0 + b_1 + b_n(N - 1))$ and the total betting pool, which is equal to $1 + b_1 + b_n(N - 1)$.

First order conditions for insider 1 maximising her expected net payoff is given by

$$\frac{\partial \pi_1(b_1)}{\partial b_1} = 0 \quad (3.1.2)$$

Which yields

$$-1 + \frac{q(b_n(N - 1) + b_1 + 1)}{b_n(N - 1) + b_1 + b_0} + \frac{b_1 q}{b_n(N - 1) + b_1 + b_0} - \frac{b_1 q(b_n(N - 1) + b_1 + 1)}{(b_n(N - 1) + b_1 + b_0)^2} = 0 \quad (3.1.3)$$

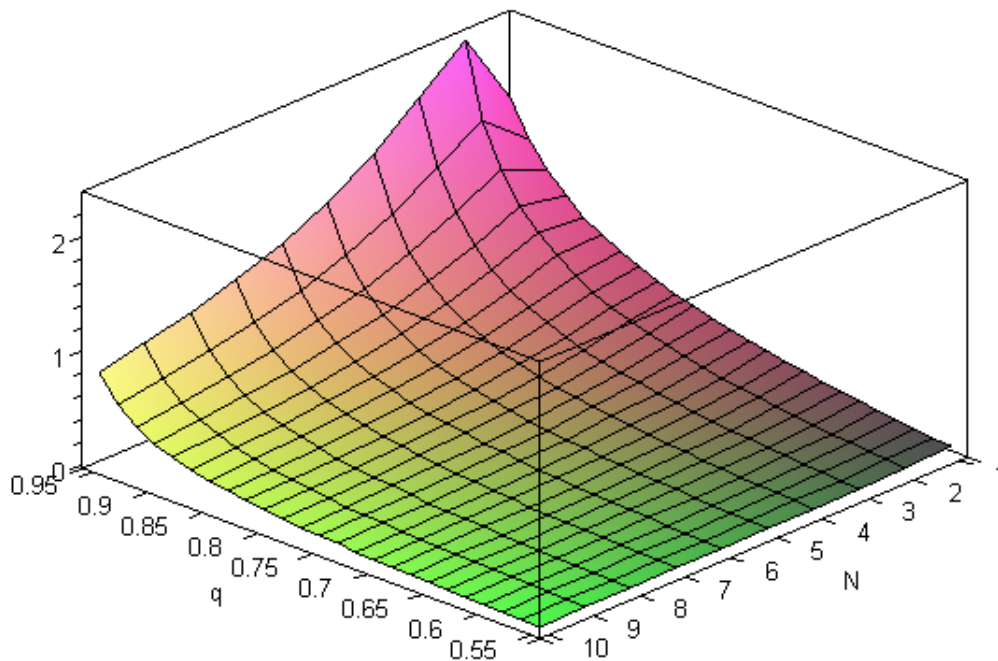
Solving simultaneously for the above and $b_1 = b_n$ (symmetry condition), we obtain the unique positive solution to b_1

$$b_1 = \frac{\sqrt{b_0^2 q^2 N^2 - 2b_0 q^2 N^2 + q^2 N^2 + 2b_0^2 q^2 N - 2q^2 N - 4b_0^2 q N + 4b_0 q N + b_0^2 q^2 - 2b_0 q^2 + q^2}}{-2(q - 1)N^2} + \frac{b_0 q N + q N - 2b_0 N + b_0 q - q}{-2(q - 1)N^2} \quad (3.1.4)$$

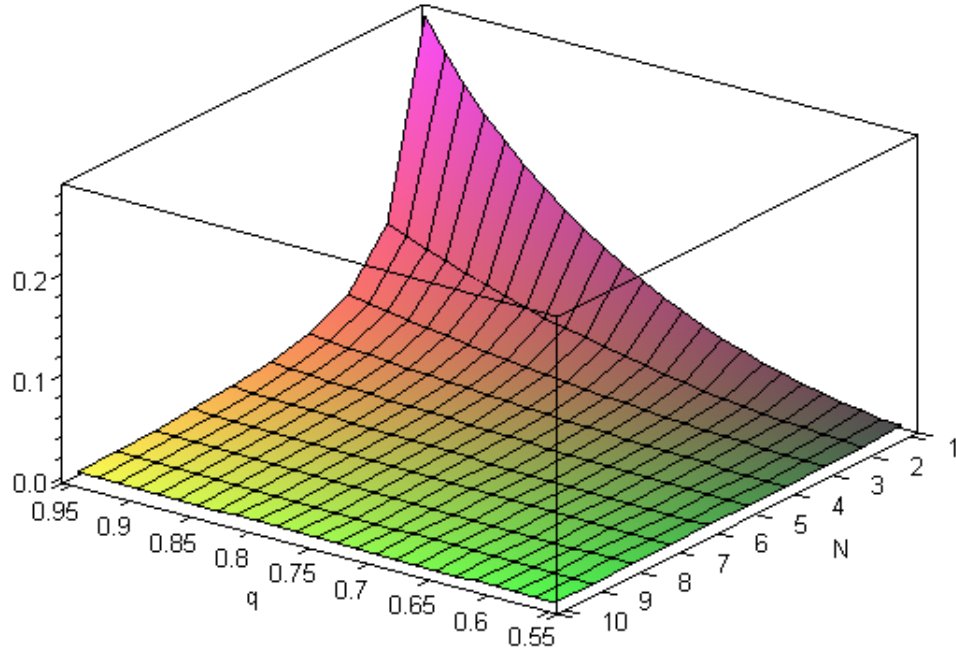
which by symmetry is also the Nash equilibrium betting amount for all the other insiders. Note that the optimal insider betting amount in this model is a function of only three variables:

1. The fraction b_0 bet by the outsider on horse 1,
2. The insiders' belief q regarding the probability of horse 1 winning the race, and
3. The number of insiders N .

To gain further insight into the result, we let $b_0 = 0.5$ and then plot the amount bet, b_1 against q ranging from 0.51 to 0.99 (on the x -axis) and N in the sequence $\{1, 2, \dots, 10\}$ in Figure 3.1.1.

Figure 3.1.1: Amount bet by individual insiders for given q and N 

Using the same variables, we also plot the net payout (ie profit) for an individual insider in Figure 3.1.2. Note that a single monopolistic insider *à la* Isaacs (1953) corresponding to $N = 1$ bets *less* than would be the case in a duopoly corresponding to $N = 2$. This is a consequence of strategic interaction in the case of $N > 1$. Competing insiders bet *more* in anticipation of other insiders' bets diluting the payoff on their own bets. As N becomes larger, however, the optimal individual insider betting amount decreases in order to avoid collective over-betting in the final pool.

Figure 3.1.2: Net payout of individual insiders for given q and N


It can be seen that the profitability of insider betting is increasing in the extent of outsider mispricing (which is proportional to q given that b_0 is fixed at 0.5) and decreasing in the number of insiders N . As the number of insiders N tends towards infinity, the optimal bet for each insider tends towards zero. This does not mean that total insider betting disappear; on the contrary, we can show that as N tends towards infinity, *total* insider betting is bounded by

$$\lim_{N \rightarrow \infty} Nb = \frac{q - b_0}{1 - q} \quad (3.1.5)$$

In other words, the total amount of insider money bet depends only on the distance between the their assessment of the winning probability and the winning probability implied by outsider betting ($q - b_0$) relative to the insiders' assessment regarding the probability of their chosen horse losing ($1 - q$).

The final market probability which corresponds to this limit is

$$m = \frac{b(1)}{B} = \frac{b_0 + \frac{q - b_0}{1 - q}}{1 + \frac{q - b_0}{1 - q}} = q \quad (3.1.6)$$

In other words, as the number of insiders tends towards infinity, the final market probability approaches the true probability. The above results are equivalent to those derived by Hurley and McDonough (1995) (except that we assume a takeout rate of zero).

3.2 Model

The model presented here extends that of Hurley and McDonough (1995) by explicitly incorporating an arbitrary risk preference (denoted by α) on the part of insiders, while retaining the strategic interaction element. The Hurley and McDonough (1995) model is therefore a special case of our model with a risk preference parameter $\alpha = 0$, indicating risk neutral insiders. The model follows the basic setup introduced in sub-section 3.1.2 (“Notation”).

3.2.1 Payoff

Let b_1 be the amount bet by insider $n = 1$ on horse $h = 1$, and let b_n be the amount bet by all the other insiders $n > 1$ on horse 1. As before, outsider bettors bet a fraction b_0 on horse 1 and $1 - b_0$ on horse 2, so that total outsider betting is normalised to unity. We calculate the best response function for insider $n = 1$. Since bettors are ex-ante identical, symmetry dictates that all the other insiders will face the same best response function *mutatis mutandis*.

All variables are real, with the exception of the number of bettors N , which is a strictly positive integer. The variables are subject to the constraints below:

$$b_0 > 0, b_0 < 1, q > b_0, q < 1, \tau \geq 0, \tau < 1, b_1 \geq 0, b_n \geq 0 \text{ and } N \geq 1$$

It follows that the total amount $B1$ bet on horse $h = 1$ should satisfy

$$B1 \equiv b_n(N - 1) + b_1 + b_0 > 0$$

and similarly, the total amount B bet on all horses should satisfy

$$B \equiv b_n(N - 1) + b_1 + 1 > 0$$

The gross winning bet payoff π is given by

$$\pi \equiv \begin{cases} \frac{b_1}{B1} (B) (1 - \tau) & h = 1 \\ 0 & h = 2 \end{cases}$$

or in expanded form

$$\pi = \begin{cases} \frac{b_1}{b_n(N-1)+b_1+b_0} (b_n (N - 1) + b_1 + 1) (1 - \tau) & h = 1 \\ 0 & h = 2 \end{cases}$$

Net final wealth w is defined as the gross winning bet payoff less the initial bet

$$w \equiv \pi - b_1$$

3.2.2 Expectation

Expected final wealth is taken with respect to the insider's belief in the winning probability of horse $h = 1$, $q \equiv Pr(h = 1)$.

$$E[w] = q(\pi - b_1) + (1 - q)(-b_1)$$

Substituting in the payoff π , the expression can be written as

$$E[w] = q \left(\left(\frac{b_1}{b_n(N-1) + b_1 + b_0} (b_n(N-1) + b_1 + 1)(1 - \tau) \right) - b_1 \right) - (1 - q)(b_1)$$

3.2.3 Variance

The variance of final wealth is calculated by reference to

$$VAR[w] = E[w^2] - (E[w])^2$$

Substituting in the expressions for w and $E[w]$ we obtain

$$VAR[w] = q(\pi - b_1)^2 + (1 - q)(-b_1)^2 - (q(\pi - b_1) + (1 - q)(-b_1))^2$$

Replacing payoff π with its definition, this becomes

$$VAR[w] = q \left(\left(\frac{b_1}{b_n(N-1) + b_1 + b_0} (b_n(N-1) + b_1 + 1)(1 - \tau) \right) - b_1 \right)^2 + (1 - q)(-b_1)^2 - \left(q \left(\left(\frac{b_1}{b_n(N-1) + b_1 + b_0} (b_n(N-1) + b_1 + 1)(1 - \tau) \right) - b_1 \right) + (1 - q)(-b_1) \right)^2$$

3.2.4 Utility

We define utility u as a linear combination of expected final wealth and the standard deviation of final wealth, with the parameter α indicating risk attitude – risk neutral if $\alpha = 0$, risk averse if $\alpha < 0$ and risk loving if $\alpha > 0$.

$$u = u_1 \equiv E[w] + \alpha \sqrt{VAR[w]}$$

Substituting in $E[w]$ and $VAR[w]$ we obtain

$$u = q(\pi - b_1) + (1 - q)(-b_1) \\ + \alpha \sqrt{q(\pi - b_1)^2 + (1 - q)(-b_1)^2 - (q(\pi - b_1) + (1 - q)(-b_1))^2}$$

Substituting in π and simplifying yields

$$u = \frac{b_1(q - q\tau - qb_1\tau - qb_nN\tau + qb_n\tau - b_0 - b_1 - b_nN + b_n + b_1q + qb_nN - qb_n)}{b_0 + b_1 + b_nN - b_n} \\ + \frac{b_1 \left(\alpha (1 + b_1 + b_nN - b_n) (1 - \tau) \sqrt{q(1 - q)} \right)}{b_0 + b_1 + b_nN - b_n}$$

3.3 Equilibrium

3.3.1 Best response function (first order conditions)

To find insider 1's best response (choice of betting amount b_1) to other insiders' best responses (choice of betting amount b_n), we formulate first order conditions for utility maximisation.

$$\frac{\partial}{\partial b_1} u = u'(b_1) = 0$$

where

$$u'(b_1) = qG - 1 + q + \frac{\alpha(2q(-b_1+\pi)G+2(1-q)b_1-2(q(-b_1+\pi)-(1-q)b_1)(qG-1+q))}{2\sqrt{q(-b_1+\pi)^2+(1-q)b_1^2-(q(-b_1+\pi)-(1-q)b_1)^2}}$$

and G equals the first derivative of final period wealth $w = \pi - b_1$ with regards to b_1

$$G = -1 + \frac{\pi}{b_1} - \frac{\pi}{B1} + \frac{\pi}{B}$$

(Recall that π denotes the gross winning bet payoff such that $\pi \equiv \frac{b_1}{B1}(B)(1 - \tau)$ if horse $h = 1$ wins, while B denotes the total amount bet $B \equiv b_n(N - 1) + b_1 + 1 > 0$ and $B1$ denotes the total amount bet on horse $h = 1$ such that $B1 \equiv b_n(N - 1) + b_1 + b_0 > 0$).

Hence a choice of $b_1^* = f(b_n, N, q, \alpha, \tau)$ such that $u'_1(b_1^*) = 0$ is a best response for insider 1 to the amount b_n individually bet by the other insiders for $n \in [2..N]$.⁴

3.3.2 Nash equilibrium

We are seeking insider betting amounts in a Nash equilibrium. Since all insiders are ex-ante identical, the amount they bet will be identical in equilibrium, that is $b_1 = b_n$, where b_1 denotes the amount bet by the first insider and b_n denotes the (equal) amount bet by each of the remaining insiders indexed by $n \in \{2..N\}$.

The symmetric Nash equilibrium is found by simultaneously solving for $u'(b_1) = 0$ and $b_1 = b_n$. (Since the insiders are ex-ante identical, the above automatically implies that $u'_n(b_n) = 0$ also for all $n \in \{2..N\}$. Hence each insider play their best response to every other insider's best response, as is required for a Nash equilibrium to hold.)

We obtain two candidate solutions (denoted s_1 and s_2) from the above:

$$s_1 = A(C + B)$$

$$s_2 = A(C - B)$$

where

$$A = \frac{1}{2N^2(\sqrt{1-q}(1-\tau)q^{3/2} - \alpha(1-\tau)q^2 + \alpha(1-\tau)q - \sqrt{q}\sqrt{1-q})}$$

$$B = \left((1 - b_0)(1 - \tau)q^{\frac{3}{2}}(1 - q)D \right)^{\frac{1}{2}}$$

$$C = (1 - \tau)q(\alpha(q - 1) - \sqrt{q}\sqrt{1 - q})(N + b_0 - 1 + b_0N) + 2\sqrt{q}\sqrt{1 - q}b_0N$$

$$D = (\alpha^2 - 1)Kq^{\frac{3}{2}} - 2\sqrt{1 - q}\alpha Kq + (4b_0N - K\alpha^2)\sqrt{q} + 4\alpha\sqrt{1 - q}b_0N$$

and

$$K = (2N - 1 + 2bN + b_0 + N^2b_0 - N^2)(1 - \tau)$$

However, to be an admissible solution, a candidate solution should be consistent with a) the initial inequality assumptions and b) second order conditions for utility maximisation. We show (below) that s_2 is the unique positive solution and that it satisfies second order conditions, so that the optimal insider betting amount is given by

$$b^* = b_1 = b_n = s_2 = A(C - B)$$

⁴Subject to second order conditions and the inequalities assumed initially.

provided that $\alpha_L < \alpha < \alpha_H$ as shown in sub-section 3.3.3 below.

Note that optimal insider betting b^* can be expressed more compactly as an implicit equation in terms of the risk preference parameter α

$$\alpha = \frac{N^2(1-q(1-\tau))b^{*2} + (2Nb_0 - (1-\tau)(b_0 + Nb_0 - 1 + N)q)b^* + b_0^2 - qb_0 + q\tau b_0}{(1-\tau)\sqrt{(1-q)(q)}(b^{*2}N^2 + (b_0 + Nb_0 - 1 + N)b^* + b_0)}$$

The final market probability of horse 1 winning (which we can also interpret as the final market price of horse 1) is given by

$$m = \frac{b_0 + Nb^*}{1 + Nb^*} = \frac{b_0 + NA(C - B)}{1 + NA(C - B)}$$

where A , B , and C retain their earlier definitions.

3.3.3 Admissibility and second order conditions

By initial assumption it is required that any solution to b_1 must be positive. (Bettors cannot bet a negative amount or “go short”). We now show that $s_2 > 0$ and $s_1 < 0$ subject to restrictions on the range of α .

3.3.3.1 Positive betting amount

First we solve for critical values of α such that $b^* = s_2 = 0$, obtaining

$$\alpha_L = \frac{b_0 - q(1-\tau)}{\sqrt{q}\sqrt{1-q}(1-\tau)}$$

and

$$\alpha_H = \frac{1 - q(1-\tau)}{\sqrt{q}\sqrt{1-q}(1-\tau)}$$

Since the denominators are always positive, and because $q(1-\tau) < 1$ (each of the factors are strictly less than 1), we have $\alpha_H > 0$ and since $b_0 < 1$, we have $\alpha_H > \alpha_L$. Closer examination shows that s_2 is positive in the interval (α_L, α_H) and negative outside it. Since s_2 is continuous in α , that means $s_2 > 0$ for all $\alpha_L < \alpha < \alpha_H$.

The critical values of α for s_1 are identical to that of s_2 , but s_1 is negative in the interval (α_L, α_H) and positive outside it. Analogous to the proof above, that means $s_1 < 0$ for all $\alpha_L < \alpha < \alpha_H$, since s_1 is also continuous in α .

3.3.3.2 Second order conditions

In order to ascertain that $b^* = b_1 = s_2$ is a maximum, we have to verify second order conditions.

$$\frac{\partial^2}{\partial b_1^2} u = u''(b_1) < 0$$

The second derivative of utility u with respect to b_1 is given by

$$u''(b_1) = \frac{2(1-\tau)(1-b_0)(q^{3/2}\alpha - \sqrt{q}\alpha - q\sqrt{1-q})(b_0 - b_n + b_n N)}{\sqrt{1-q}(b_0 + b_1 + b_n N - b_n)^3}$$

We are required to show that $u''(b^*) < 0$ as required for second order conditions, subject to the constraints on α to ensure a positive betting amount ($\alpha_L < \alpha < \alpha_H$).

Recall that we initially assumed

$b_0 > 0$, $b_0 < 1$, $q > b_0$, $q < 1$, $\tau \geq 0$, $\tau < 1$, $b_1 \geq 0$, $b_n \geq 0$ and $N \geq 1$

Additionally, we have shown that optimal b^* is positive for α in the range $\alpha_L < \alpha < \alpha_H$. Substituting b^* for b_1 and b_n , we obtain

$$u''(b^*) = \frac{2(1-\tau)(1-b_0)(q^{3/2}\alpha - \sqrt{q}\alpha - q\sqrt{1-q})(b_0 + b^*(N-1))}{\sqrt{1-q}(b_0 + b^*N)^3}$$

It follows directly from the initial assumptions above that all factors of $u''(b_1)$ are strictly positive save for

$$J = \left(q^{3/2}\alpha - \sqrt{q}\alpha - q\sqrt{1-q} \right)$$

Hence we need to prove $J < 0$ in order to establish $u''(b^*) < 0$.

Note that

$$\frac{\partial}{\partial \alpha} J = q^{3/2} + q^{1/2} > 0 \text{ since } 0 < b_0 < q$$

Therefore J is strictly increasing in α for all q . Now let J_H be J with α_H substituted for α , then

$$J_H = q^{3/2}\alpha_H - \sqrt{q}\alpha_H - q\sqrt{1-q} > \left(q^{3/2}\alpha - \sqrt{q}\alpha - q\sqrt{1-q} \right) = J$$

which follows from $\alpha_H > \alpha$ and the fact that J is strictly increasing in α .

Substituting $\alpha_H = \frac{1-q(1-\tau)}{\sqrt{q}\sqrt{1-q(1-\tau)}}$ into J_H and simplifying, we obtain

$$J_H = - \left(\frac{2q(1-\tau) + 1 + q}{\sqrt{1-q}(1-\tau)} \right)$$

Since all the factors in $\frac{2q(1-\tau)+1+q}{\sqrt{1-q}(1-\tau)}$ are strictly positive from our initial assumptions, it follows that $J_H < 0$ and since $J < J_H$ we have $J < 0$. As we have already established that $J < 0$ implies $u''(b^*) < 0$ as required, our proof is complete.

3.4 Discussion

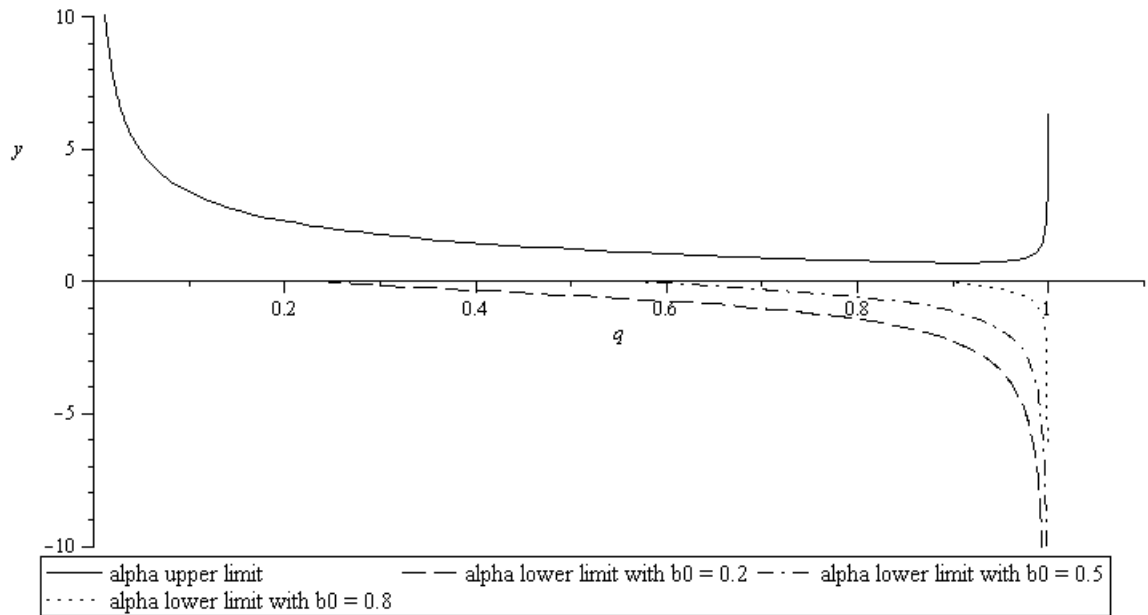
3.4.1 Symmetry

This model is inherently symmetrical as there are only two mutually exclusive outcomes. To illustrate this symmetry, consider an equilibrium solution for insider betting amounts $b^* = f(N, q, \tau, \alpha, b_0(1))$ where $b_0(1) < q$ is assumed. Now, how do we cope with a situation where $b_0(1) > q$ instead? We can deal with this by re-writing the equation so that $b^* = f(N, 1 - q, \tau, \alpha, 1 - b_0(1))$ is re-interpreted to mean the amount an insider should bet on horse 2, rather than horse 1. We can do this since $1 - b_0(1) < 1 - q$ as required if $b_0(1) > q$. But this is really just equivalent to switching the labels on the horses around, so that horse 1 becomes horse 2 and *vice versa*. If we do that, $b^* = f(N, 1 - q, \tau, \alpha, 1 - b_0(1))$ becomes $b^* = f(N, q, \tau, \alpha, b_0(1))$ as before with $b_0(1) < q$ as originally required. By way of example, let $N = 5, q = 0.3, \alpha = 0.1$ and $\tau = 0.15$ with $b_0(1) = 0.7$ and $q = 0.3$. Since $b_0(1) = 0.7 \not< q = 0.3$, we cannot directly use the equilibrium solution for the optimal amount b^* for insiders to bet on horse 1, which requires $b_0(1) < q$. However, we can use the formula to calculate the amount that would be bet by insiders on horse 2. Note that $b_0(2) = 1 - b_0(1) = 1 - 0.7 = 0.3$ and insiders' belief that horse 2 will win is given by $1 - q = 1 - 0.3 = 0.7$. Now we have $b_0(2) = 0.3 < (1 - q) = 0.7$ as is required for the equilibrium insider betting amount, but with the understanding that this insider betting amount now relates to horse 2 rather than horse 1. We then calculate the equilibrium amount insiders will bet individually on horse 2 as $b^* = f(N, (1 - q), \tau, \alpha, b_0(2)) = f(5, 0.7, 0.15, 0.1, 0.3) \approx 0.1480$. (Recall that all amounts are expressed as a fraction of total outsider betting so that $b_0(1) + b_0(2) = 1$. So that means each of the 5 insiders will bet 14.80% of total outsider betting, or 73.98% of total outsider betting for all 5 insiders together.)

3.4.2 No equilibrium zones

Another feature of our model is that it places restrictions on the range of the risk preference parameter α in order to enforce positive betting amounts by insiders. We showed that the amount bet by insiders will be positive only if $\alpha_L = \frac{b_0 - q(1-\tau)}{\sqrt{q}\sqrt{1-q}(1-\tau)} < \alpha < \frac{1-q(1-\tau)}{\sqrt{q}\sqrt{1-q}(1-\tau)} = \alpha_H$. The two critical values are not necessarily symmetric. This is illustrated in the Figure 3.4.1 below where we plot the critical bounds of α as a function of q assuming that the takeout rate is zero ($\tau = 0$).

Figure 3.4.1: Critical values of α as a function of q , assuming different values of b_0



Note that in plotting the lower bounds, the plots are only shown for $q(1 - \tau) > b_0$, which is one of the initial model assumptions $q > b_0$, adjusted for the takeout rate. As the insider’s belief in the winning probability of horse 1 approaches certainty, the higher and lower bound on alpha tends towards positive and negative infinity respectively. In other words, the “better” the quality of the insiders’ inside information, the greater the range of risk preferences across which they are willing to participate. On the other hand, if q is close to b_0 indicating that insiders’ beliefs do not diverge much from outsiders, the lower bound of α approaches zero. This suggest that when there is no inside information, only risk loving insiders have an incentive to participate.

We can obtain further intuition regarding this result by solving $\alpha_L = \alpha = \frac{b_0 - q(1-\tau)}{\sqrt{q}\sqrt{1-q}(1-\tau)}$ for q . This allows us to frame the lower bound constraint on α as a constraint on insider’s beliefs q instead. The critical value of q thus obtained is

$$q_{cr} = \frac{\alpha^2(1 - \tau) + \frac{|\alpha|}{\alpha} \sqrt{\alpha^2(\alpha^2\tau^2 - 2\alpha^2\tau - 4b_0\tau + \alpha^2 - 4b_0^2 + 4b_0)} - 2b_0}{2(\alpha^2 + 1)(1 - \tau)}$$

Hence, enforcing a positive betting amount requires $q > q_{cr}$. (Note that the formula is sensitive to the sign of α , with the square root term being positive for positive α and negative for negative α due to the presence of the sign function $\frac{|\alpha|}{\alpha}$.) In the special case of $\alpha = 0$ (risk neutrality), this constraint simplifies to

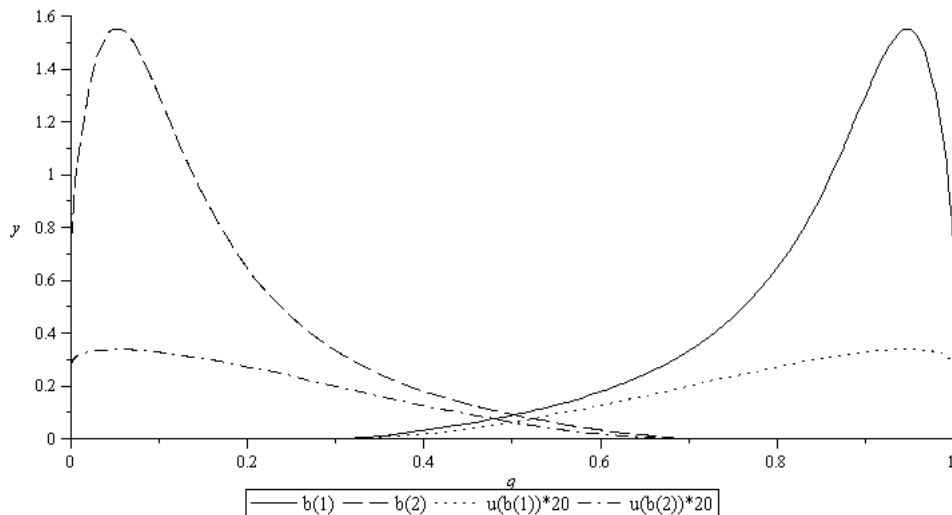
$$q_{cr} = \frac{b_0}{1 - \tau}$$

which is the same as the constraint imposed in Ottaviani and Sørensen (2008) discussing a simplified version of the model set forth in Isaacs (1953).

3.4.3 Insider beliefs

Combining the ideas explored in sub-section 3.4.1 (“Symmetry”) and sub-section 3.4.2 (“No equilibrium zones”), we now consider the individual insider betting amount on both horses as the winning belief regarding horse 1 ranges from 0 to 1. First, we consider the case of risk-loving insiders where $N = 5$, $\tau = 0.1$, $b_0 = 0.5$ and $\alpha = 0.5$.

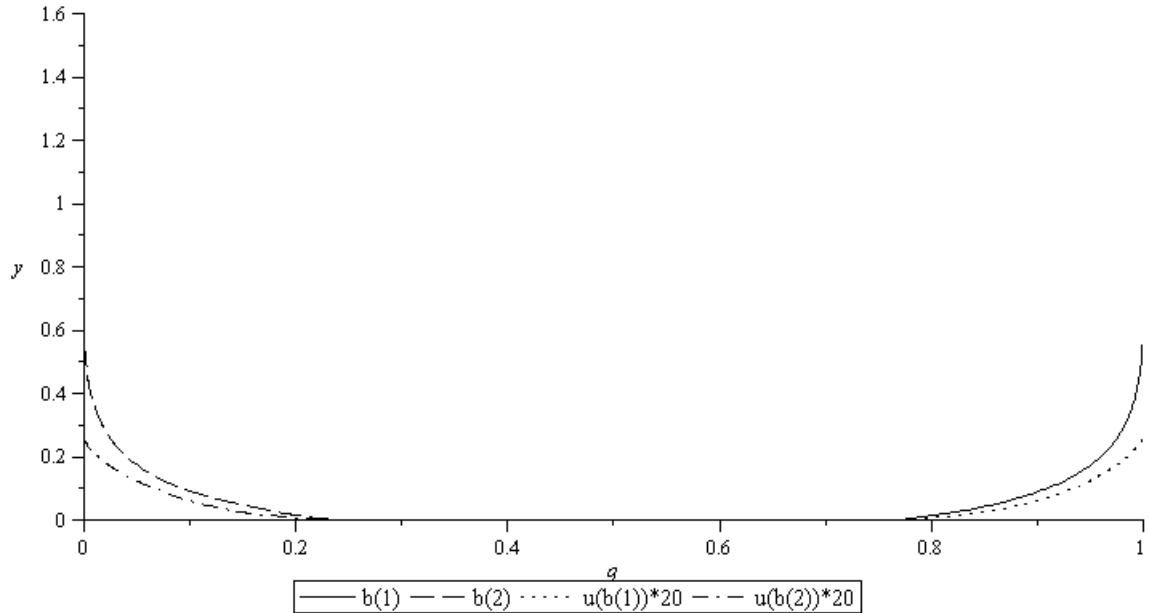
Figure 3.4.2: Risk loving insider betting amounts on either horse as a function of q , with associated utility (utility scaled 20x)



Note that the cross-over point, where insiders change the horse they should optimally bet on, corresponds to $b_0 = 0.5$, the winning probability that outside bettors associate with horse 1. The switch happens because at $q = 0.5$ the utility of betting on horse 1 becomes higher than betting on horse 2 as q increases. Because insiders

are risk lovers in this example, they never abstain from betting. Now consider the same example, but with the risk aversion parameter α set to -0.5 , so that insiders are risk averse.

Figure 3.4.3: Risk averse insider betting amounts on either horse as a function of q , with associated utility (utility scaled 20x)



This paints a very different picture – there is a large range of beliefs where insiders are not prepared to bet. Due to their risk aversion, insiders require a large amount of risk adjusted return, in the form of a large difference between the beliefs of outsiders represented by b_0 and their own beliefs, denoted by q , before they are willing to bet. The examples above suggest that in considering participation by insiders, we can view inside information and risk preference as substitutes, in the sense that better inside information can compensate for higher risk aversion and *vice versa*. Similarly, with lower risk aversion insiders are willing to participate on the basis of weaker inside information.

3.4.4 Equilibrium final market probabilities

In equilibrium, combined insider and outsider betting gives rise to a total betting pool in which $B1 = b_0 + Nb^*$ is bet on horse 1 and $B2 = 1 - b_0$ is bet on horse 2. Total betting is then simply $B = B1 + B2 = 1 + Nb^*$. As noted earlier, the implied market probability of a horse winning is simply the ratio of the total amount bet on that horse to the total amount bet on all horses. Thus the final market probability of horse 1 is given by

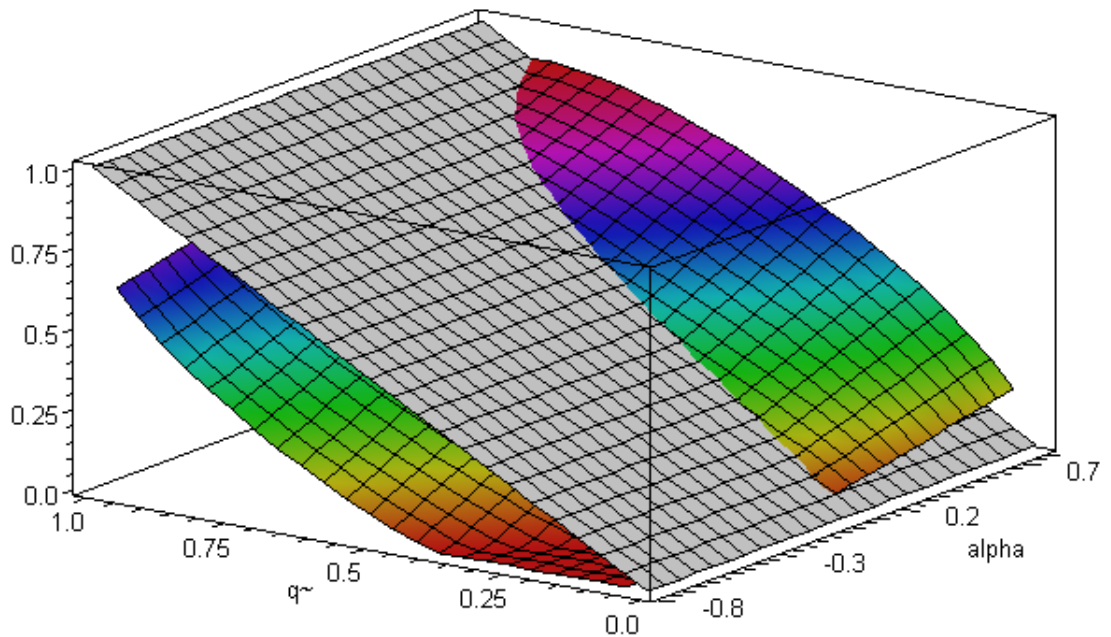
$$m = \frac{b_0 + Nb^*}{1 + Nb^*}$$

We could substitute in our solution for b^* in the general case to frame the total market probability in terms of the original model parameters, but the result is unwieldy. Instead, we let $\tau = 0.2$ which is close to the typical takeout typically observed and let $N \rightarrow \infty$. The expression we obtain for m then becomes a function of only two model parameters – α and q .

$$m = \frac{0.8 \left(\sqrt{q}\sqrt{1-q} - \alpha q + \alpha + \sqrt{1-q}\sqrt{-\alpha^2 q + q + 2\alpha\sqrt{q}\sqrt{1-q} + \alpha^2} \right) q}{\alpha q^2 - \alpha q - q^{3/2}\sqrt{1-q} + q\sqrt{1-q}\sqrt{-\alpha^2 q + q + 2\alpha\sqrt{q}\sqrt{1-q} + \alpha^2} + 2\sqrt{q}\sqrt{1-q}}$$

Graphically, this is represented in Figure 3.4.4.

Figure 3.4.4: Final market probabilities where $N \rightarrow \infty$ and $\tau = 0.2$



Interestingly, the amount b_0 bet by outsiders does not enter into this formula – this suggests that inside information (the difference between q and b_0) is not the cause of the divergence between insider beliefs and final market probabilities when there are a large number of insiders. It would also seem that, with a large number of insiders, some degree of risk love on the part of insiders are required in order to induce over-betting on horse 1.

If instead we consider the special case where $N \rightarrow \infty$ and $q = \frac{1}{2}$ (a fair game), then m reduces to

$$\lim_{N \rightarrow \infty} m = \begin{cases} \frac{1}{2}(1 - \tau)(1 + \alpha) & \alpha > -1 \\ 0 & \textit{otherwise} \end{cases}$$

Ignoring the degenerate case where $\alpha \leq -1$, we note that the final market probability of horse 1 depends crucially on the take out rate τ and the risk aversion parameter α . That is to say that, if the takeout rate is zero and insiders are risk averse, then a large number of insider bettors will drive the final market probability to their ex-ante belief in the winning probability of horse 1 of $q = \frac{1}{2}$, ie there would be no favourite longshot bias. Inverting the above, we can say that in a race where insiders believe either horse is equally likely to win, horse 1 will be over-bet if $(1 - \tau)(1 + \alpha) > 1$. Solving for α , this is equivalent to $\alpha > \frac{1}{1 - \tau} - 1$. Given that the typical takeout rate is not much above 20%, over-betting on horse 1 would require a risk preference of $\alpha > \frac{1}{1 - 0.2} - 1 = 0.25$; in other words, substantial risk love on the part of insiders.

3.4.5 Special case 1: Risk neutral insiders

We can recover the equilibrium betting amounts for risk neutral insiders as a special case of the model with $\alpha = 0$. The equilibrium betting amount is then given by

$$b^* = \frac{(1 - \tau)q^{3/2}(-b_0 + 1 - N(b_0 + 1)) + 2b_0N\sqrt{q}}{2N^2((1 - \tau)q^{3/2} - \sqrt{q})}$$

$$\frac{\sqrt{(1 - b_0)(1 - \tau)q^{3/2}(1 - q)((N - 1)^2 - b_0(N + 1)^2)(1 - \tau)q^{3/2} + 4b_0N\sqrt{q}}}{2N^2\sqrt{1 - q}((1 - \tau)q^{3/2} - \sqrt{q})}$$

And the final market probability of the winning horse is then given, as before, by

$$m = \frac{b_0 + Nb^*}{1 + Nb^*}$$

It can be shown that, in the limit as $N \rightarrow \infty$, $m = (1 - \tau)q$. (This result corresponds to Proposition 1 of Hurley and McDonough (1995)). In other words, with a large number of risk neutral insiders the only source of mispricing (relative to insider's beliefs) is due to the track takeout rate. With a zero track takeout rate, we obtain $m = q$, so that final market probabilities are equal to the common private beliefs of insiders.

Interestingly, risk neutral insiders collectively bet a finite amount Nb^* as $N \rightarrow \infty$.

$$\lim_{N \rightarrow \infty} Nb^* = \frac{q(1-\tau) - b_0}{1 - q(1-\tau)}$$

This is despite the fact that each individual insider bet approaches zero as the number of insiders tends towards infinity.

3.4.6 Special case 2: Risk sensitive representative insider

Similarly, we can reduce the model to a single risk sensitive insider parameterised by α by letting $N = 1$ in our model. We find that a single risk sensitive insider will optimally bet

$$b^* = A(C - B)$$

where

$$A = \frac{1}{2(\sqrt{1-q}(1-\tau)q^{3/2} - \alpha(1-\tau)q^2 + \alpha(1-\tau)q - \sqrt{q}\sqrt{1-q})}$$

$$B = \left((1 - b_0)(1 - \tau)q^{\frac{3}{2}}(1 - q)D \right)^{\frac{1}{2}}$$

$$C = (1 - \tau)q(\alpha(q - 1) - \sqrt{q}\sqrt{1 - q})(2b_0) + 2\sqrt{q}\sqrt{1 - q}b_0$$

$$D = (\alpha^2 - 1)Kq^{\frac{3}{2}} - 2\sqrt{1 - q}\alpha Kq + (4b_0 - K\alpha^2)\sqrt{q} + 4\alpha\sqrt{1 - q}b_0$$

and

$$K = (4b)(1 - \tau)$$

With $\alpha = 0$ this further simplifies to

$$b^* = \frac{q^{3/2}b_0 - q^{3/2}b_0\tau - b_0\sqrt{q} + q\sqrt{1 - b_0}\sqrt{1 - \tau}\sqrt{b_0}\sqrt{1 - q} + q\tau}{\sqrt{q}(1 - q + q\tau)}$$

3.4.7 Special case 3: A simplified model

In order to illustrate the model we now consider a special case in which the number of free parameters is reduced to two. We do this by fixing

- the number of insiders $N = 4$. This allows us to preserve the strategic interaction element of the model
- the takeout rate $\tau = 0$. This allows us to isolate the impact of risk preferences and strategic interaction from the impact of taxes and costs.

- the insider’s common private belief regarding the winning probability of horse 1 is set to $q = \frac{1}{2}$ (a *fair game*).

Our free parameters are then the risk preference α and the amount of outsider betting b_0 . Since q is fixed, we can use b_0 to mimic the degree of insider information, since the difference $q - b_0$ represents the quantum by which the insider’s beliefs regarding the winning probability of horse 1 differs from that of the outsiders.

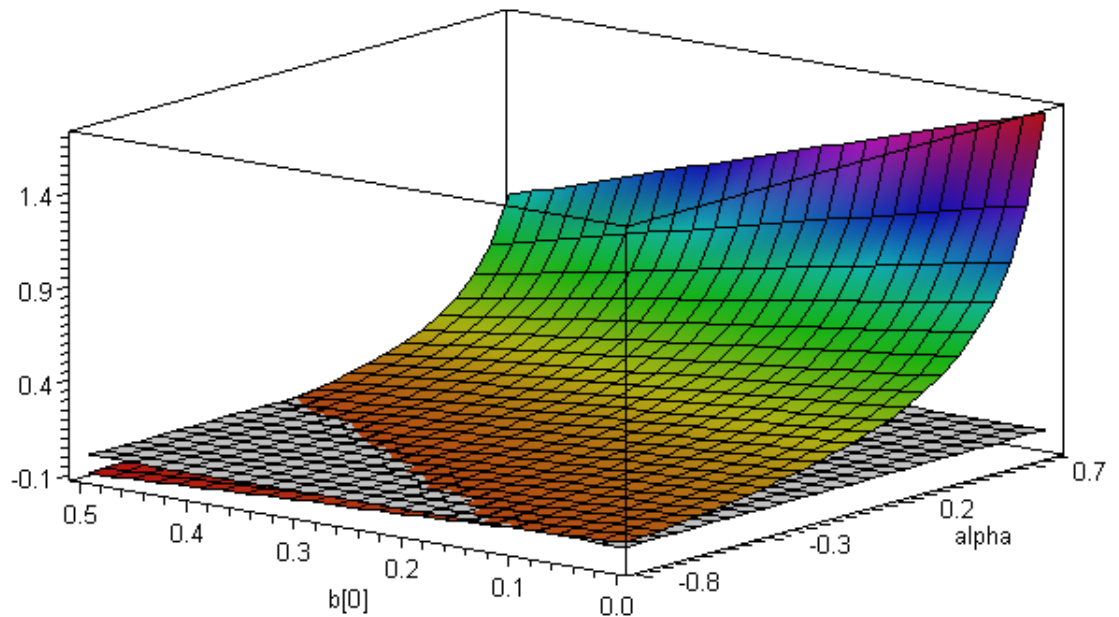
Despite adding a number of simplifying assumptions, this model retains strategic interaction (since $N = 4 > 1$), parameterised risk preferences (α) and inside information (represented by $q - b_0$).

In our simplified model the amount bet by insiders reduces to

$$b^* = \frac{b_0(11 - 5\alpha) - 3\alpha - 3 - \sqrt{(1 - b_0)(\alpha + 1)(9\alpha - 25\alpha b_0 + 7b_0 + 9)}}{32(\alpha - 1)}$$

which we plot in Figure 3.4.5 below.

Figure 3.4.5: Insider betting with $N = 4$, $\tau = 0$ and $q = \frac{1}{2}$



The horizontal grey plane in Figure 3.4.5 represents $b^* = 0$, hence the surface above this plane represents combinations of α and b_0 which result in positive insider betting. Notably, there is a small area in the left hand corner where insider betting is never positive – this area corresponds to situations where insiders are both risk averse ($\alpha < 0$) and have little inside information, that is, b_0 is close to $q = \frac{1}{2}$. This

non-participation of insiders is not driven by the takeout rate, since we assumed the takeout rate to be zero.

3.4.8 Empirical statics

The complexity of the general equilibrium solution in our model means that analysing the model in terms of analytically derived statics is not feasible⁵. Instead we take a more pragmatic approach, in which we use real data to calculate the final market probabilities of each race according to our theoretical model. We then estimate a reduced form representation of the final market probability by regressing the model final market probability directly on the model parameters using OLS. The coefficients thus obtained can be interpreted as empirical aggregate statics for each model parameter. The data that this analysis is based on is discussed in more detail in sub-section 4.2 and sub-section 4.5. A more detailed discussion of our method and the results we obtain is set out in the appendix in section 6.3.

⁵Although statics can be calculated, their analytical complexity precludes meaningful interpretation.

Chapter 4

Empirical tests

4.1 Introduction

In this section we analyse a dataset of 1.6 million horse race bets through the lens of our theoretical model. In particular, we are interested in what the data has to say about insider risk preferences in the context of strategic interaction. Virtually all strategic interaction models of parimutuel betting assume risk neutrality – we would like to test whether this assumption is borne out in the data and, if so, which factors drive this result.

4.2 Data description and analysis

The dataset we use is sourced from the Totalisator Agency Board (TAB) of New Zealand. Under the Racing Act of 2003 the TAB has a statutory monopoly on all sports betting in New Zealand, including horse racing. The dataset records the detail of 5,701,478 individual sport bets from 1 August 2006 to 31 July 2009, including 1,634,808 straight win horse racing bets. All horse race bets administered by the TAB pays out winnings according to the parimutuel method. The data is summarised in Table 4.1 below.

Table 4.1: Data summary – horse racing bets from 1 Aug 2006 to 31 July 2009

Variable	N	Mean	Std. Dev.	Min	Max	Sum
Transaction Value	1,634,808	42.205	202.277	5	40,000	68,997,071
Odds	1,634,808	10.002	11.941	1.15	501	n/a
Bet Payout	255,511	249.852	970.026	5.5	75,000	63,839,934

A quick calculation shows that 15.63% of bets resulted in a winning payout. However, total payouts comprised 92.53% of total bets.

4.3 Insiders vs outsiders

4.3.1 Cut-off point

We operationalise the distinction between insiders and outsiders based on the dollar amount bet. Insiders are defined as those bettors that bet an amount equal to or above a specified cut-off point (denoted by c) in a single bet. Initially, we use $c = 100$ or one hundred New Zealand dollars as our cut-off point. We recognise that our choice of cut-off point is necessarily arbitrary – we have tried to choose an amount that corresponds to the idea of “real money” to the average bettor¹. The data suggest that our classification rule may be capturing some of the essence of the distinction between insiders and outsiders, as insiders are significantly more successful in their betting – this can be seen from Table 4.2.

Although insiders, as we define them, account for slightly less than 10% of betting volume by number of bets, they generate more than two-thirds of the betting by dollar volume. As we can surmise from Table 4.2, they also tend to have a higher winning ratio than outsiders, whether measured by number of bets or dollars. In terms of dollar returns, insiders have a winning percentage that is 9.2% higher than that of outsiders.

4.3.2 Payout ratio and average bet size

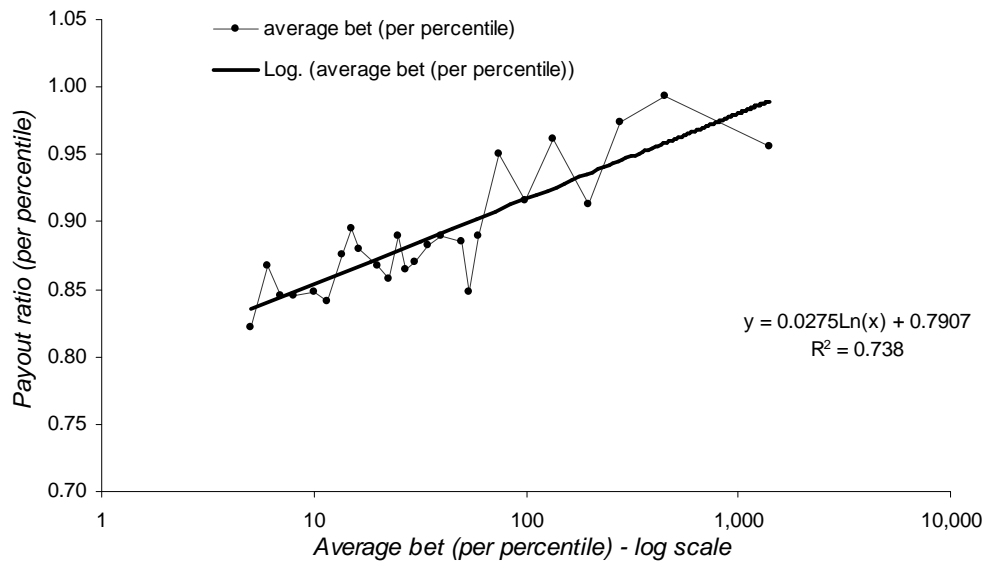
Such a finding naturally suggests that higher betting amounts may proxy for bettors with better inside information, which in turn should lead to higher average payouts. To investigate this idea, we partitioned our dataset into percentiles based on individual betting amount. For each percentile, we calculated both the average amount bet and the payout ratio (= total payouts/total amount bet). The results are shown in Figure 4.3.1 below and appear to support our intuition that higher betting proxies for better inside information.

¹However, we conduct a robustness check in sub-section 4.6.3.2 where we consider three alternative cut-off points, namely NZD 50, NZD 200 and NZD 500.

Table 4.2: Data summary – insiders vs outsiders

Variable	N	Mean	Std. Dev.	Min	Max	Sum
Outsiders						
Transaction Value	1,474,663	16.11	15.05	5.00	99.00	23,756,216
Odds	1,474,663	10.36	12.35	1.15	501.00	n/a
Bet Payout	220,845	93.85	109.68	5.50	4,080.00	20,727,299
Winning % (bets)	15%					
Winning % (dollars)	87%					
Insiders						
Transaction Value	160,145	282.50	592.95	100.00	40,000.00	45,240,202
Odds	160,145	6.68	6.32	1.15	201.00	n/a
Bet Payout	34,666	1,243.66	2,390.87	120.00	75,000.00	43,112,579
Winning % (bets)	22%					
Winning % (dollars)	95%					

Figure 4.3.1: Payout ratio (per percentile) against average bet (per percentile)



Interestingly, the payout ratio appears to increase linearly with the *log* of the average bet size. In other words, the payout rate is increasing in the average bet amount, but at a decreasing rate, suggesting the omnipresent law of diminishing returns may also be at work at the racetrack. An OLS regression in the form

$$Payout\ Rate = \beta_0 + \beta_1 \ln(Avg\ Bet) + \varepsilon$$

is fitted to the data points. We find that this simple specification explains 73.8% of the variation in average payout rates. (A table containing the data that Figure 4.3.1 is based on is located in section 6.2 in the appendix)

The above analysis provides support for our notion that insiders self-select by wagering larger amounts than outsiders, and in our subsequent analysis we use the cut-off point c to classify bettors as insiders and outsiders respectively.

4.4 Race level data

In order to test the final market price predicted by our theory against the data, we need to collapse the individual level data to race level data – this is because the final market price is a race level metric. In aggregating the data, we keep track separately of insider and outsider betting volumes (using our classification rule that insiders are those betting NZD 100 or more). This results in a dataset with 9,543 individual races. The dataset is summarised in Table 4.3. Variable names suffixed with “amt” refer to the sum of individual dollar bets – all other variables are counts of individual bets.

Table 4.3: Race level data summary

Variable	Description	N	Mean	Std. Dev.	Min	Max	Sum
cbets	Count of bets	9,543	171.31	862.59	1	49,740	1,634,807
sbets	Sum of bet amounts	9,543	7,230.05	40,674.88	5	2,448,529	68,996,329
payout	Count of payouts	9,543	6,689.71	39,748.15	0	2,434,716	63,839,874
insider	Count of insiders	9,543	16.78	91.95	0	5,469	160,144
winner	Count of winners	9,543	26.77	129.50	0	6,979	255,511
inwin	Count of insider winners	9,543	3.63	18.38	0	1,059	34,666
outwin	Count of outsider winners	9,543	23.14	111.43	0	5,920	220,845
inlose	Count of insider losers	9,543	13.15	74.27	0	4,410	125,478
outlose	Count of outsider losers	9,543	131.39	664.06	1	38,351	1,253,818
inwinamt	Sum of insider winner bet amounts	9,543	1,188.05	6,583.18	0	385,468	11,337,552
outwinamt	Sum of outsider winner bet amounts	9,543	429.67	2,099.96	0	113,814	4,100,350
inloseamt	Sum of insider loser bet amounts	9,543	3,552.61	21,827.41	0	1,318,809	33,902,557
outloseamt	Sum of outsider loser bet amounts	9,543	2,059.72	10,741.70	5	630,438	19,655,870
inpayoutamt	Sum of insider payout amounts	9,543	4,517.72	28,307.26	0	1,806,495	43,112,573
outpayoutamt	Sum of outsider payout amounts	9,543	2,171.99	11,745.65	0	628,221	20,727,301

We will use this data to calculate proxy model parameters in order to formally test the theoretical model we constructed in chapter 3.

4.5 Operational definition of model parameters

In order to apply our theory to the data, we need to operationally define our theoretical model parameters in terms of the race-level dataset in hand. For each race we determine the theoretical model parameters as follows:

- Number of insiders, N : The count of insider bettors, where any bet of c or above ($c = 100$ New Zealand dollars unless otherwise noted) is considered an insider bet – smaller bets are considered outsider bets.
- Fraction of outsiders betting on horse 1, b_0 : The sum of outsider betting on the winning horse expressed as a fraction of total outsider betting in that race.²
- The takeout rate, τ : The takeout rate is set at 15.5% prior to November 2006 and at 14.5% thereafter – this corresponds to the takeout rates applicable to straight win horse race bets at the time.
- The belief of insiders regarding the probability of horse 1 winning the race, q : We proxy insiders' *ex-ante* belief by the implied market price of their betting in the data; that is, by the ratio of the amount they bet on the winning horse to the total amount they bet in the race.
- The final market probability, \bar{m} : We calculate the empirical final market probability for horse $h = 1$ as the ratio of the total amount bet on horse 1 to the total amount bet on all horses.

The remaining risk preference parameter α is our main variable of interest. Using our theoretical model, we are able to derive an explicit analytical solution for α based on the parameters above. The theoretical final market probability for horse 1 is given by $m = \frac{b_0 + Nb^*}{1 + Nb^*}$. Substituting in the solution for the optimal insider betting amount b^* in the above and solving for α yields an expression for the model implied insider risk preference:

²Our dataset does not allow us to identify the particular horse that a bet relates to. We are able, however, to identify those bets that relate to the winning horse as those are the bets for which we observe a payout, which is recorded in our dataset. We therefore associate horse 1 of the theory with the winning horse in the dataset. There is no additional assumptions implicit in this, as we are concerned with the final market probability of the winning horse, which is determined *before* it is known that a particular horse is in fact the winner. In other words, we are using the winning payout to identify a single horse in the dataset which we identify with horse 1 in the theory.

$$\alpha = \frac{\sqrt{1-q}((b_0N + q - q\tau - N)\bar{m}^2 + (\tau - 1)q(b_0N + b_0 - N + 1)\bar{m} + qb_0 - q\tau b_0)}{\sqrt{q}(\tau - 1)(q - 1)(\bar{m}b_0 - b_0 + \bar{m} - N\bar{m} + \bar{m}Nb_0 - \bar{m}^2)}$$

Thus we calculate the model implied risk preference α using the empirically observed values of N, b_0, τ, q and \bar{m} for each race.

4.6 Empirical tests and results

4.6.1 Overview

We wish to test the hypothesis that insiders are individually risk neutral while allowing for strategic interaction within a parimutuel setting (our null hypothesis – $H_0 : \alpha = 0$). The alternative hypothesis is that insiders are either risk loving or risk averse ($H_a : \alpha \neq 0$). It is important to note that our hypothesis concerns risk preferences at the *individual* agent level – this is in contrast with earlier work (eg. Ali (1977)) in which the risk preferences of a *representative* agent is the subject of investigation. The second point to emphasise is that our hypothesis applies to a strategic interaction setting – that is to say, a situation in which individual payoffs are determined according to a game theoretic equilibrium. We take three different approaches to analysing the data:

The first approach is also the most direct. We calculate the implied risk sensitivity of insiders for each race, using $c = 100$ as the cut-off point between insiders and outsiders. Realising that the cut-off point is somewhat ad-hoc, we then proceed to calculate implied risk sensitivities for cut-off points ranging from NZD 10 to NZD 500. This allows us to measure the effect of the cut-off point on implied risk sensitivity.

In our second approach, we fix the beliefs of insiders at a fixed increment Δq above that of outsiders; that is, we let $q = b_0 + \Delta q$. We then proceed to calculate implied risk sensitivities for Δq ranging from 0 to 0.2 in increments of 0.01. This allows us to measure how risk sensitivity changes as our assumptions regarding insider beliefs relative to that of outsiders (ie Δq) changes.

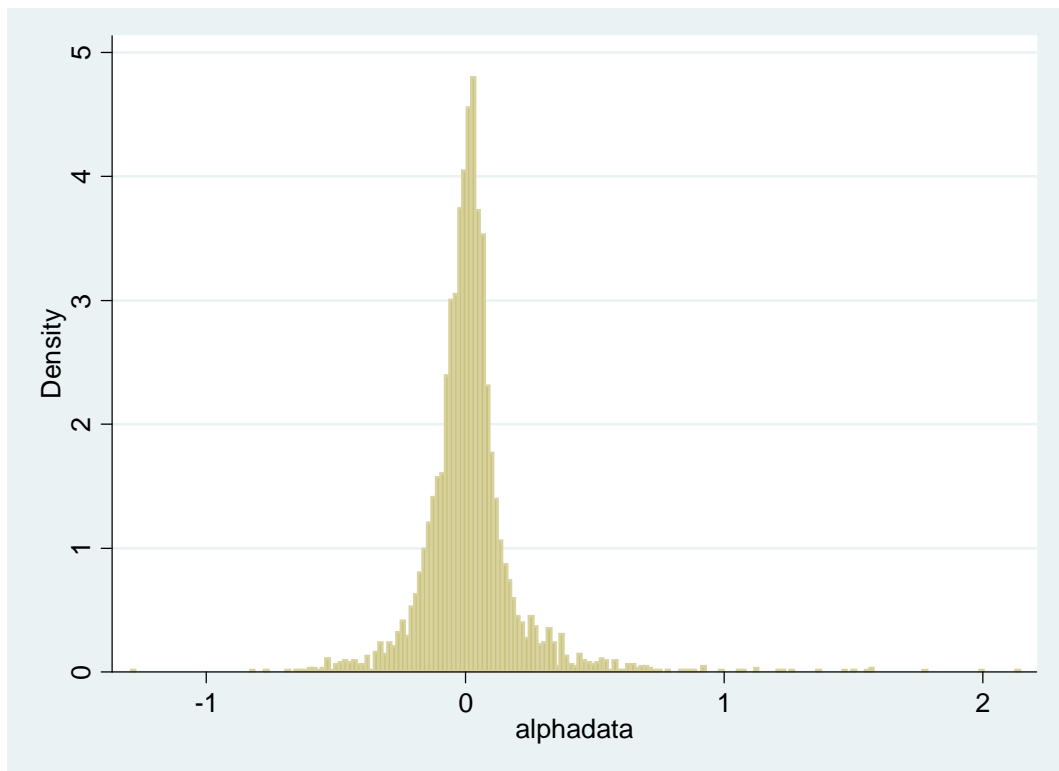
The third approach takes the final market probability m as the starting point. By assuming different values for risk sensitivity (α) and insider beliefs relative to outsiders (Δq), we can calculate the theoretical final market probability predicted by our model, $m_{\alpha,q}$, as well as that observed directly from betting data, \bar{m} . By regressing \bar{m} on $m_{\alpha,q}$, we can identify those combinations of risk sensitivity (α) and

inside information (Δq) that yields theoretical predictions that best corresponds with empirical observations.

4.6.2 Direct estimation of risk sensitivity

Implied risk sensitivity is calculated for each race as set out in sub-section 4.5 above, assuming a cut-off point of $c = 100$ between insiders and outsiders. We restrict ourselves to those races that conform with the assumptions of our theoretical model³ and find that our implied risk sensitivity α is approximately normally distributed with a mean of 0.0188 and a standard deviation of 0.1894. A histogram of implied risk sensitivities are shown in Figure 4.6.1 below.

Figure 4.6.1: Histogram of implied α , at an insider/outsider cut-off point of $c = 100$ (Model compliant observations only)



The mean implied risk sensitivity of 0.0188 indicates slight risk love on the part of insider bettors – a dollar of standard deviation is valued at 1.88% of a dollar of expected return. Although the mean implied risk sensitivity is statistically significant⁴, it is not economically significant given that it amounts to less than 2% of the

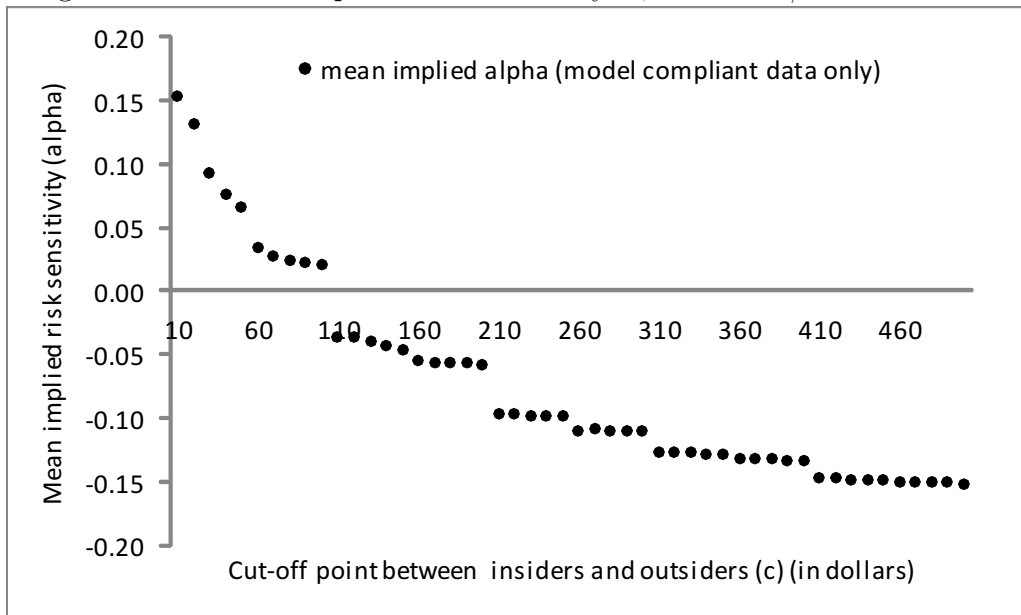
³In particular, the theoretical model requires $q > b_0$, which holds for 3,646 of the 5,654 races for which α can be calculated.

⁴The t-statistic for the mean risk sensitivity against a null hypothesis of zero is given by $\frac{\alpha}{stdev/\sqrt{n}} = \frac{0.0188}{0.1894/\sqrt{3646}} = 5.998$, which indicates that the mean is statistically significant at a confidence level of 1% or above.

expected net payout of the insider.

We noted earlier the rather ad-hoc nature our decision to use a bet size of NZD 100 as the cut-off point between insiders and outsiders. It would be interesting to see how implied risk sensitivity varies with different choices of c as the cut-off point between insiders and outsiders. To address this, we recalculated our implied risk sensitivities for different cut-off points, with c ranging from NZD 10 to NZD 500 in NZD 10 increments. The result of this analysis, which includes is illustrated in Figure 4.6.2 below.

Figure 4.6.2: Mean implied risk sensitivity α_c vs insider/outsider cut-off point c



We can see that implied risk sensitivity is influenced by the cut-off point we select. In particular, implied risk sensitivity tends to decrease as we impose higher cut-off points, but at a decreasing level. This pattern is consistent with smaller, risk-loving recreational bettors and larger, relatively more risk averse, sophisticated insider bettors. As we move the cut-off point higher (and to the right in the graph), we are capturing fewer recreational risk-loving bettors in our definition of insiders and consequently the implied risk-love of our insiders starts to decrease.

4.6.3 Estimation conditional on specified insider beliefs

4.6.3.1 Method

As before, we wish to test the hypothesis that insiders are individually risk neutral. In contrast with the approach taken in sub-section 4.6.2, we calculate implied risk sensitivity based on an assumed level of insider beliefs. More exactly, we set the

level of insider beliefs equal to that of outsiders plus a positive adjustment Δq , that is $q = b_0 + \Delta q$. As before, b_0 is calculated directly from the data. The quantity Δq is formally the difference in belief between insiders and outsiders regarding the probability of horse 1 winning the race. Informally, we can think of Δq as an indication of the inside information insiders have relative to outsiders; this is based on the idea that if insiders have a stronger belief in the winning probability of horse 1, then this must be due to their having better information. We will often refer to Δq as “inside information” hereafter; however keep in mind that formally $\Delta q = q - b_0$.

This approach suffers from the drawback that one must assume some value of Δq – hence, each estimate of risk sensitivity α will be conditional on the assumed value of Δq (to emphasise this point we will denote the implied risk sensitivity conditional on Δq as $\alpha_{\Delta q}$). The benefit of this approach is that it allows us to analyse more directly the relationship between inside information and risk preferences. Another benefit is that we avoid some of the data issues in the previous section – our analysis can utilise a much larger set of races since we avoid problems posed by races where $q < b_0$. We assume a cut-off point of $c = 100$ in classifying insiders and outsiders, unless otherwise noted.

4.6.3.2 Results

Using the formula for α set out in sub-section 4.5 above, we calculate the implied risk sensitivity $\alpha_{\Delta q}$ for each race in the race level dataset for a range of values of $q = b_0 + \Delta q$ where $\Delta q \in \{0, 0.01, \dots, 0.19, 0.20\}$. The results are shown in Table 4.4.

Table 4.4: Implied risk sensitivity $\alpha_{\Delta q}$ for a range of Δq

Δq	N	Implied risk sensitivity $\alpha_{\Delta q}$				t-Statistic*
		Mean	Std. Dev.	Min	Max	
0.0000	8577	0.2338	0.6450	-1.4557	8.6528	33.57
0.0100	8685	0.2243	0.7689	-1.5395	11.6543	27.19
0.0200	8685	0.1743	0.6543	-1.6335	8.2114	24.83
0.0300	8685	0.1361	0.6040	-1.7403	6.6804	21.00
0.0400	8684	0.1029	0.5739	-2.3319	5.7644	16.71
0.0500	8682	0.0737	0.5536	-2.6068	5.1370	12.40
0.0600	8678	0.0449	0.5430	-7.6602	4.6722	7.70
0.0700	8674	0.0197	0.5262	-2.3983	4.3096	3.49
0.0800	8672	-0.0039	0.5316	-2.6770	11.5234	-0.68
0.0900	8666	-0.0293	0.5164	-3.2297	8.2036	-5.28
0.1000	8657	-0.0537	0.5012	-3.6322	3.5653	-9.97
0.1100	8653	-0.0770	0.5004	-4.6551	3.3865	-14.31
0.1200	8647	-0.1006	0.5150	-13.1717	3.2299	-18.16
0.1300	8633	-0.1206	0.4867	-3.9475	3.0912	-23.02
0.1400	8627	-0.1449	0.5359	-21.0616	3.2980	-25.11
0.1500	8617	-0.1641	0.4808	-6.7713	2.8554	-31.68
0.1600	8615	-0.1849	0.4812	-5.6823	2.7539	-35.66
0.1700	8607	-0.2034	0.4737	-3.2428	2.7528	-39.84
0.1800	8599	-0.2251	0.4940	-11.2598	6.4871	-42.25
0.1900	8582	-0.2477	0.4840	-6.1525	5.2286	-47.41
0.2000	8569	-0.2687	0.4988	-14.4026	2.4240	-49.87

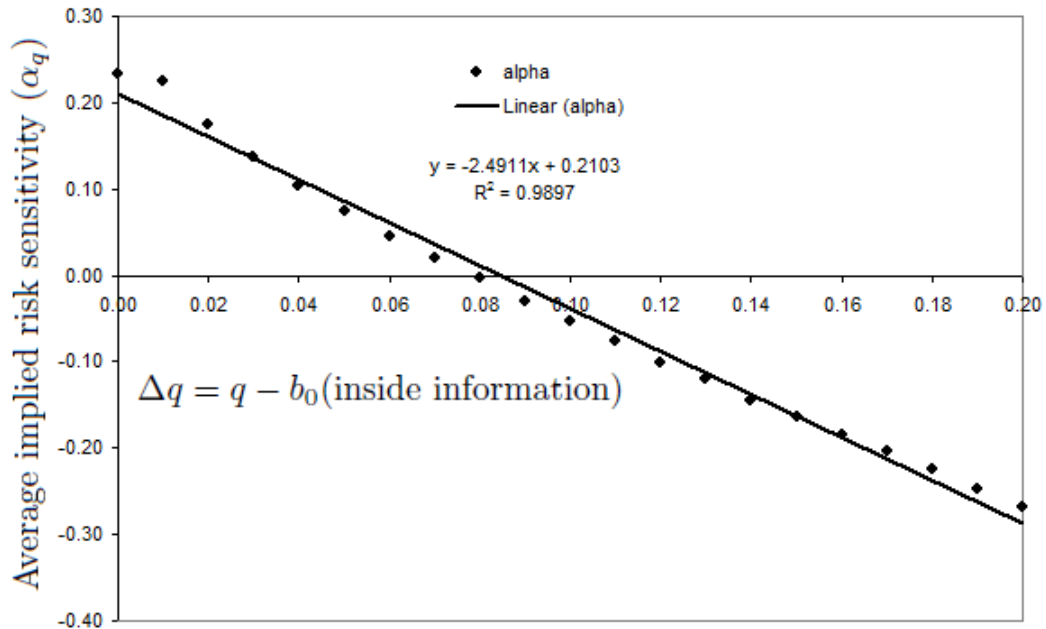
* Mean of $\alpha_{\Delta q}$ vs null hypothesis of $\alpha_{\Delta q} = 0$; t-stat = $\frac{\alpha_{\Delta q}}{stdev/\sqrt{n}}$

From the table above we can see that risk sensitivity is strictly decreasing in inside information (Δq), with risk neutrality ($\alpha_{\Delta q} = 0$) corresponding to $\Delta q = 0.08$. This suggests that inside information and risk preference are in a sense substitutes (an intuition we first hinted at in sub-section 3.4.3 when we discussed insider beliefs). The actual values of $\alpha_{\Delta q}$ for individual races are quite widely distributed, as can be seen from the associated standard deviations. The last column in the table shows the t-statistic calculated for the mean of $\alpha_{\Delta q}$ against the null hypothesis of $\alpha_{\Delta q} = 0$ – this indicates how statistically different from zero our estimates of the mean of $\alpha_{\Delta q}$ are. The mean risk sensitivity $\alpha_{\Delta q}$ is statistically different from zero at the 5% confidence level for each level of assumed inside information Δq save one: conditional on $\Delta q = 0.08$, the implied risk sensitivity is statistically indistinguishable from zero. In other words, we cannot reject the hypothesis that insiders are risk neutral, conditional on $\Delta q = 0.08$. It is intriguing to note the close correspondence between $\Delta q = 0.08 = 8\%$ and 9.2%, the degree by which insiders outperform outsiders in the data (see sub-section 4.3.1). We may interpret this as providing tentative support

for the notion that insiders are close to risk neutral on average.

In Figure 4.6.3 we plot the mean of model implied risk preferences α_q against Δq based on the data presented in 4.4.

Figure 4.6.3: Mean of model implied risk sensitivity α_q as a function of Δq



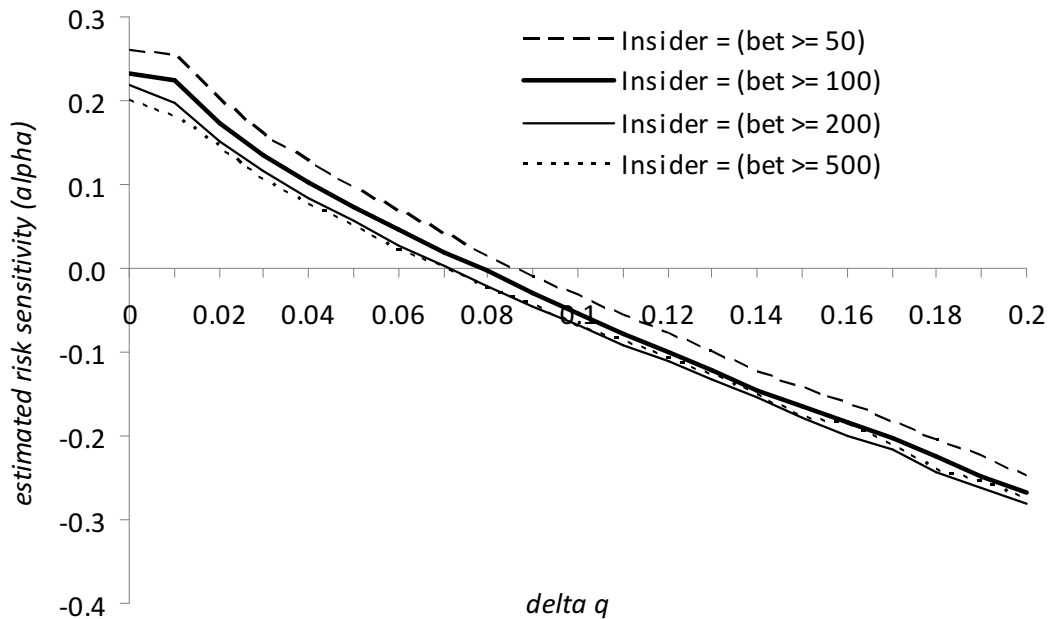
The strong linear relationship between the mean risk sensitivity $\alpha_{\Delta q}$ and Δq is striking. The association between risk aversion and inside information suggested by the data is given by the linear relationship

$$\alpha_{\Delta q} = -2.4911(\Delta q) + 0.2103$$

We could interpret this line as an indifference curve between higher risk love ($\alpha_{\Delta q}$) and better inside information (Δq). A cursory examination of the formula we use to calculate $\alpha_{\Delta q}$ does not immediately suggest that we should expect such a linear relationship only on the basis of the theoretical model. This suggests that this relationship is a feature of the data, rather than of our theoretical model.

It is reasonable to ask whether our results are driven by our choice of NZD 100 as the cut-off point between insiders and outsiders. To address this concern, we re-did the above analysis using three additional cut-off points, namely NZD 50, NZD 200 and NZD 500, in addition to our assumed cut-off point of NZD 100. We plot the results in Figure 4.6.4.

Figure 4.6.4: Estimated risk sensitivity α_q against Δq for different insider bet cut-off points



We find that the linear relationship described above remains intact. We also find that for the lower cut-off point (NZD 50) risk neutrality is associated with a slightly higher Δq of 0.085, and similarly, for the higher cut-off points (NZD 200 and NZD 500) risk neutrality is associated with a slightly lower Δq of 0.07. Hence the actual level of inside information (Δq) that corresponds to implied risk neutrality is sensitive to the cut-off point between insiders and outsiders. The pattern we observe suggests that raising the cut-off point lowers the degree of inside information that is consistent with risk neutrality. We can only speculate why this appears to be the case – one possible explanation might be that recreational bettors bet smaller amounts of money and are also on average risk loving. If we lower the cut-off point from NZD 100 to NZD 50, we may be capturing a larger number of recreational risk-loving bettors in our definition of insiders, thereby resulting in a higher (more risk loving) implied risk sensitivity α at each level of inside information Δq for such insiders. We leave this intriguing possibility as a topic for further research.

4.6.4 Model vs empirical final market probabilities

4.6.4.1 Method

We first calculate the empirical (observed) final market probability, which we denote by \bar{m} . We then proceed to calculate the theoretical final market probability according to our model using the model parameters N, b_0, τ as determined above, along

with *assumed* values of α and q to arrive at a theoretical final market probability $m_{(\alpha,q)}$. (We explicitly subscript $m_{(\alpha,q)}$ to remind ourselves that this theoretically calculated value is dependent on the assumed values for α and q). By choosing a suitable range of assumed values for α and q , we proceed to generate a set of candidate theoretical solutions such that:

$$M = \{m_{(\alpha,q=b_0+\Delta q)} \mid \alpha \in [-0.5, -0.48, \dots, 0.48, 0.50], \Delta q \in \{0, 0.05, 0.1, 0.15, 0.2\}\}.$$

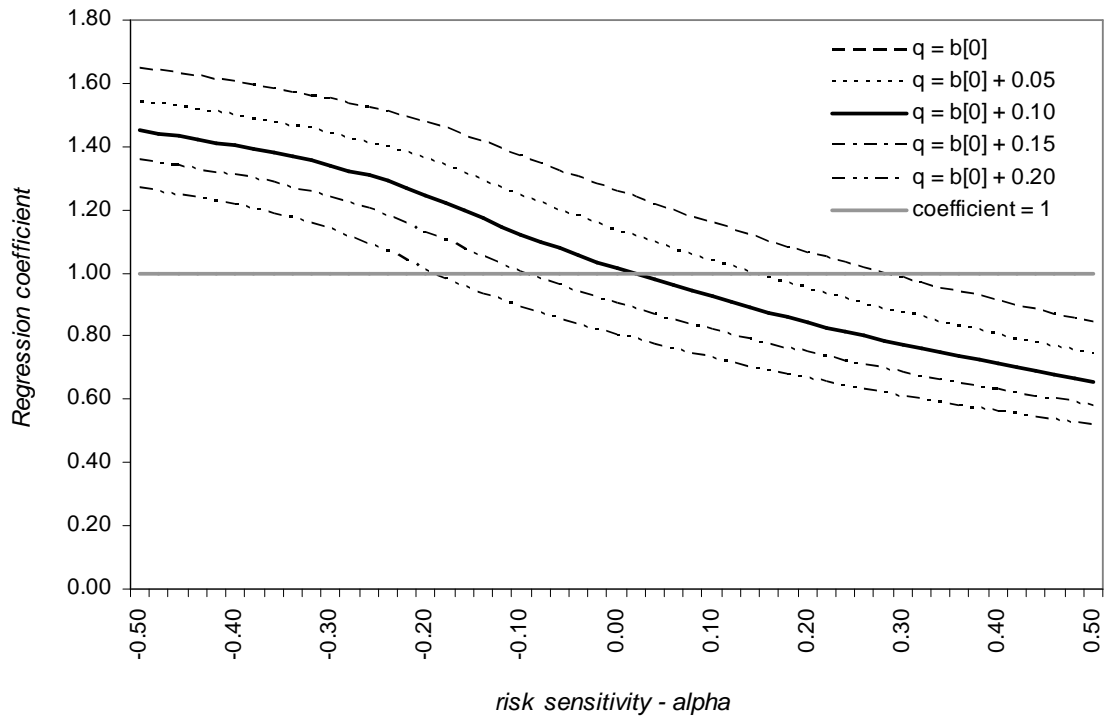
We would like to test which combinations of α and q generate a theoretical final market probability $m_{(\alpha,q)}$ which is most congruent with the empirically observed final market probability \bar{m} . To do this we estimate an OLS equation in the form

$$\bar{m} = \gamma m_{(\alpha,q)} + \varepsilon$$

which yields an estimate of the empirical final market probability $\hat{m} = \gamma m_{(\alpha,q)}$ for each element of M . (Note that we do not include an intercept in our model). If some $m_{(\alpha,q)}$ is congruent with \bar{m} then we would expect to obtain $\gamma = 1$ in the above regression. This allows us to test all of the candidate theoretical solutions in M and to identify for each q the corresponding α for which $\gamma = 1$ in a statistical sense, or equivalently, $\bar{m} = m_{(\alpha,q)} + \varepsilon$. Let $M_f \subset M$ be those elements of M for which the regression coefficient γ is statistically closest to 1 for a given choice of q . We can think of M_f as the set of feasible theoretical solutions, in the sense that they accord most closely with the observed final market probability \bar{m} in a statistical sense. Equivalently, we can interpret M_f as a set of combinations of α and q that corresponds most closely with \bar{m} . That is, $\{(\alpha, q) \mid m_{(\alpha,q)} \in M_f\}$ is the set of combinations of α and q which are statistically most congruent with the observed pattern of final market probabilities.

4.6.4.2 Results

The results of the above analysis is presented in Figure 4.6.5 below. The graph plots the regression coefficient γ against different assumed values of risk sensitivity α . Each of the different lines in the graph is associated with a level of q determined according to the key shown in the graph. Each point at which a line intercepts the horizontal grey line $y = 1$ shows a combination of risk preference α (according to the x-axis) and insider beliefs q (according to the specific line) which resulted in a regression coefficient of $\gamma = 1$. The intersection points are, in a sense, combinations of α and q jointly estimated from the data.

Figure 4.6.5: Regression coefficient γ against risk sensitivity α for different q 

It is intriguing to note that the middle line, which corresponds to $q = b_0 + 0.1$, intersects with the grey line ($\gamma = 1$) very close to the point where $\alpha = 0$. We could interpret this as evidence that the data fails to reject our hypothesis that insiders are risk neutral, or $\alpha = 0$. This interpretation depends, of course, on the admittedly strong assumption that $q = b_0 + 0.1$. More generally, this would suggest that where insiders have better inside information (say $q = b_0 + 0.2$) they are willing to be more risk loving (intercept with $y = 1$ at $\alpha = 0.3$) and conversely, where insiders have little inside information (say $q = b_0$) they will tend to be more risk averse (intercept with $y = 1$ at $\alpha = -0.2$).

For each choice of q we present below in Table 4.5 the choice of α for which the regression did not reject $\gamma = 1$. We present both in-sample results (based on 80% of the data randomly chosen) and out-of-sample results (based on the remaining 20% of the data); the out-of-sample results show the regression of predicted final market probabilities using in-sample estimated coefficients ($\hat{m}_{IS} = \gamma_{IS} m_{(\alpha, q)}$) against the out-of-sample empirical final market probabilities (\bar{m}).

Table 4.5: Best fit regression results for α and q

q	b_0	$b_0 + 0.05$	$b_0 + 0.10$	$b_0 + 0.15$	$b_0 + 0.20$
α	0.28	0.14	0.02	-0.1	-0.2

In-sample ($\bar{m} = \gamma_{IS}m_{(\alpha,q)} + \varepsilon$)					
γ_{IS}	0.999	1.002	0.997	1.003	1.009
t-Stat	176	175	170	167	166
Lower CI (5%)	0.988	0.990	0.985	0.991	0.997
Upper CI (5%)	1.010	1.013	1.008	1.015	1.020
Adj R-square	0.8191	0.8048	0.7962	0.7900	0.7893
N	6815	7440	7423	7391	7350

Out-of-sample ($\bar{m} = \eta_{OS}\hat{m}_{IS} + \varepsilon$, where $\hat{m}_{IS} = \gamma_{IS}m_{(\alpha,q)}$)					
η_{OS}	0.997	0.997	0.994	1.003	1.002
t-Stat	88	88	84	83	83
Lower CI (5%)	0.975	0.974	0.971	0.979	0.978
Upper CI (5%)	1.020	1.019	1.018	1.027	1.025
Adj R-square	0.8131	0.7992	0.7868	0.7839	0.7847
N	1760	1924	1916	1908	1901

The out-of-sample results provide some comfort that the results we obtained are not the result of data mining. Again, the above results are consistent with the idea that risk preference and inside information are “substitute goods” from the perspective of insiders.

Chapter 5

Conclusion

A broad survey of the literature dealing with parimutuel betting was conducted. Based on this survey, we identified a gap in the existing literature – the lack of any theoretical models that combine both individual risk preferences on the part of bettors and the strategic interaction between bettors. We construct a theoretical model of parimutuel betting which includes both strategic interaction between insiders and arbitrary common risk preferences of individual insiders. By solving this model for the static symmetric Nash equilibrium in continuous strategies, we obtain a closed-form solution to the optimal insider betting amount and the associated final market probabilities. Within the context of this model, we find that individual insider risk preferences and insider inside information operate as close substitutes, in the sense that a decrease in risk appetite can be compensated for by an increase in inside information and vice versa.

Using a unique dataset of 1.6 million individual horse race bets, we document a positive empirical association (not previously reported in the literature) between the log of average bet size and average payout rates in bet level data aggregated in percentiles by bet size. On the basis of this regularity, we operationalise the distinction between insider and outsider bettors by classifying all bets of NZD 100 or above as insider bets and the remainder as outsider bets. In so doing, we are able to construct a race-level dataset containing all of the model parameters with the exception of risk preference and inside information. This allows us to formally test our hypothesis that insiders are risk neutral within the context of our theoretical model. We are unable to reject the hypothesis that insiders are risk neutral, conditional on assuming a degree of inside information that corresponds to the degree to which insiders actually outperform outsiders in our dataset. This provides some empirical support to existing strategic interaction models of parimutuel markets that assume risk neutrality on the part of informed bettors.

Bibliography

- Adams, B.R., Rusco, F.W. and Walls, W.D., ‘Professional bettors, odds arbitrage competition and betting market equilibrium’, *Journal of Singapore Economic Review*, 47 (2002), pp. 11–27.
- Ali, M.M., ‘Probability and utility estimates for racetrack bettors’, *The Journal of Political Economy*, 85 (1977):4, pp. 803–815.
- Asch, P., Malkiel, B.G. and Quandt, R.E., ‘Racetrack betting and informed behavior’, *Journal of Financial Economics*, 10 (1982):2, pp. 187–194.
- Busche, K. and Hall, C.D., ‘An exception to the risk preference anomaly’, *Journal of business*, 61 (1988):3, pp. 337–346.
- Busche, K. and Walls, W.D., ‘Decision costs and betting market efficiency’, *Rationality and Society*, 12 (2000):4, p. 477.
- Chadha, S. and Quandt, R.E., ‘Betting bias and market equilibrium in racetrack betting’, *Applied Financial Economics*, 6 (1996):3, pp. 287–292.
- Cochrane, J.H., *Asset pricing*, (Princeton University Press Princeton, NJ, 2001).
- Coleman, L., ‘New light on the longshot bias’, *Applied Economics*, 36 (2004):4, pp. 315–326.
- Conlisk, J., ‘The utility of gambling’, *Journal of Risk and Uncertainty*, 6 (1993):3, pp. 255–275.
- Eisenberg, E. and Gale, D., ‘Consensus of subjective probabilities: The pari-mutuel method’, *The Annals of Mathematical Statistics*, 30 (1959):1, pp. 165–168.
- Fama, E.F., ‘Efficient capital markets: A review of theory and empirical work’, *Journal of finance*, (1970), pp. 383–417.
- Fama, E.F., ‘Efficient capital markets: II’, *Journal of finance*, 46 (1991):5, pp. 1575–1617.

- Feeney, R. and King, S.P., 'Sequential parimutuel games', *Economics Letters*, 72 (2001):2, pp. 165–173.
- Gandar, J.M., Zuber, R.A. and Dare, W.H., 'The search for informed traders in the totals betting market for national basketball association games', *Journal of Sports Economics*, 1 (2000):2, p. 177.
- Golec, J. and Tamarkin, M., 'Bettors love skewness, not risk, at the horse track', *Journal of Political Economy*, 106 (1998):1, pp. 205–225.
- Gramm, M. and McKinney, C.N., 'The effect of late money on betting market efficiency', *Applied Economics Letters*, 16 (2009):4, pp. 369–372.
- Gramm, M. and Owens, D.H., 'Determinants of betting market efficiency', *Applied Economics Letters*, 12 (2005):3, pp. 181–185.
- Griffith, R.M., 'Odds adjustments by American horse-race bettors', *The American Journal of Psychology*, 62 (1949):2, pp. 290–294.
- Hausch, D.B. and Ziemba, W.T., *Efficiency of sports and lottery betting markets*, (Elsevier, 1995), pp. 545–580.
- Hausch, D.B., Ziemba, W.T. and Rubinstein, M., 'Efficiency of the market for race-track betting', *Management science*, (1981), pp. 1435–1452.
- Hurley, W. and McDonough, L., 'A note on the Hayek hypothesis and the favorite-longshot bias in parimutuel betting', *The American economic review*, (1995), pp. 949–955.
- Isaacs, R., 'Optimal horse race bets', *American Mathematical Monthly*, 60 (1953):5, pp. 310–315.
- Kahneman, D. and Tversky, A., 'Prospect theory: An analysis of decision under risk', *Econometrica: Journal of the Econometric Society*, 47 (1979):2, pp. 263–291.
- Koessler, F., Noussair, C. and Ziegelmeyer, A., 'Parimutuel betting under asymmetric information', *Journal of mathematical Economics*, 44 (2008):7-8, pp. 733–744.
- Levitt, S.D., 'Why are gambling markets organised so differently from financial markets?' *Economic Journal*, 114 (2004), pp. 223–246.
- McGlothlin, W.H., 'Stability of choices among uncertain alternatives', *The American Journal of Psychology*, 69 (1956):4, pp. 604–615.
- Ottaviani, M. and Sørensen, P.N., *Parimutuel versus fixed-odds markets*, 2005.

- Ottaviani, M. and Sørensen, P.N., *The Timing of Parimutuel Bets*, 2006.
- Ottaviani, M. and Sørensen, P.N., *The Favorite-Longshot Bias: An Overview of the Main Explanations*, (Citeseer, 2008), pp. 83–101.
- Ottaviani, M. and Sørensen, P.N., ‘Surprised by the Parimutuel Odds?’ *American Economic Review*, 99 (2009):5, pp. 2129–2134.
- Ottaviani, M. and Sørensen, P.N., ‘Noise, Information, and the Favorite-Longshot Bias in Parimutuel Predictions’, *American Economic Journal: Microeconomics*, 2 (2010):1, pp. 58–85.
- Qiu, J., ‘Loss aversion and mental accounting: the favorite longshot bias in parimutuel betting’, *Jena Economic Research Papers*, (2007), p. 017.
- Quandt, R.E., ‘Betting and equilibrium’, *The Quarterly Journal of Economics*, 101 (1986):1, pp. 201–207.
- Sauer, R.D., ‘The economics of wagering markets’, *Journal of Economic Literature*, 36 (1998):4, pp. 2021–2064.
- Sauer, R.D., ‘The state of research on markets for sports betting and suggested future directions’, *Journal of Economics and Finance*, 29 (2005):3, pp. 416–426.
- Shin, H.S., ‘Optimal betting odds against insider traders’, *The Economic Journal*, 101 (1991):408, pp. 1179–1185.
- Shin, H.S., ‘Prices of state contingent claims with insider traders, and the favourite-longshot bias’, *The Economic Journal*, 102 (1992):411, pp. 426–435.
- Snowberg, E. and Wolfers, J., ‘Explaining the favourite-longshot bias: is it risk-love or misperceptions?’ *NBER Working Paper* April (2010).
- Thaler, R.H. and Ziemba, W.T., ‘Anomalies: Parimutuel betting markets: Race-tracks and lotteries’, *The Journal of Economic Perspectives*, 2 (1988):2, pp. 161–174.
- Vaughan Williams, L., ‘Information efficiency in betting markets: A survey’, *Bulletin of Economic Research*, 51 (1999):1, pp. 1–39.
- Walls, W.D. and Busche, K., ‘Broken odds and the favourite-longshot bias in parimutuel betting: a direct test’, *Applied Economics Letters*, 10 (2003):5, pp. 311–314.

Watanabe, T., Nonoyama, H. and Mori, M., 'A model of a general parimutuel system: characterizations and equilibrium selection', *International Journal of Game Theory*, 23 (1994):3, pp. 237–260.

Weitzman, M., 'Utility analysis and group behavior: An empirical study', *The Journal of Political Economy*, 73 (1965):1, pp. 18–26.

Chapter 6

Appendix

6.1 Maple code to verify derivation of equilibrium insider betting amounts

```
restart; _EnvExplicit := true; printlevel:=1;_EnvTry := normal;
assume(b[0] > 0, b[0] < 1, b[0]::real, q > b[0], q < 1, q::real,
tau >= 0, tau < 1, tau::real, b[1] >= 0, b[1]::real,
b[n] >= 0, b[n]::real, N >= 2, N::integer,
b[0]+b[1]+b[n]*N-b[n]>0, 1+b[1]+b[n]*N-b[n]>0);
pi := ( b[1]/(b[0]+b[1]+(N-1)*b[n]) ) *
(1+(b[1])+(N-1)*(b[n])) * (1-tau);
E_u := q*(-b[1]+pi)+(1-q)*(-b[1]);
VAR_u := q*((-b[1]+pi)^2)+(1-q)*((-b[1])^2) - ((E_u)^2);
u := E_u + alpha*(VAR_u^(1/2));
d_u := diff(u, b[1]);
foc_b:=d_u=0; foc_b:=simplify(foc_b);
d2_u := diff(u, b[1]$2);
soc_b:=d2_u; soc_b:=simplify(soc_b);
sol := solve([foc_b, b[n]=b[1]], [b[1],b[n]]);
b_star1:=simplify(rhs(sol[1][1]));
b_star2:=simplify(rhs(sol[2][1]));
b_star:=b_star2;
# solves foc(b_star)=0, meets b_star > 0 and soc(b_star) < 0
m:=(b[0]+N*b_star)/(1+N*b_star); m:=simplify(m);
mispricing:=simplify(m-q);
```

6.2 Average payout rate and average bet size by percentile

Table 6.1 below is a result of the analysis described in sub-section 4.3. Note that although the analysis is based on a percentile partition of the bet amount data, we do not have exactly 100 data points as one might expect – this is because of the large amount of small, “round” bets such as NZD 5 and NZD 10, each of which accounts for much more than one per cent of the total betting volume. Since there is no point in having multiple data points relating to NZD 5 bets, we group bets of the same size in the same percentile grouping, even if this results in fewer data points.

Table 6.1: Bets and payouts by percentile

Percentile	N	Total bet amt	Total payout amt	Minimum bet	Maximum bet	Average bet	Payout ratio
1	461,076	2,305,380	1,893,979	5	5	5.00	0.8215
29	34,853	209,118	181,404	6	6	6.00	0.8675
31	22,085	154,595	130,705	7	7	7.00	0.8455
32	21,664	173,312	146,409	8	8	8.00	0.8448
34	366,844	3,658,188	3,103,879	9	10	9.97	0.8485
56	16,934	196,802	165,542	11	12	11.62	0.8412
57	11,684	157,914	138,300	13	14	13.52	0.8758
58	40,145	602,175	538,574	15	15	15.00	0.8944
60	8,038	131,924	116,131	16	17	16.41	0.8803
61	200,469	3,998,540	3,470,055	18	20	19.95	0.8678
73	10,130	228,145	195,788	21	24	22.52	0.8582
74	27,168	679,200	604,341	25	25	25.00	0.8898
75	5,441	146,736	126,885	26	28	26.97	0.8647
76	69,009	2,069,180	1,799,637	29	30	29.98	0.8697
80	12,701	435,781	384,669	31	37	34.31	0.8827
81	38,523	1,538,300	1,369,299	38	40	39.93	0.8901
83	89,764	4,454,158	3,944,864	41	50	49.62	0.8857
88	2,870	154,263	130,915	51	55	53.75	0.8487
89	15,929	960,119	854,437	56	65	60.27	0.8899
90	16,093	1,210,089	1,149,748	66	84	75.19	0.9501
91	72,196	7,187,603	6,585,179	85	100	99.56	0.9162
95	15,897	2,146,172	2,062,618	101	150	135.00	0.9611
96	27,055	5,324,042	4,862,173	151	200	196.79	0.9132
98	16,954	4,683,929	4,563,626	201	300	276.27	0.9743
99	18,352	8,221,399	8,164,420	301	500	447.98	0.9931
100	12,934	18,000,000	17,200,000	501	40000	1391.68	0.9556
Sum	1,634,808	69,027,064	63,883,577		Average		0.9255

6.3 Statics based on an empirical approximation

In chapter 3 we obtained an analytic solution to our model. Unfortunately, the complexity of this solution renders the calculation of analytic model statics infeasible. Instead, we pursue an alternative approach – applying an OLS regression to our race level dataset to estimate a reduced form approximation to the analytic solution. Such an approximate solution can then be used to gain an understanding of aggregate model statics. In constructing the race level dataset, we operationalised the model parameters as set out in detail in sub-section 4.5. Using these parameters we calculated a “model implied” final market probability m based on the foregoing parameters (as detailed more fully in sub-section 4.6.3). We then proceeded to estimate two specifications. The first is simply a linear combination of model parameters (excluding q , since it is a linear function of b_0).

$$\hat{m}_1 = \beta_1 N + \beta_2 b_0 + \beta_3 \tau + \beta_4 \alpha + \varepsilon$$

The second includes squared model parameter terms in order to capture any second order effects in the data.

$$\hat{m}_2 = \beta_1 N + \beta_2 b_0 + \beta_3 \tau + \beta_4 \alpha + \beta_5 N^2 + \beta_6 b_0^2 + \beta_7 \tau^2 + \beta_8 \alpha^2 + \varepsilon$$

The results of these regressions are set out in in Table 6.2 below. Notably, these simple specifications are able to account for 96.75% and 98.42% of the variation in model final market probabilities respectively. This compares with 91.98% when regressing the model implied final market probability on outsider betting implied market probability, that is, $\hat{m}_0 = \beta_1 b_0 + \varepsilon$. As one would expect, the betting patterns of outsiders have a substantial impact on the model implied final market probability; that said, the additional model parameters clearly add further explanatory power to the specification.

Table 6.2: Empirical approximation of model final market probabilities

Variable	\hat{m}_0		\hat{m}_1		\hat{m}_2	
	Coefficient	t-Stat	Coefficient	t-Stat	Coefficient	t-Stat
N			0.000071	1.75	0.000245	2.63
b_0	1.087512	248.91	1.018508	254.26	1.041681	132.82
τ			-0.059058	-7.85	-0.320130	-1.21
α			0.123297	89.08	0.209928	139.20
N^2					-0.0000003	-2.51
b_0^2					-0.120900	-11.20
τ^2					1.693191	0.93
α^2					-0.025373	-74.34
N	5,402		5,402		5,402	
Adj R ²	91.98%		96.75%		98.42%	

Based on the regression results above, we could draw the following conclusions regarding the *aggregate* statics of our theoretical model based on the race level dataset:

1. Final market probability is increasing in outsider betting b_0 , but at a decreasing rate (negative coefficient of b_0^2). (This effect is statistically insignificant in the first specification which does not include squared terms)
2. Final market probability is increasing the number of insiders N , but at a slowly decreasing rate (small negative coefficient of N^2)
3. Final market probability is decreasing in the takeout rate τ , however, this effect is statistically insignificant when including τ^2 in the specification.
4. Final market probability is increasing in insider risk preference α at a decreasing rate (negative coefficient of α^2)

6.4 Stata code for estimating reduced form final market probability (\hat{m})

```

set matsize 500

cd D:\Paul\Masters\Data
use horse racing_collapsed, clear

// ** tau [t] **
// takeout rate was 15.5% before Nov 2006, and 14.5% after
gen t=0.155 if tdate_e < date("1/11/2006","dmy")
replace t=0.145 if tdate_e >= date("1/11/2006","dmy")

// ** M DATA [mdata] **
gen mdata = (inwinamt+outwinamt)/(sbets)

// ** insider m **
gen mdatainsider = (inwinamt)/(inwinamt+inloseamt)

// ** N [n] **
gen n = insider

// ** b0 [b] **
gen b = (outwinamt)/(outwinamt+outloseamt)

// ** alpha [a] **
gen a = (1-q)^(1/2)*((b*n+q-q*t-n)*mdata^2+(-1+t)*q*(b*n+b-n+1)*mdata+q
    *b-q*t*b)/q^(1/2)/
(-1+t)/(-1+q)/(mdata*b-b+mdata-n*mdata+mdata*n*b-mdata^2)

// ** b_star [b_star] **

gen K = (b+2*b*n+n^2*b-1-n^2+2*n)*(1-t)
gen A = 1/(2*n^2*(-a*(1-t)*q^2+(1-q)^(1/2)*(1-t)*q^(3/2)+a*(1-t)*q-q
    ^ (1/2)*(1-q)^(1/2)))
gen BB = ((1-b)*(1-t)*q^(3/2)*(1-q)*((-1+a^2)*K*q^(3/2)-2*(1-q)^(1/2)*a
    *K*q+(-K*a^2+4*b*n)
    *q^(1/2)+4*a*(1-q)^(1/2)*b*n))^(1/2)
gen C = (1-t)*q*(a*(-1+q)-q^(1/2)*(1-q)^(1/2))*(-1+b+b*n+n)+2*q^(1/2)
    *(1-q)^(1/2)*b*n

// optimal insider betting amount = bstar = bstar2
gen b_star = A*(C-BB)

// drop 5% outliers
_pctile b_star, percentiles(5(5)95)

```

```
drop if b_star<r(r1)
drop if b_star>r(r19)

// ** THEORETICAL M **
gen m=(b+b_star*n)/(1+b_star*n)

gen n_n = n*n
gen b_b = b*b
gen t_t = t*t
gen a_a = a*a

// approximate theoretical m using model parameters
reg m n b t a, noconst
reg m n b t a n_n b_b t_t a_a, noconst
```

6.5 Stata code for direct empirical testing of predicted final market probability

```
// STATA script to test data
set matsize 500
cd D:\Paul\Masters\Data
use horse racing_collapsed, clear
// add takeout rate "tau"
// takeout rate was 15.5% before Nov 2006, and 14.5% after
gen tau=0.155 if tdate_e < date("1/11/2006","dmy")
replace tau=0.145 if tdate_e >= date("1/11/2006","dmy")
// generate empirical final market price
gen mdata = (inwinamt+outwinamt)/(sbets)

foreach qdelta of numlist 0(0.01)0.2 {

    // # Generate parameters (N, b0, tau, q) and theoretical values
    // (blt, mt — postfixed w "t" for theory)
    // N is the number of insiders — since CAPITAL N is reserved by
    // stata we use small n
    gen n = insider
    // b0 is the ratio of winning horse bets to all bets for
    // outsiders only
    gen b0 = (outwinamt)/(outwinamt+outloseamt)
    // Adds qdelta to q in order to ensure that q > critical_q as
    // required
    gen q = min(1,b0+'qdelta')
    local qdeltalabel = round(('qdelta')*100,1)

    gen alphadata 'qdeltalabel' = (1-q)^(1/2)*((b0*n+q-q*tau-n)*
    mdata^2+(-1+tau)*q*(b0*n+b0-n+1)*mdata+q*b0-q*tau*b0)/q
    ^ (1/2)/(-1+tau)/(-1+q)/(mdata*b0-b0+mdata-n*mdata+mdata*n*
    b0-mdata^2)

    drop n b0 q

} // end foreach qdelta
// save regression estimates to file

summarize alphadata0–alphadata100
save horse_racing_direct_completed, replace
```

6.6 Stata code for indirect empirical testing of predicted final market probability

```
// STATA script to test data
set matsize 500
cd D:\Paul\Masters\Data
use horse racing_collapsed, clear
// add takeout rate "tau"
// takeout rate was 15.5% before Nov 2006, and 14.5% after
gen tau=0.155 if tdate_e < date("1/11/2006","dmy")
replace tau=0.145 if tdate_e >= date("1/11/2006","dmy")
// generate empirical final market price
gen mdata = (inwinamt+outwinamt)/(sbets)
// flag hold out sample
set seed 85559683
gen holdout = (uniform())>0.8

foreach qdelta of numlist 0(0.05)0.2 {
    eststo clear
    foreach alpha of numlist -0.5(0.02)0.5 {

        // # Generate parameters (N, b0, tau, q) and theoretical values (
        // b1t, mt — postfixed w "t" for theory)
        // N is the number of insiders — since CAPITAL N is reserved by
        // stata we use small n
        gen n = insider
        // b0 is the ratio of winning horse bets to all bets for
        // outsiders only
        gen b0 = (outwinamt)/(outwinamt+outloseamt)
        // Use critical_q formula from maple file "Strategic with Risk —
        // Non-core"
        // Adds qdelta to q in order to ensure that q > critical_q as
        // required
        gen q = min(1,b0+'qdelta')
        // bstar is calculated using the formula in maple file "Strategic
        // with Risk — Formulae"
        // replace "b[0]" with "b0", "N" with "n" and "alpha" with "'
        // alpha'" — rest should work
        gen K = (b0+2*b0*n+n^2*b0-1-n^2+2*n)*(1-tau)
        gen A = 1/(2*n^2*(-'alpha'*(1-tau)*q^2+(1-q)^(1/2)*(1-tau)*q
            ^((3/2)+'alpha'*(1-tau)*q-q^(1/2)*(1-q)^(1/2)))
        gen BB = ((1-b0)*(1-tau)*q^(3/2)*(1-q)*((-1+'alpha'^2)*K*q^(3/2)
            -2*(1-q)^(1/2)*'alpha'*K*q+(-K*'alpha'^2+4*b0*n)*q^(1/2)+4*'
            alpha'*(1-q)^(1/2)*b0*n))^(1/2)
        gen C = (1-tau)*q*('alpha'*(-1+q)-q^(1/2)*(1-q)^(1/2))*(-1+b0+b0*
            n+n)+2*q^(1/2)*(1-q)^(1/2)*b0*n
    }
}
```

```

// optimal insider betting amount = bstar = bstar2
gen bstar = A*(C-BB)
// generate a labels to save results
local alphalabel = ('alpha')*50+25
local qdeltalabel = ('qdelta')*20
// theoretical final market price
gen mt = (b0+n*bstar)/(1+n*bstar)
// regress empirical on theoretical for non-holdout sample
reg mdata mt if !holdout, noconst
predict pred_m_ 'alphalabel' _'qdeltalabel'
eststo alpha'alphalabel' // record regression estimates
// drop old variables at start to be ready for new loop
drop n b0 q K A BB C bstar mt

} // end foreach alpha
// save regression estimates to file
estout using q'qdelta'.txt, cells(b t ci_l ci_u) stats(r2_a N)
replace
} // end foreach qdelta

save horse_racing_completed, replace

// out of sample regressions
eststo clear
foreach optimalreg in 39_0 32_1 26_2 20_3 15_4 {
    reg mdata pred_m_'optimalreg' if holdout, noconst
    eststo osalpha'optimalreg'
}
estout using outofsample.txt, cells(b t ci_l ci_u) stats(r2_a N)
replace

```