

# **Market Efficiency in Person-to-Person Betting**

**August 2005**

**Michael A. Smith**

Senior Lecturer in Economics  
Canterbury Christ Church University  
North Holmes Road,  
Canterbury CT2 8DN  
United Kingdom  
Tel: +44 1227 76 7700  
Fax: +44 1227 47 0442  
Email: Mas34@Canterbury.ac.uk

**David Paton\***

Professor of Industrial Economics  
Nottingham University Business School  
Wollaton Road  
Nottingham NG8 1BB  
United Kingdom  
Tel: +44 115 846 6601  
Fax: +44 115 846 6667  
Email: David.Paton@Nottingham.ac.uk

and

**Leighton Vaughan Williams**

Professor of Economics and Finance  
Nottingham Business School  
Nottingham Trent University  
Burton Street  
Nottingham NG1 4BU  
United Kingdom  
Tel: +44 115 848 6150  
Email: Leighton.Vaughan-Williams@ntu.ac.uk

\* corresponding author

## **Acknowledgements**

We would like to thank an anonymous referee, participants at the 2005 International Equine Industry Program Academic Conference in Louisville, at the 2004 Western Economic Association Conference in Vancouver, at the 2003 Money, Investment and Risk Conference in Nottingham, and at a staff seminar at Nottingham Trent University for many helpful comments and suggestions.

# **Market Efficiency in Person-to-Person Betting**

## **Abstract**

Established gambling operators have argued that person-to-person wagering on Internet ‘betting exchanges’ represents unfair competition. In this paper we suggest that, in fact, betting exchanges have brought about significant efficiency gains by lowering transaction costs for consumers. We test this hypothesis using matched data on UK horse racing from betting exchanges and from traditional betting media. In contrast to traditional betting media, we find that betting exchanges exhibit both weak and strong form market efficiency. Further, we find evidence that an information based model explains the well documented favourite-longshot bias more convincingly than traditional explanations based on risk preferences.

**Keywords:** betting exchanges, market efficiency, favourite-longshot bias.

**JEL Classification:** D82, G12, G14.

# Market Efficiency in Person-to-Person Betting

## 1. Introduction

The growth in importance of person-to-person wagers on the Internet ('betting exchanges') represents an interesting phenomenon for researchers studying the efficiency of financial markets. Having been introduced as recently as 2000, Internet betting exchanges, which give bettors the opportunity to bet directly with each other, have grown rapidly in terms of turnover. Betfair currently claim to match about 500,000 bets per day (<http://www.betfairpromo.com/1soccer/index.asp?rfr=235&sid=35>).

Existing gambling operators have lobbied strongly for tougher regulation of betting exchanges on the grounds that they permit traders on the exchanges to act as bookmakers without having to register and pay tax as such. Because of their much lower margins, the current betting tax structure, which is levied on margins, also benefits the exchanges disproportionately, it is argued, compared with traditional bookmakers.

An alternative perspective is that betting exchanges represent an innovation that has improved information flows to consumers and lowered barriers to entry for producers. We might expect that, in this environment, the implied reduction in transaction costs would lead to an increase in both productive and allocative efficiency relative to other wagering markets.

There is, in fact, a long and established literature examining the efficiency of betting markets, much of it focusing on the existence of weak form inefficiencies such as the 'favourite-longshot bias' whereby bets placed at shorter odds ('favourites') tend to yield a higher expected return than bets at longer odds ('longshots') - see Sauer (1998), Vaughan Williams (1999) for surveys of the literature.

Hurley and McDonough (1995) offer a theoretical model of the favourite-longshot bias based on the existence of positive transaction and information costs faced by bettors.

Sobel and Raines (2003) go further, seeking to test an information model empirically against

the risk preference alternative, using an extensive dataset of prices drawn from nearly three thousand races at two US greyhound tracks. We test this information model using new data from betting exchanges and traditional betting markets, and in so doing compare the bias in these two competing markets.

First, however, we explain (in Section 2) more fully the operation of betting exchanges. In Section 3, we consider explicitly the importance of transaction costs and information in market efficiency. We introduce our data and empirical methodology in Sections 4 and 5 respectively and present our empirical results in Section 6. We make some concluding remarks in the final section.

## **2. Internet Betting Exchanges**

Betting exchanges exist to match people who want to bet on a future outcome at a given price with others who are willing to offer that price. The person who bets on the event happening at a given price is the backer. The person who offers the price is known as the layer.

The advantage of this form of betting for the bettor is that, by allowing anyone with access to a betting exchange to offer or lay odds, it serves to reduce margins in the odds compared to the best odds on offer with traditional bookmakers. Exchanges allow clients to act as a backer or layer at will, and indeed to back and lay the same event at different times during the course of the market.

The way in which this operates is that the major betting exchanges present clients with the three best odds and stakes which other members of the exchange are offering or asking for. For example, for England to beat Australia at cricket the best odds on offer to those wishing to back England might be 3 to 1, to a maximum stake of £80, 2.5 to 1 to a further stake of £100 and 2 to 1 to a further stake of £500. This means that potential backers can stake up to a maximum of £80 on England to beat Australia at odds of 3 to 1, a further

£100 at 2.5 to 1 and a further £500 at 2 to 1. These odds, and the staking levels available, may have been offered by one or more other clients who believe that the true odds were longer than they offered.

An alternative option available to potential backers is to enter the odds at which they would be willing to place a bet, together with the stake they are willing to wager at that odds level. This request (say £50 at 4 to 1) may be accommodated by a layer or layers at any time until the event takes place.

The margin between the best odds on offer and the best odds sought tends to narrow as more clients offer and lay bets, so that in popular markets the real margin against the bettor (or layer) tends towards the commission levied (normally on winning bets) by the exchange.

### **3. A Cost-based Model of the Favourite-longshot Bias**

Early models of the favourite-longshot bias suggested that bettors are 'risk loving' (see, for example, Rosett 1965; Weitzman 1965). More recent studies, however, have attributed the bias to the existence of transactions and information costs. In particular, Hurley and McDonough (1995) suggest that the extent of any bias may be positively related to the transaction costs faced by bettors as a class in acquiring information concerning the true probabilities of runners, as well as by the magnitude of the 'take' or deductions, i.e. the profit margin or administrative costs of market operators.

In their model, Hurley and McDonough consider the case of risk-neutral bettors occupying a parimutuel betting market. In the absence of transactions or information costs, bettors are able to calculate the 'true' probabilities of each outcome, so that the subjective probabilities about each potential outcome (as contained in the odds) will equal the objective probabilities about each outcome, i.e. no bias. The presence of positive transactions and information costs, however, causes the subjective probability that the 'favourite' (defined as

the horse with the highest objective win probability) wins to diverge systematically from the objective probability. In the limit, bettors will be totally uninformed and so will bet equal amounts on each outcome, regardless of the objective probabilities, i.e. they will bet relatively too much on the options with a low probability of success and too little on those with a high probability of success. This is the classic favourite-longshot bias and the bias will exist insofar as transactions and information costs discourage bettors from becoming totally informed. It follows also that the bias would increase as these costs increase.

Although the Hurley and McDonough proposition was not supported by their experimental evidence, there is, in fact, an emerging body of empirical evidence gathered from horse race markets that is consistent with their hypothesis. For example, Vaughan Williams & Paton (1997) find that the favourite-longshot bias is more pronounced in low-grade races than in high class races. This finding is consistent with a reasonable assumption that the cost of acquiring information relevant to the race outcomes is higher for low-grade races than high class contests, because there is likely to be less public and media scrutiny of low grade runners.

Sobel and Raines (2003) offer further supporting evidence for an information-based explanation, identifying a lower bias in high volume betting markets, assumed to be better informed, than low volume markets, assumed to be proportionately more heavily populated by casual bettors. The starting point for building their information model is to show that in the absence of any information regarding race outcomes, the expected proportion of public bets made on each runner in a pari-mutuel market will be  $1/N$ , where  $N$  is the number of race entrants. This represents the limiting case of extreme bias. To the extent that the betting public acquire race specific information to inform their assessment of the true chances of individual runners, the actual degree of bias will depart from this limiting case and the proportions bet will approach more closely the distribution of objective probabilities. The

degree of bias is therefore largely a function of the amount of information available to bettors and the number of runners in the race. Using a substantial dataset of greyhound racing pari-mutuel prices, Sobel and Raines measure the bias in a subset of races associated with a high proportion of ‘serious’ bettors (weekday races), and compare with a subset associated with a high proportion of ‘casual’ bettors (weekends), in order to test the model. They find evidence of a conventional favourite-longshot bias associated with a high proportion of casual bettors, and of an opposite favourite-longshot bias (due to over-reaction to information) in the presence of a high proportion of ‘serious’ bettors, substantiating their information model.

Sobel and Raines also demonstrate a clear relationship between the degree of bias and the number of race entrants, and show that this finding is at odds with the predictions of risk preference models. They usefully specify testable models of risk and information explanations of bias as functional relationships between subjective and objective probabilities, enabling empirical arbitration between the two models. Their tests of the models, including controls for race grade, time of day, and bet complexity, suggest that the information model explains the markets they examine better than the risk preference alternative.

In this paper we seek to build upon the work of Hurley and McDonough and Sobel and Raines in order to examine the influence of transactions and information costs on the existence of the favourite-longshot bias. We use Shin’s methodology (Shin 1991, 1992, 1993) to calculate the bias for a sample of races, firstly in respect of prices from traditional bookmaking markets, and secondly in respect of betting exchange prices for the same races. The Shin approach has been employed in other recent papers studying the structural and behavioural characteristics of betting markets (see, for example, Cain, Peel and Law, 2001a, 2001b, 2003). Shin developed a systematic theoretical model that accounts for the bias by reference to insider activity, specifying an informational hierarchy comprising of insiders

(who are assumed to have certain knowledge of race outcomes), price setters (bookmakers, who are monopoly price setters) and outsiders (relatively casual recreational bettors). Shin models the behaviour of bookmakers, arguing that they engineer the favourite-longshot bias to pass the cost of losses due to insider activity on to outsiders (for a concise, formal summary of the Shin model, see Law and Peel 2002, appendix).

The level of commission levied across markets in the betting exchange that we consider (Betfair, the world's largest betting exchange) is normally set at a maximum of 5% of winnings. This is considerably less than the notional profit margin of bookmakers implied in the 'over-round', i.e. the sum of probabilities implied in the odds minus 1, which averages at 25.63% in our 700 race sample (based on mean bookmaker prices). If the costs-based explanation of the bias is correct, therefore, we should expect the favourite-longshot bias to be more pronounced in the bookmaker data.

In addition we test the effects of information costs associated with race class, adopting a procedure similar to that used by Vaughan Williams and Paton (1997), whereby races are classified by betting volume/grade as a proxy for information intensity. In particular, we measure the favourite-longshot bias for each information class within the exchange data, and separately within the bookmaker data. The null hypothesis is that the degree of bias across information classes is equal. Finally, we seek to arbitrate between information and risk preference explanations of bias by employing the Sobel and Raines methodology, testing functional relationships between subjective and objective probabilities associated with these alternative models against our data.

#### **4. Data used in this study**

The first set of prices collected were those offered by bookmakers. Unlike pari-mutuel prices, these odds are fixed, regardless of subsequent fluctuations in the market; the only

exception to this is when there are withdrawals of runners in the race, in which case a differential reduction is applied, based on the probability of success of the withdrawn runner or runners.

Bookmakers' prices were gathered for 799 horse races run in the UK during 2002. Sample races were drawn from the second half of the 2001-02 National Hunt season, the 2002 Flat season, and the beginning of the 2002-03 National Hunt season. In order to minimise liquidity issues, sampling was restricted to Saturdays and other days where overall betting turnover was likely to be vigorous. One advantage of sampling over the full calendar year 2002 is that our data should not suffer in aggregate from seasonal bias. Prices were taken from the Internet site of the *Racing Post*, the major daily publication dealing with horse racing and gambling in the U.K. Taking prices from the Internet site allows for a real-time comparison with betting exchange data, also acquired online.

Our aim was to establish as complete a set of prices as possible as early in the market as possible. To ensure that prices were not merely nominal, a trial was conducted whereby bets were placed to establish that the prices stated could be obtained. Actual bets were small (ranging from £5 to £20), but enquiries were also made with individual bookmakers as to whether much larger bets would be accepted. There was evidence of some limits to bet size set by bookmakers on occasions, but not frequently enough to raise concerns about the integrity of prices in general or to suggest a lack of liquidity that might require qualification of the results presented here.

We calculated the mean of bookmakers' prices for each runner in each race, to enable us to develop a measure of bias that could be compared directly with that of previous studies. In addition, we identified the most favourable price for each horse (the outlier), as this is an important competitive benchmark against which betting exchange prices are compared by bettors.

The bookmaker data were matched with corresponding betting exchange prices, collected at the same time each day, 10.30 a.m. Bet limits on betting exchanges are explicit, and evidenced by the amounts layers state that they are prepared to accept in bets on individual runners (as outlined above). Where bet limits were small, the prices offered were ignored, and races where overall betting volume was trivially low were excluded from the sample of races, on the grounds that the market did not have sufficient liquidity to warrant treating such observations as representative<sup>1</sup>. A minimum acceptable aggregate turnover threshold (£2000 per race, by 10.30 am) was applied as a filter to the races in the sample in respect of Betfair prices; races where this turnover threshold was not met were screened out of the analysis. After exclusion of races on grounds of turnover or recent withdrawals, we were left with exactly 700 races for the analysis.

As a measure of information intensity, we divided our information classes according to betting volume, which is highly correlated with other relevant qualitative criteria such as racecourse grade, information on runners, media coverage, and prize money. The classification, including typical associated qualitative race criteria, was as follows:

**Class 1: Races with low betting volume.** These are mostly at low grade racetracks and for small monetary prizes; often unexposed or unknown form for a number of runners; minimal media coverage.

**Class 2: Races with moderate betting volume.** These usually attract middle ability horses; form is more exposed than Class 1 races; average prize money.

---

<sup>1</sup> To avoid sample bias, we were careful to exclude only races where turnover was low with *both* Betfair *and* bookmakers, as evidenced explicitly on the Betfair website, and by inference from bet sizes in trade press results sections in the case of bookmakers, and enquiries made with bookmakers as to bet limits.

**Class 3: Races with higher than average betting volume.** These are competitive races with a high degree of betting interest, generated by characteristics of the race or its contenders likely to attract public interest and enhanced media coverage; higher than average reported betting volumes in the press.

**Class 4: Races with very high betting volume.** High profile and top class races; high profile contending horses; high degree of competition and media interest, speculation on runners often extending weeks before the contest.

These classes are not official race categories. They are primarily distinguished by betting turnover, as a proxy for the degree of media publicity and other qualitative aspects outlined in the class descriptions above. Official industry race classes were not used for our purpose, as it is far from clear that these are closely correlated to the amount of public information about runners. Many horses in high-class two-year old races, for example, are relatively unexposed to prior public scrutiny. Table 1 summarises the distribution of races in our sample between the four information categories.

In order to distinguish between information and risk preference explanations of the observed market structures, we needed a further classification of prices to facilitate estimation of the functional relationship between subjective and objective probabilities in our data. We employed a method of classification traceable to Weitzman (1965), whereby normalised odds probabilities are categorised according to a measure of the monetary return to a nominal winner at given odds to a unit bet, including stake. This largely solved the problem of specifying classes having an insignificant number of runners, especially in the shortest odds categories. Table 2 summarises the normalised odds probabilities of horses in our sample, categorised by Weitzman category.

## 5. Empirical Methodology

Our measure of the favourite-longshot bias is derived from the model constructed by Shin (1993, pg.1148), which explains the favourite-longshot bias as a result of bookmaker behaviour in the face of insider trading. Shin shows that, in equilibrium, the sum of price probabilities offered by bookmakers will exceed 1, such that:

$$D_i = z(n-1)_i + \sum a_{k,n;k} \text{Var} (p)_i + \sum b_{k,n;k} [\text{Var} (p)]_i^2 \quad (1)$$

where  $D$  is the sum of prices in race  $i$ , expressed as probabilities minus one;  $n$  is the number of runners; and  $\text{Var} (p)$  is the “variance” of price probabilities for the runners in the race<sup>2</sup>.

The coefficient of  $n-1$ ,  $z$ , is the measure of insider trading, and the higher the value of  $z$ , the greater the degree of bias. For his sample of 178 races in the early 1990s, Shin estimated  $z$  to be 0.025, i.e. 2.5% of betting turnover could be attributed to insiders.

Shin uses an iterative ordinary least squares method to estimate  $z$  in his sample of races, beginning the process with an initial estimate of  $z$  from the observed variance of prices within the races, used as a proxy for variance of probabilities. The value of  $z$  is re-estimated by using this initial value, and the iterative process is repeated until convergence is achieved. This estimating procedure, replicated in the later studies by Vaughan Williams and Paton (1997), and by Cain, Law and Peel (2001a), is also adopted here. In each case, we restrict the polynomial in equation (1) to  $k = 2$ . In accordance with Shin (1992) and Vaughan Williams and Paton (1997), using higher values of  $k$  has very little impact on our results.

Shin’s  $z$  offers a robust method of analysing the degree of bias in specific races and has the further useful property that it does not require an estimate of true probabilities based on results, permitting application to a much smaller set of races.

---

<sup>2</sup> Shin does not use the term variance in its usual sense; rather,  $\text{Var} (p)$  is a measure of distance of vector  $p$  from the vector  $1/n$ .

We wish to obtain estimates of the value of  $z$  in equation 1 for the three different prices. The prices all pertain to the same set of races. Given this, we exploit potential correlations in the residual terms by estimating equation 1 for each set of prices using the seemingly unrelated regression (SUR) technique. SUR enables us to achieve gains in efficiency in the presence of correlations in error term across the three models. A further advantage of the SUR approach is that it enables us to test directly the equality of coefficients (in our case, the estimate of  $z$ ) across the three equations.

Recall that the cost based model predicts that the level of bias will be systematically higher in cases when transaction costs are higher. In terms of the Sobel and Raines model this is because fewer ‘serious’ bettors will wish to be involved when transaction costs are high. The first empirical consequence of this proposition is that the estimated value of  $z$  for the outlier bookmaker prices should be lower than that for the mean bookmaker prices.<sup>3</sup>. Secondly, the estimated value of  $z$  for the Betfair prices should be lower than for either set of bookmaker prices. To the extent that our classification of races into different classes proxies for the costs to consumers of obtaining information about form, a third empirical consequence is that the estimated value of  $z$  should be lower, the higher is the class of race.

In addition to testing the transaction cost/information based model we wish to consider the adequacy of the alternative risk preference explanation of the favourite-longshot bias in relation to the same data, employing the methodology of Sobel and Raines.

Building on Rosett (1965), Sobel and Raines specify the risk model as:

$$\log(\rho_i) = \alpha + \beta \log(\pi_i) \quad (2)$$

---

<sup>3</sup> The outlier prices may also offer ‘quasi-arbitrage’ trading opportunities for the bettor, i.e. opportunities to trade at prices better than the objective probabilities, assuming that the mean of the prices on offer is a good reflection of the true chances of the runners (see Paton and Vaughan Williams, 2005; Smith, Paton and Vaughan Williams, 2005).

where  $\rho_i$  is the subjective probability of horse  $i$  winning, and  $\pi_i$  is the corresponding objective probability. Further, they derive an information model using a process of Bayesian updating, specified as:

$$\rho_i = \alpha + \beta\pi_i \quad (3)$$

In both cases  $\beta$  is a measure of the favourite-longshot bias, with  $0 < \beta < 1$  indicating over betting of longshots relative to runners at short odds, and  $\beta > 1$  indicating an opposite favourite-longshot bias.

In estimating these models for the data in our sample, we adopt normalised mean bookmaker prices for each horse as the observed values of  $\rho_i$ . We exploit our finding of very little bias in the betting exchanges (see next section) to estimate objective probabilities, utilising normalised betting exchange prices for runners as a proxy for  $\pi_i$ . Mean values for  $\rho_i$  and  $\pi_i$  are established for the Weitzman categories, as summarised in Table 2, and subsequently used in a weighted least squares regression to estimate equations (2) and (3).

A further test performed by Sobel and Raines is to establish the point at which the subjective and objective probability lines cross; they demonstrate that for the information model this crossing point is at  $\pi = 1/N$ , whereas for the risk preference model the corresponding crossing point is at  $\pi = \lambda^{1/1-\beta}$ , suggesting that the degree of bias is independent of the number of entrants in a race.

Should the estimation of equation (3) fit the data in our sample better than a similar estimation of equation (2), this would provide empirical support for the proposition that the information model better explains the favourite-longshot bias than the risk preference model.

Further, if the crossing point of subjective and probability lines corresponding to our data is close to  $1/N$ , this will also provide empirical support for the information model.

## 6. Results and discussion

In respect of the Shin  $z$  estimates, our results are reported in Table 3 to 8. In all cases, the Breusch-Pagan test rejects the null that the residuals in the three equations are independent and this provides support for our use of the SUR methodology.

We report the SUR results of equation 1 with the three different sets of prices for the whole sample in Table 3. Recall that the bias (Shin's  $z$  value) is given by the estimated coefficient on the variable  $n-1$ . For the mean fixed odds data, the bias is estimated to be 2.17%, a figure broadly comparable with the estimates derived from starting prices by Shin (1993) and Vaughan Williams and Paton (1997), using a similar methodology<sup>4</sup>. When we use the outlier fixed odds data, the bias reduces to 1.19%. These values are significantly higher than the corresponding figure for the betting exchange data, where the bias is just 0.9%, offering support for the Hurley and McDonough proposition that the degree of favourite-longshot bias will be more pronounced when trading costs are higher. In terms of the Sobel and Raines model, a lower conventional bias in the exchanges relative to traditional betting markets is consistent with a higher proportion of 'serious' bettors on the exchanges than with bookmakers<sup>5</sup>. The formal tests that pairs of these coefficients are equal (reported in Table 8), confirm that the estimated bias in the exchange data is significantly lower than both the mean and outlier fixed odds data.

---

<sup>4</sup> An alternative approach to that of Shin in which the favourite-longshot bias is estimated directly is suggested in Schnytzer and Shilony (1995). Using this approach led to a very similar ordering of the bias across each of the three price formats to that reported here.

<sup>5</sup> We have no empirical data to suggest that the proportion of casual bettors is greater with bookmakers than the exchanges but recent evidence submitted to a UK parliamentary committee considering gambling legislation suggests a significant degree of non- recreational trading on the betting exchanges.

We now consider the degree of bias associated with the four information classes. As discussed above, an important implication of the Hurley and McDonough and the Sobel and Raines hypotheses is that there will be a positive relationship between costs involved in acquiring race specific information, and the degree of favourite-longshot bias. We therefore expect the values of  $z$  to decrease as we progress from the subset of races associated with the least public information (Class 1) through to that associated with most public information (Class 4). This proposition is borne out by the  $z$  values across information sets for all three price formats (reported in Tables 4-7 and in Figure 1). Unequivocal support for this element of the hypothesis would require that  $z$  decreases monotonically from Class 1 to Class 4. Whereas only for the betting exchange data is this strict condition met, the overall trends in  $z$  for the bookmaker data are also clearly decreasing. The tests of equality of coefficients across the three sets of prices (summarised in Table 8) give confidence in the systematic nature of these overall structural trends. The only exception is the Betfair/outlier test of equality at Class 1. As Class 1 in our information hierarchy coincides with lower turnover races, it is likely that this anomaly is caused by a reduced degree of price competition between layers on the exchange, with a number of prices filtered out by our turnover rule; this lack of competitive pricing would not be as apparent in respect of outliers as bookmakers may feel obliged to offer a full set of competitive prices, despite low public interest, to maintain credibility. In these circumstances one could expect to encounter a liquidity limit to further reductions in the degree of bias in the betting exchange. Nonetheless, exchange prices remain competitive with outliers, emphasising the importance of the latter as a benchmark for exchange layers.

The evidence is, therefore, consistent overall with an information and cost-based modelling of the favourite-longshot bias.

Intuitively, it may seem paradoxical that the extent of insider trading should be lowest in the exchanges, where traders possessing inside information have arguably the greatest opportunity to exploit this knowledge by, for example, laying high odds against non-triers winning. Our results are not inconsistent with this intuitive reasoning, although they do suggest that insider trading is not as commonplace on exchanges as sometimes portrayed in the media. To demonstrate this consistency, if we consider the standard errors of the estimates of Shin  $z$  presented in Tables 4-7, they are without exception highest for the betting exchange values, despite lower estimates of  $z$ . This implies that although the degree of bias is overall less in the structure of exchange prices, the impact of specific items of insider information is likely to be more evident and exaggerated in particular races than is the case for the bookmaker data. In other words, our results do not deny the existence of isolated cases of insider activity in the exchanges – rather, they suggest that such activity is not widespread.

A further implication of the interpretation of the favourite-longshot bias suggested here is that it would be unwise to attribute the observed bias in bookmaker prices solely to bookmaker insurance against asymmetric information, as in the Shin model. The transaction cost and information models explain the observed pattern of bookmaker prices equally well. A question arises as to whether it is legitimate to use a measure of bias (Shin's  $z$ ) explicitly derived from a model of bookmaker behaviour, to propose a model not based on this initial premise. Shin models bookmaker competition, whereas the person-to-person exchange consists of individuals who do not need to maintain a credible market structure embracing all runners. Aside from the degree of bias and level of transaction costs, however, the competitive structure of exchange markets resembles that of Shin's bookmaking market quite closely. For example, the dynamics of the market are such that, for individual runners, exchange layers have to offer competitive prices to attract bettors, with the sum of price

probabilities usually exceeding one by only a few percent. On the other hand, the occasions when the sum of probabilities falls below one are extremely rare (these characteristics are confirmed by observation of our sample races). We feel justified, therefore, in applying the Shin  $z$  measure in this non-bookmaker betting medium.

In judging whether an information model explains the favourite-longshot bias better than a risk preference model, we estimated the Sobel and Raines specifications expressed in equations (2) and (3) above. The coefficients  $\alpha$  and  $\beta$  corresponding to each equation, estimated from observed values of  $\rho$  (based on bookmaker mean prices), and  $\pi$  (exchange prices employed as a proxy), are summarised in Table 9. Figure 2 shows fitted and actual values of  $\rho$  plotted against  $\pi$  for the risk model, whilst Figure 3 plots the corresponding values for the information model. The information model appears to fit the data much more closely than the risk alternative. Figure 3, in fact, suggests an almost perfect fit. This empirical finding corresponds very closely to the Sobel and Raines result, supporting their conclusion that an information based explanation of the favourite-longshot bias is more robust than one based on risk preference.

Finally, we consider the crossing points of the subjective and objective probability functions (in Figures 2 and 3 the latter would be represented by a 45% line through the origin). Recall that the information model predicts that the crossing point will be at  $\pi = 1/N$ . As the mean number of race entrants in our sample is 11.8, we therefore expect the lines to cross at a value of  $1/11.8 = 0.0847$ . Substituting  $\rho = \pi$  into equation (3) gives a resulting crossing point at  $\pi = \alpha / (1 - \beta)$ , which from Table 9 yields a value of 0.0849. This is virtually identical to  $1/N$ , offering further empirical support for the information model.

## 7. Conclusions

In this study we have shown that the established favourite-longshot bias is demonstrably lower in person-to-person (exchange) betting than in traditional betting markets. As these exchange betting markets are characterised by relatively low transactions costs, our findings are consistent with models in which such costs can help to explain the favourite-longshot bias. We further find that, in both exchange and traditional betting markets, the level of bias is lower the greater the amount of public information that is available to traders. Additional empirical support for an information based model is found by employing an alternative methodology which enables us to arbitrate between information based and risk preference models of the favourite-longshot bias in relation to our data. Our results suggest that an information model explains the favourite-longshot bias better than a risk preference model.

## References

- Cain, M., Law, D. and Peel, D. A. (2001a). The Relationship between two Indicators of Insider Trading in Racetrack Betting. *Economica*, 68, 97-104.
- Cain, M., Law, D., and Peel, D.A. (2001b). The Incidence of Insider Trading in Betting Markets and the Gabriel and Marsden Anomaly. *The Manchester School*, 69 (2) 197-207.
- Cain, M., Law, D. and Peel, D.A. (2003). The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets. *Bulletin of Economic Research*, 55, 263-73.
- Hurley, W. and McDonough, L. (1995). A Note on the Hayek Hypothesis and the Favourite-Longshot Bias in Parimutuel Betting. *American Economic Review*, 85 (4), 949-955.
- Law, D. and Peel, D.A (2002). Insider Trading, Herding Behaviour and Market Plungers in the British Horse-race Betting Market. *Economica*, 69, 327-338.

- Paton, D and Vaughan Williams, L. (2005). Forecasting Outcomes in Spread Betting Markets: Can Bettors Use 'Quarbs' to Beat the Book? *Journal of Forecasting*, 24 (2), 139-154
- Rosett, R. N. (1965). Gambling and Rationality. *Journal of Political Economy*, 73, 595-607.
- Sauer, R. D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 36, 2021-2064.
- Schnytzer, A. and Shilony, Y. (1995). Inside Information in a Betting Market. *Economic Journal*, 105 (July), 963-971.
- Shin, H. S. (1991). Optimal Betting Odds against Insider Traders. *Economic Journal*, 101, 1179-1185.
- Shin, H. S. (1992). Prices of State Contingent Claims with Insider Traders, and the Favourite-longshot bias. *Economic Journal*, 102, 426-435.
- Shin, H. S. (1993). Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims. *Economic Journal*, 103, 1141-1153.
- Smith, M.A., Paton, D. and Vaughan Williams, L. (2005). An Assessment of Quasi-Arbitrage Opportunities in Two Fixed Odds Horse Race Betting Markets. *Information Efficiency in Financial and Betting Markets*, ed. L. Vaughan Williams, Cambridge: Cambridge University Press, 159-171.
- Sobel, R. S. and Raines, S. T. (2003). An Examination of the Empirical Derivatives of the Favorite-Longshot Bias in Racetrack Betting. *Applied Economics*, 35, 371-385.
- Vaughan Williams, L (1999). Information Efficiency in Betting Markets. *Bulletin of Economic Research*, 51 (1), 1-30.
- Vaughan Williams, L. and Paton, D (1997). Why is there a Favourite-Longshot Bias in British Racetrack Betting Markets? *Economic Journal*, 107, 150-158.
- Weitzman, M. (1965). Utility Analysis and Group Behaviour: An Empirical Study. *Journal of Political Economy*, 73, 18-26.

**Table 1: Distribution of races by information class**

Information class	Races
1	171
2	265
3	137
4	127
<b>Total</b>	<b>700</b>

**Table 2: Normalised odds probabilities (Weitzman classes)**

Return <sup>(i)</sup>	No. of horses <sup>(ii)</sup>	Subjective probability <sup>(iii)</sup>	Objective probability <sup>(iii)</sup>
1	13	0.5512	0.5933
2	49	0.4481	0.4666
3	134	0.3272	0.3391
4	261	0.2492	0.2571
5	305	0.1991	0.2064
6	344	0.1672	0.1732
7	364	0.1434	0.1481
8	373	0.1253	0.1295
9	356	0.1113	0.1129
10	376	0.0999	0.1030
11	321	0.0911	0.0923
12	306	0.0835	0.0848
13	421	0.0756	0.0751
15	496	0.0668	0.0668
17	667	0.0574	0.0563
21	856	0.0476	0.0452
26	721	0.0380	0.0340
34	629	0.0300	0.0256
41	429	0.0241	0.0184
51	361	0.0196	0.0135
67	233	0.0146	0.0095
101	124	0.0102	0.0060
151 and over	119	0.0054	0.0036
<b>N</b>	<b>8258</b>		
Average no. of race entrants (N)	11.8		
<b>1/N</b>	<b>0.0848</b>		

**Notes:**

(i) Return to a unit stake bet on a nominal winner, inclusive of stake.

(ii) Categorised according to mean bookmaker odds.

(iii) Mean value for all horses in the category.

**Table 3: SUR Estimates of the Shin over-round function D: all races**

	<b>Mean</b>	<b>Outlier</b>	<b>Betfair</b>
<b>n-1</b>	0.0217*** (0.0003)	0.0119*** (0.0003)	0.0090*** (0.0005)
<b>V</b>	-26.856*** (2.189)	-18.042*** (2.170)	-11.366*** (2.718)
<b>Nv</b>	8.095*** (0.4282)	4.662*** (0.4300)	3.758*** (0.5233)
<b>n<sup>2</sup>v</b>	-0.4276*** (0.0217)	-0.2380*** (0.0221)	-0.2625*** (0.0264)
<b>V<sup>2</sup></b>	1111.340*** (115.761)	622.428*** (109.765)	632.404*** (138.865)
<b>Nv<sup>2</sup></b>	-310.588*** (34.453)	-160.578*** (32.554)	-196.288*** (41.824)
<b>n<sup>2</sup>v<sup>2</sup></b>	16.135*** (2.584)	8.256*** (2.442)	13.017*** (3.048)
<b>R<sup>2</sup></b>	0.9698	0.9042	0.6049
<b>N</b>	700	700	700
<b>Independence</b>		760.52***	

**Notes:**

(i) Estimates are from the final stage of the iterative process as described in the text.

(ii) The dependent variable is D = sum of price probabilities minus one (equation 1).

(iii) Figures in brackets are standard errors. \*\*\* indicates significance at the 1% level; \*\* at the 5% level; \* at the 10% level.

(iv) Independence indicates the Breusch-Pagan test that the equations are independent. The test statistic is distributed as  $\chi^2(3)$ .

**Table 4: SUR Estimates of the Shin function D: Class 1 races**

	Mean	Outlier	Betfair
<b>n-1</b>	0.0258*** (0.0007)	0.0132*** (0.0008)	0.0182*** (0.0016)
<b>V</b>	-4.709 (5.458)	-6.809 (6.412)	14.978 (10.953)
<b>Nv</b>	2.851** (1.118)	1.078 (1.317)	-2.086 (2.204)
<b>N<sup>2</sup>v</b>	-0.2092*** (0.0560)	0.0149 (0.0665)	-0.1087 (0.1063)
<b>V<sup>2</sup></b>	272.314 (240.570)	-47.958 (268.694)	196.135 (458.517)
<b>nv<sup>2</sup></b>	-115.416 (81.873)	71.602 (94.806)	-181.077 (166.986)
<b>N<sup>2</sup>v<sup>2</sup></b>	8.102 (7.488)	-10.122 (8.812)	24.533 (15.652)
<b>R<sup>2</sup></b>	0.9828	0.9278	0.7548
<b>N</b>	171	171	171
<b>Independence</b>		94.40***	

See Table 3, notes (i) to (iv)

**Table 5: SUR Estimates of the Shin function D: Class 2 races**

	Mean	Outlier	Betfair
<b>n-1</b>	0.0261*** (0.0006)	0.0147*** (0.0006)	0.0124*** (0.0009)
<b>V</b>	-21.045*** (3.575)	-14.684*** (3.625)	-6.920 (4.608)
<b>Nv</b>	6.981*** (0.8128)	4.115*** (0.8308)	2.611** (1.033)
<b>N<sup>2</sup>v</b>	-0.4641*** (0.0483)	-0.2767*** (0.0500)	0.2547*** (0.0600)
<b>V<sup>2</sup></b>	1020.799*** (230.061)	564.807** (232.228)	539.283* (273.593)
<b>nv<sup>2</sup></b>	-312.503*** (72.273)	-170.358** (73.450)	-181.400* (85.851)
<b>N<sup>2</sup>v<sup>2</sup></b>	20.298*** (5.677)	12.243** (5.786)	14.491* (6.635)
<b>R<sup>2</sup></b>	0.9755	0.9156	0.6907
<b>N</b>	265	265	265
<b>Independence</b>		312.86***	

See Table 3, notes (i) to (iv)

**Table 6: SUR Estimates of the Shin function D: Class 3 races**

	<b>Mean</b>	<b>Outlier</b>	<b>Betfair</b>
<b>n-1</b>	0.0186*** (0.0005)	0.0105*** (0.0006)	0.0052*** (0.0008)
<b>V</b>	-20.879*** (4.850)	-14.002*** (5.314)	-9.025 (6.013)
<b>Nv</b>	7.165*** (0.8985)	4.031*** (0.9918)	2.803*** (1.075)
<b>N<sup>2</sup>v</b>	-0.2767*** (0.0424)	-0.1670*** (0.0473)	-0.1351*** (0.0502)
<b>V<sup>2</sup></b>	1154.01*** (407.163)	755.033* (412.894)	152.903 (446.454)
<b>nv<sup>2</sup></b>	-280.511** (124.598)	-181.760 (123.729)	-21.920 (137.868)
<b>N<sup>2</sup>v<sup>2</sup></b>	2.368 (9.189)	4.443 (9.200)	-3.203 (10.561)
<b>R<sup>2</sup></b>	0.9856	0.9414	0.6891
<b>N</b>	137	137	137
<b>Independence</b>		116.76***	

See Table 3, notes (i) to (iv)

**Table 7: SUR Estimates of the Shin function D: Class 4 races**

	Mean	Outlier	Betfair
<b>n-1</b>	0.0173*** (0.0006)	0.0089*** (0.0006)	0.0038*** (0.0007)
<b>V</b>	-38.961*** (8.283)	-20.106** (8.995)	-16.812** (7.210)
<b>Nv</b>	8.981*** (1.912)	4.191*** (1.318)	3.312*** (1.029)
<b>N<sup>2</sup>v</b>	-0.3580*** (0.0447)	-0.1546*** (0.0502)	-0.1442*** (0.0396)
<b>V<sup>2</sup></b>	787.121*** (245.831)	342.393 (251.637)	454.209** (205.910)
<b>nv<sup>2</sup></b>	-154.164** (78.255)	-54.260 (78.174)	-105.455 (66.063)
<b>N<sup>2</sup>v<sup>2</sup></b>	2.995 (5.893)	0.2528 (5.855)	5.517 (4.947)
<b>R<sup>2</sup></b>	0.9725	0.8966	0.5333
<b>N</b>	127	127	127
<b>Independence</b>		104.14***	

**Notes:**

See Table 3, notes (i) to (iv)

**Table 8: Results of null hypotheses tests: equality of z values across price formats Classes 1 to 4 (information sub-sets)**

<b>H<sub>0</sub>:</b>	<b>Mean = Outlier</b>	<b>Mean = Betfair</b>	<b>Betfair = Outlier</b>
<b>All races</b>	1848.38***	960.68***	42.26***
<b>Class 1</b>	946.27***	348.26***	0.09
<b>Class 2</b>	1291,36***	715.75***	29.48***
<b>Class 3</b>	1191.89***	853.05***	79.28***
<b>Class 4</b>	1007.80***	670.06***	60.10***

**Notes:**

(i) Tests are of the null hypothesis that the value of z (i.e. the coefficient on n-1) is the same for the respective samples.

(ii) \*\*\* indicates significance at the 1% level.

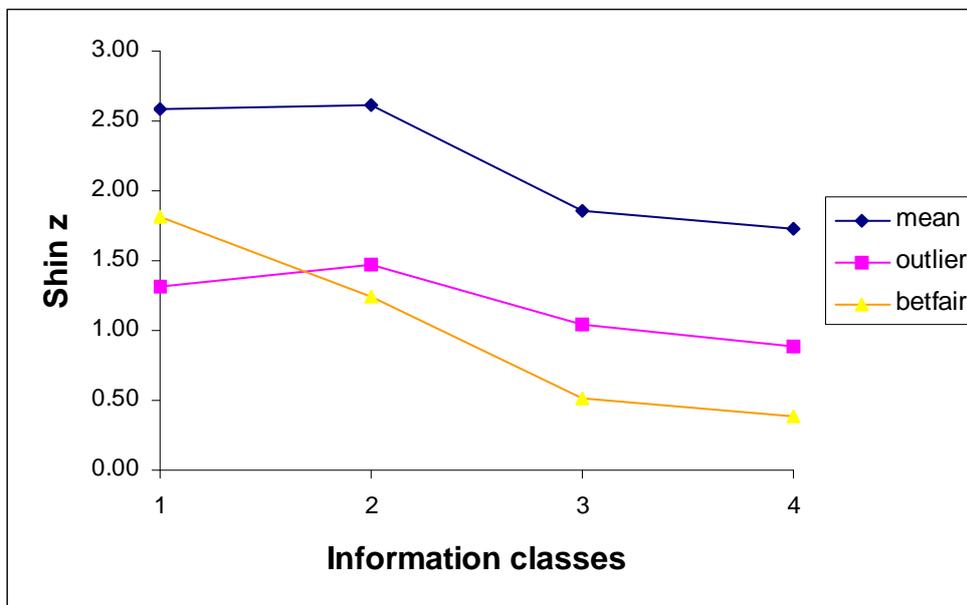
**Table 9: Coefficients for the Risk Preference and Information based models**

	<b>Risk model</b>	<b>Information model</b>
$\alpha$	-0.3103*** (0.0399)	0.0050*** (0.0004)
$\beta$	0.8655*** (0.0132)	0.9405*** (0.0038)
$R^2$	0.9976	0.9998
N	23	23

**Note:**

\*\*\* indicates significance at the 1% level.

**Figure 1: Shin z coefficients for bookmakers mean, bookmakers outlier, and betting exchange prices**

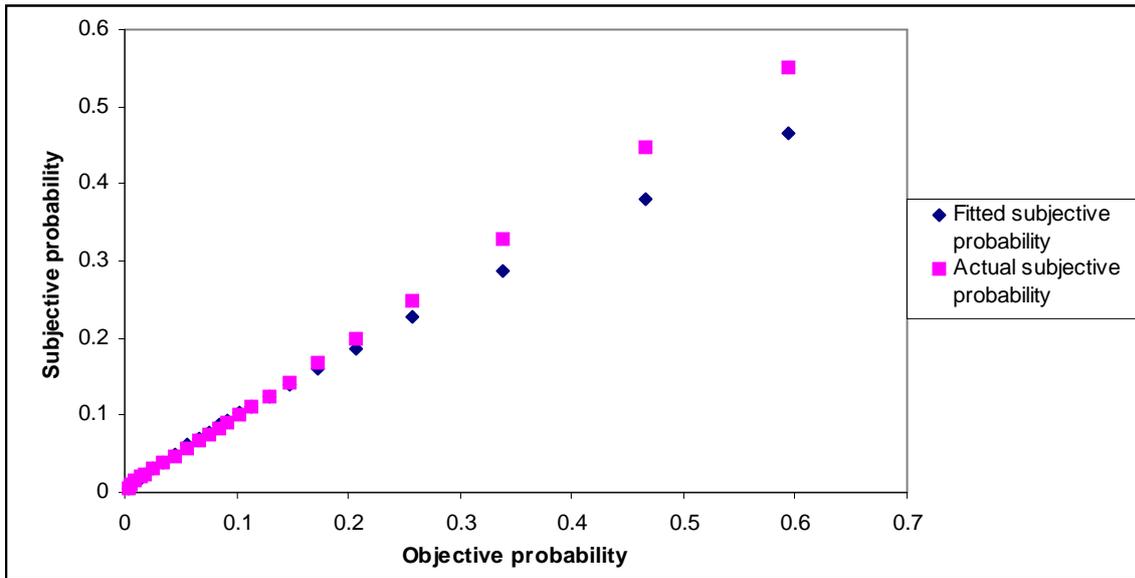


**Notes:**

(i) The y axis shows the coefficient of  $n-1$ , or Shin's  $z$ , multiplied by 100. The interpretation of this value is that it indicates the percentage of insider trading volume in the market concerned, and also acts as a direct proxy measure of the degree of bias.

(ii) Class 1 = least public information; Class 4 = most public information.

**Figure 2: Estimated relationship between subjective and objective probabilities: the risk preference model**



**Figure 3: Estimated relationship between subjective and objective probabilities: the information model**

