Limited Cognition and Clustered Asset Prices: Evidence from Betting Markets

Alasdair Brown*

University of East Anglia

Fuyu Yang†

University of East Anglia

July 30, 2015

*Corresponding Author. School of Economics, University of East Anglia, Norwich NR4 7TJ, U.K. Email: alasdair.brown@uea.ac.uk Tel: 44 1603 591131

†School of Economics, University of East Anglia, Norwich NR4 7TJ, U.K. Email: fuyu.yang@uea.ac.uk Tel: 44 1603 591058

Abstract

Asset prices tend to cluster at round numbers. We examine betting exchange data on U.K. horse races to establish whether limited cognition is partially responsible for this clustering. The key tool in this study is the stark increase in cognitive load faced by traders during races compared to prior to races. Using an approach that is part regression discontinuity and part difference-in-difference, we find that traders exhibit a substantially higher propensity to quote round numbers during races. This result is robust to a series of placebo tests, and also to the use of bounds to deal with missing data.

JEL Classification: G02, G12, G14

Keywords: limited cognition, price clustering, regression discontinuity, difference-in-difference

1 Introduction

There is long-standing evidence that stock prices cluster at round numbers, see Niederhoffer (1965, 1966), Harris (1991), Ahn *et al.* (2005), Sonnemans (2006), Ikenberry and Weston (2007), Bhattacharya *et al.* (2012), and Kuo *et al.* (2015). It is unlikely that the fundamental values of stocks cluster at round numbers, and therefore - at least recently, as models of bounded rationality (Simon, 1955) have become more popular in financial economics - price clustering has been attributed to limited cognition on the part of traders. The story is that attention-constrained traders latch onto cognitive reference points - such as round numbers - when they do not have the mental capacity to make finer price distinctions.

Yet, if limited cognition is indeed causing price clustering, we should observe greater (lower) levels of clustering when traders are more (less) cognitively constrained. In this paper, we use a U.K. betting exchange as a unique laboratory to test this hypothesis, and therefore more firmly establish the role of limited cognition in the clustering of asset prices.

We examine Betfair betting exchange trading on the two major horse race meetings in the U.K: Royal Ascot and the Cheltenham Festival. We compare the propensity for bettors to quote odds at a range of round numbers with their propensity to quote odds at the nearest non-round number. The round numbers and their nearest neighbour reflect almost identical win probabilities: the only difference is that one price is arguably a cognitive reference point, and therefore more likely to be quoted by the cognitively constrained bettor, and one is not. Importantly, we compare these propensities prior to races, when there is little new information arriving, with the same propensities during races, when important information - on the relative positions of horses - arrives in rapid succession. The idea is that bettors are more likely to be cognitively constrained - or to put it another way, their constraints are more likely to be binding - during races, when the cognitive load is greater. We should therefore observe more round number quotes during races.

We find that bettors are substantially more likely to quote a round number during races compared to prior to races. This result is perhaps most vividly captured by Figure 1. This figure contains a simple bar chart of quoted odds (including the stake) at Royal Ascot in 2011, both pre-race (in the left panel) and during races (in the right panel). While there is some evidence that certain round numbers are quoted more frequently than other odds prior to races, all round numbers are substantially more likely to be quoted once races get under way. This suggests that the cognitively constrained trader, forced to process the rapid unfolding of each race, latches on to round numbers as a heuristic method for pricing assets.

The logic of the empirical approach in this paper can be thought of as part regression discontinuity (RD) design, and part difference-in-difference (DID). The decision to quote a round number, or not, can be thought of as the 'running' or 'assignment' variable in an RD. In common with RD analysis, we assume that horses quoted at a round number, or just below, have almost identical characteristics (in this case, win probabilities). Therefore, in the absence of cognitive constraints, the bettor should be equally likely to quote a round number as its nearest neighbour. We are then interested in the extent to which the assignment decision (i.e. quoting a round number) is determined by the treatment. In our setting the treatment is the running of each race and the increase in cognitive load faced by bettors at this time. As with DID analysis, we compare the propensity to quote a round number relative to its nearest neighbour prior to races (pre-treatment), with the same propensity during races (during the treatment). We observe that the propensity to quote a round number - compared to its nearest neighbour - substantially increases as the cognitive load increases.

To ensure the robustness of our results, we conduct a series of placebo tests. One possible criticism with the analysis described thus far is that we have chosen the odds (i.e. the price) immediately below the round number as the nearest neighbour (so as to avoid undercutting). Therefore, if the running of each race causes the odds on the majority of horses to drift out, our results may subtly pick up this effect and assign it to cognitive constraints. In the placebo tests, therefore, we re-designate the nearest neighbour as the 'placebo' round number, and designate the odds just below the placebo point as the new nearest neighbour. If the greater propensity to quote a round number was simply capturing the drifting of odds during races, a similar effect should be observed in the placebo tests. In fact, no such effect is observed in the placebo tests, supporting the view that cognitive constraints were indeed driving the increased use of round numbers during races.

One other concern is over missing data in our samples. Missing data arises when there are no quotes on a horse to win, and is a particular problem during races. For example, there are no quotes in 3.6% of time periods in one of the markets during races at Royal Ascot in 2011. To deal with this 'attrition', we construct lower and upper bounds for the 'in-running' effect (during races). For the lower bound, we assume that all missing values in-running would have otherwise been a nearest neighbour. For the upper bound, we assume that all missing values in-running would have otherwise been a round number. In our earlier analysis we had implicitly assumed that missing values were equally likely to have been round numbers or nearest neighbours. These bounds therefore indicate how robust the results are to such an assumption. In fact, we find that even at the lower bound the in-running effect is substantial and highly significant. The analysis of bounds therefore confirms the effect of limited cognition on asset price clustering.

The analysis that we have described suggests that round numbers are used as heuristic methods for pricing assets when traders are cognitively constrained. There is still the possibility, however, that these round numbers are instead used to take advantage of the cognitive limitations of other bettors. For example, perhaps a trader quotes odds of 2, rather than 1.99, as they expect other traders to exhibit a 'left-digit bias'. This left-digit bias (observed in other economic settings to be discussed in the next section) means that cognitively constrained traders do not process all of the digits in a number, instead focusing only on the left-most digit. Therefore, odds of 2 would seem disproportionately larger than odds of 1.99, and therefore would be more likely to attract an order. Perhaps bettors quote round numbers during races so as to take advantage of this bias. There is indeed some evidence of a left-digit bias prior to races. A quote at a round number is slightly more likely to receive an order than a quote just below the round number. However, there is no evidence that this effect increases during races. We therefore conclude that the clustering of asset prices at round numbers is due to the cognitive limitations of the traders providing the quote, rather than an attempt to take advantage of the limitations of any trader that may order at that quote.

The setting for this study is chosen primarily for reasons of internal validity. To be specific, the separation of activity into pre-race and in-running trading creates stark variation in the cognitive load facing traders. Such variation is more difficult to identify in many core financial markets. That being said, results obtained on the betting exchange can still carry significant external validity. Betfair is the world's largest betting exchange, with 11.8 million GBP traded on a single race (*The Grand National*) in 2012 (Betfair Annual Report 2012). Secondly, the market structure is a limit order book (more on this in Section 3), which is similar to that used in the majority of financial markets. Finally, the exchange has been in operation since the year 2000, and while this does not guarantee that subjects are experienced, this does at least give bettors the opportunity to have become familiar with the market structure.

As asset price clustering is a form of mispricing, this work is related to a growing literature on the effect of limited cognition on asset mispricing. This literature includes, among others, DellaVigna and Pollet (2009) and Hirshleifer *et al.* (2009). These studies present evidence to suggest that mispricing (i.e. the post-earnings announcement drift) is more pronounced when traders are otherwise distracted, either because the weekend is imminent, or because there are a large number of announcements on the same day.

In using betting exchange data - and the concentration of information arrival during sporting events - we use a similar approach to that in Brown (2014). In that paper, Wimbledon tennis betting data was used to establish the role that information processing constraints play in asset mispricing. Mispricing between two equivalent assets - a bet on a player to win in one market, and a replicating portfolio of bets on the same player in another market - increases substantially as information arrives during each tennis match. Part of the reason for this increase is that information is not synchronously incorporated into the two prices, thereby suggesting that information processing constraints are binding. In this paper, we are interested in the effect of limited cognition on mispricing between one asset and its fundamental value, rather than relative mispricing between two equivalent assets. 'Absolute' mispricing, considered in this paper, can be measured by the extent to which asset prices cluster at round numbers.

While limited cognition is now arguably the dominant explanation for the clustering of asset prices at round numbers, ours is the first paper (to our knowledge) to show that the frequency of round number quotes increases when traders are subject to a greater cognitive load. In part, we are able to make this contribution due to the great contrast between the amount of information that needs to be processed during races, compared to prior to races. While earnings announcements and other announcements may increase the cognitive load on financial market traders, the contrast between these trading periods and other periods is less stark, as financial information arrives in an almost continuous fashion.

In addition, we also argue that our setting allows us to distinguish between different channels through which limited cognition may lead to round number prices. On the one hand, round numbers may arise due to the cognitive limitations of the trader who originally proposes the price. Kuo *et al.* (2015) show that traders on the Taiwan Futures Exchange who frequently quote round numbers are amongst the worst performers (in terms of trading profits). This suggests that traders' own cognitive limitations lie behind their predilection for round number quotes. On the other hand, round numbers may be used as a 'trick' by one set of traders to induce orders from other cognitively constrained traders. Bhattacharya *et al.* (2012) find pronounced selling at round numbers and pronounced buying just below these round numbers, suggesting a left-digit bias that can be exploited by traders if they quote round numbers. To distinguish between these two channels, the variation in cognitive load - pre-race vs in-running - in our setting is again useful. During races (when the cognitive load is greater), we observe a greater propensity to quote round numbers, but not a greater propensity to order at round numbers. Our analysis therefore suggests that round number quotes arise more often because of the cognitive limitations of the trader who proposes the price, rather than the trader that takes the price.

The rest of the paper is ordered as follows. In Section 2 we provide some psychological foundations for the analysis and survey the relevant economic literature. In Section 3 we describe the data, and in Section 4 we present the analysis. Section 5 concludes.

2 Psychological Foundations

In this paper we test whether an increase in the cognitive load facing traders - stimulated by the arrival of information during each race - increases the probability that traders quote round numbers. The designation of round numbers as 'cognitive reference points', likely to be latched on to by traders when they are cognitively constrained, is motivated by the work of Rosch (1975).

Rosch conducted a series of experiments, one of which showed that non-round numbers are evaluated in the context of their relation to round numbers. The experimental method was to give subjects a statement with blanks, such as _ is essentially _. Subjects would then be given two numbers to fill the blanks, one a round number such as 1000, and the other a non-round number such as 996. If 1000 was seen as the cognitive reference point, we should observe more subjects returning the statement '996 is essentially 1000' than the opposite statement. This was indeed the result that was found. To ensure the robustness of the results, Rosch also asked the same question with two non-round numbers and found that the two numbers were equally likely to be placed in either space.[1]

The designation of round numbers as cognitive reference points is not restricted to the laboratory. Pope and Simonsohn (2011) find that baseball season batting averages cluster just above 0.300, suggesting extra effort is exerted to push or keep the average above this level. Similarly, Pope and Simonsohn also find that students are more likely to request to resit their SATs if their initial score fell just below a round number, rather than just above. In other words, round numbers appear to be more salient than other numbers across a range of environments.

[1]We are not aware of any study in experimental psychology that examines whether the propensity to choose the round number as the cognitive reference point is affected by the amount of time pressure that subjects are under. Even if there were such a study, we may wish to see whether this effect also occurs in an economic setting (e.g. the betting market in our project).

Later in this study we will attempt to establish whether round numbers are used by cognitively constrained traders as a heuristic method for pricing assets, or are used to take advantage of the cognitive limitations of their potential trading counterparty. For example, those ordering at quotes already in the limit order book may exhibit what is termed, at least in the economics literature if not the psychology literature, the 'left-digit bias'. The identification of this bias is based on the work of Hinrichs *et al.* (1981) and Poltrock and Schwartz (1984) who found that subjects evaluate numbers sequentially, focusing on the left digit first. A cognitively constrained subject, therefore, may not get round to evaluating the second and third digits. The implication for our setting is that there may be a jump in the probability that a bettor places an order at odds of 2, compared to 1.99, but no such jump in the probability of an order at odds of 2.02 compared to 2. It is possible that traders quote a higher proportion of round numbers during races as this is when their potential counterparty is subject to a greater cognitive load, and therefore more likely to exhibit such a left-digit bias. We will test this hypothesis in the latter part of Section 4.

There is in fact quite prominent evidence consistent with a left-digit bias in other economic settings. For example, one explanation for the prevalence of 99 cents (pence) pricing is that consumers concentrate on the left digit, so, for example, a $1.99 product is regarded as disproportionately cheaper than a $2 product (Ginzberg, 1936). More recently, Lacetera *et al.* (2012) and Busse *et al.* (2013) identified a left-digit bias in the used car market. They observed large discontinuities in the prices of cars depending on the number of miles driven. For example, there were large drops in the price of a car once it hit 10,000 miles. Lacetera *et al.* (2012) also conducted a placebo test using Canadian data - where distance travelled is calculated in kilometres - and found a similar discontinuity at the 10,000 kilometre mark. Car owners, in the U.S. at least, appear to respond to this bias by bringing a sizeable number of cars with just less than 10,000 miles driven to the market.

3 Data

The data in this study is taken from Betfair, a betting exchange based in the U.K. The exchange allows bettors to bet on, or against, a horse to win. Bets on a horse to win are labelled 'back' bets, and bets on a horse to lose are labelled 'lay' bets. (A bettor laying a horse is allowing their counterparty to back the horse). Bettors are also able to wager on a horse to place (a horse places if they do not win but finish in a high-ranking position).

The exchange operates as a standard limit order book. This means that bettors can provide quotes (where they specify the odds), or execute their bets at odds quoted by others. When a bettor provides a quote, this is labelled a 'limit order', and when a bettor trades at a quote provided by another bettor, this is labelled a 'market order'. These betting options reflect a widening of choice for punters. Prior to the inception of these types of exchanges, bettors could only take long positions on a horse to win or place, and could only execute at prices quoted by a bookmaker. Now each individual bettor can submit quotes themselves.[2]

Figure 2 provides a screenshot of the Betfair limit order book. This example is taken from the 13.45 race at Wolverhampton on the 29th April 2015, a race outside of our sample. All back bet quotes (wagering on a horse to win) can be found on the left, with all lay bet quotes (betting on the horse to lose) on the right. The best back and lay quotes are nearest the centre, highlighted in blue and pink respectively. Prices on the exchange are quoted in the form of odds. Odds of 7, as quoted on Black Truffle for example, would return 7 GBP for every 1 GBP staked if the horse in question won. If the horse did not win, the trader that quoted that limit order would pocket the amount staked by their counterparty.

Horses are listed in each market from the expected favourites (those with the lowest odds) at the top, to the expected longshots at the bottom. Bettors who wish to provide quotes can specify the direction of their bet (back or lay), the odds that they wish to offer (or request), and the volume involved in the bet. If these specific requirements are not fully met by existing quotes, then the punter's quote enters the book as a limit order.

The pricing grid - which determines the precision of quotes - becomes more coarse for higher odds. For example, odds can be quoted in increments of 0.01 between 1 and 2 (i.e. 1.01, 1.02 and so on), but only in increments of 0.02 between 2 and 3.[3] The increased coarseness of

[2] Betfair also operate a parallel service, the Sportsbook, where they operate as a traditional bookmaker. The service is intended for less popular sporting events, where liquidity on the main exchange may be limited. Historical data for Betfair's Sportsbook pricing is not available.

[3] Odds are quoted in increments of 0.05 between 3 and 4, of 0.1 between 4 and 6, of 0.2 between 6 and 10, of 0.5 between 10 and 20, of 1 between 30 and 50, of 5 between 50 and 100, and of 10 between 100 and 1000.

the pricing grid for higher odds is one reason why we choose fewer round numbers for higher odds (more on this in Section 4). If a bettor enters an impermissible quote, such as 2.01 for example, then the exchange automatically rounds up the quote to the nearest permissible number (in this case, 2.02). As with quotes at permissible odds, the bettor can then decide whether to proceed with the bet or not.[4]

Once a quote is in the book, other bettors can take the other side of the bet. However, if a bet has not yet been taken, the bettor providing the quote is free to cancel or update the price and volume that they wish to bet. This freedom is particularly important during races, as quotes will quickly become outdated.

The data in this paper are sourced from Fracsoft, a third-party vendor of Betfair limit order book data. We have chosen to study two prominent horse race meetings across two years. We examine the premier flat racing meeting in the U.K., Royal Ascot, which takes place in June each year, and also the premier jump racing event, the Cheltenham Festival, which takes place in March each year. Royal Ascot lasts for 5 days, with 6 races per day, while the Cheltenham Festival lasts for 4 days, with 7 races per day.[5] At least 394 horses take part in each of these meetings.

For this study, we examine the odds available on a horse to win and also the odds on a horse to place. These odds are displayed in separate markets. For the most part we examine quotes for 'back' bets in each market, as this is typically the more liquid side of a market and also where the majority of bets are executed. For every horse, the best-priced quotes are sampled every second for 20 minutes prior to each race (the pre-running phase), and every second for the full duration of each race (the in-running phase). As there is a 30-40 minute

[4]It is possible that Betfair's procedure for dealing with impermissible odds contributes to the increase in the frequency of round numbers quoted during races. If bettors are more likely to make mistakes in-running - and quote impermissible odds - then we will see quotes of 2.99, for example, rounded up to 3 (our designated round number) rather than rounded down to 2.98 (our designated nearest neighbour). However, this potential confound should not be such an issue when we designate 2 as our round number, and 1.99 as our nearest neighbour. Without going beyond 2 decimal places, and therefore expecting payment in a fraction of pennies, it is not possible for bettors to enter an impermissible quote in this range. And, as we shall see in Table 8, quotes of 2 still become much more likely than quotes of 1.99 during races.

[5]Data is available on the Fracsoft database for all days, except the 17th June 2011 and the 19th June 2012, both at Ascot.

gap between the start times of each race, pre-race bettors are not distracted by the running of a previous race.

In Figure 1 we displayed a bar chart of odds at Royal Ascot in 2011 prior to races (left panel) and during each race (right panel). We now replicate this display for the other three meetings in our sample: Royal Ascot in 2012 (Figure 3), and the Cheltenham Festival in 2011 (Figure 4) and 2012 (Figure 5). These bar charts only display odds (including the stake) of below 11, so as to cram as many important round numbers into the plot. There is a similar pattern in Figures 3-5 to that observed in Figure 1. To be specific, while there is some evidence of clustering at round numbers prior to races, there is clear evidence of clustering at round numbers when races go 'in-running'. This suggests traders use round numbers as a heuristic method to price assets when under severe time constraints during races. In the next section, we will define the full classification of round numbers in this study, and then proceed to the formal tests.

4 Analysis

Betfair allow for quotes from 1.01 (i.e. 1 penny returned for each pound staked), to 1000 (999 pounds returned for each pound staked). We classify the round numbers in this study as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, and 50. This includes the majority of horses that are in contention, but not so certain to win or place that pricing their bets becomes a simple exercise. The round numbers reflect implied probabilities from 0.5 (odds of 2), to 0.02 (odds of 50). We pick fewer round numbers from 10 onwards as the pricing grid becomes significantly more coarse at this point.

We then classify the price immediately below each of these round numbers as the nearest neighbours. These odds are 1.99, 2.98, 3.95, 4.9, 5.9, 6.8, 7.8, 8.8, 9.8, 14.5, 19.5, 29, 38, and 48. We pick the odds just below the round number as a control price because quotes just above the round number could become prevalent because of undercutting. To be specific, a cognitive constrained trader may quote a round number - because of its role as a cognitive reference point - but be immediately undercut by another (perhaps algorithmic) trader simply using the initial quote as a reference point and improving upon it. In Figures 1 and 3-5 there is visible tapering of the frequency of quotes to the right of a round number, particularly during races, which suggests that systematic undercutting does take place.

Before we proceed, we should clarify why we are interested in the prevalence of round numbers relative only to their nearest neighbours, rather than relative to all other numbers. The main issue is that the distribution of prices (odds) changes greatly between pre-race and in-running periods. Prior to the races in our 4 meetings, 80.4% of quotes can be found in the interval between 2 and 50, our smallest and largest designated round numbers. Contrast this with the in-running period, when only 68.7% of quotes can be found in the interval between 2 and 50. As each race nears completion, the imminent winners will be quoted at 1.01 (if at all), and the imminent losers will be quoted at 1000 (if at all). This means that non-round numbers - outside of our range between 2 and 50 - will likely increase in prevalence during races, simply because horses drift to extreme odds and the fundamental value of the bet either becomes very low or very high.

With this change in the distribution of odds in mind, we wanted to find appropriate odds to compare to our round numbers. We therefore chose only the odds closest to the round numbers, their nearest neighbours. Round numbers and their nearest neighbours reflect almost identical win probabilities and therefore, in the absence of cognitive constraints - we should be just as likely to observe a round number as its nearest neighbour. This argument applies both pre-race and in-running. The result is that if a round number becomes more prevalent than its nearest neighbour during races, after controlling for any differences pre-race, we can more reliably apportion this effect to limited cognition.

By comparing the incidence of round numbers only to their nearest neighbours, we also reduce the importance of our choice of round numbers. We chose every whole number from 2 to 10, as many horses will find themselves quoted in this range, and then tapered the number of round numbers from then on (15, 20, 30, 40 and 50), as the pricing grid becomes more coarse for higher odds. One might ask, however, why not 25? Any why not 100, or 200? And going the other way, why not 1.1 through to 1.9, as these numbers are round compared to 1.11 and 1.91? A case could be put for each of these numbers, but the important thing is that by comparing our round numbers only to their nearest neighbours, we negate the effect of any drift in fundamental values during races. A quote of 50, for example, could perhaps become more likely as races progress and horses drop out of contention, but, in the absence of cognitive constraints, such a price is almost equally likely as a price of 48 or 55. Similarly a quote of 2 may become more likely as a horse nears the finish line in the lead, but, again, in the absence of cognitive constraints, such a quote should be just as likely as a quote of 1.99 or 2.02.

In Table 1 we provide summary statistics on the four race meetings separately. We examine the percentage of seconds where there is at least one quote, and the percentage of seconds where there is a round number quote, or a nearest neighbour quote. These percentages are compared for back and lay bets, and for pre-race and in-running phases. (For the nearest neighbours of lay bets, we consider the odds immediately above the round number, as undercutting on lay bets will work in the opposite direction. This means odds of 2.02, 3.05, 4.1, 5.1, 6.2, 7.2, 8.2, 9.2, 10.5, 15.5, 21, 32, 42, and 55).

First, it is apparent that liquidity declines, particularly on the lay side, during races. For example, we only have lay quotes 54.96% of the time for the win market at Royal Ascot in 2011 during races. This motivates the predominant use of back bets in this study. Second, across both markets, and across all events, the pre-race proportion of round numbers and nearest neighbours is approximately 10%. Once we move in-running, however, there are marked differences between the two markets. While the proportion of round numbers and nearest neighbours declines slightly in the win market, there are substantial increases in round numbers in the place market. As argued a little earlier, round numbers and nearest neighbours can be found in the central mass of the implied probability distribution. Therefore, once a race begins, the probability that a horse wins drifts towards the tails - either as victory becomes likely, or defeat inevitable - thereby leaving fewer observations at round numbers or nearest neighbours. This does not necessarily occur in the place market, where a greater proportion of horses may remain in contention for much of the race. These summary statistics illustrate why it is sensible to compare the frequency of round numbers, not with all other quotes, but only with their nearest neighbours.

The analysis that follows therefore excludes all quotes except the round numbers and their nearest neighbours. This means discarding a great deal of data. For example, the Royal Ascot 2011 place market data includes 720,880 observations in total, which is reduced to 114,967 when we consider only round numbers and their nearest neighbours. By focusing on this small set of numbers, we are only picking quotes which reflect almost identical win or place probabilities. For example, odds of 4 in the win market imply a win probability of 0.25, while the nearest neighbour odds of 3.98 reflect an implied win probability of 0.251. Therefore, we assume that 'nature' is almost equally likely to draw a horse with a win probability of 0.25 as one with a probability of 0.251. The key difference is that one of the odds is a round number - and therefore, we hypothesise, more likely to be latched onto by the cognitively constrained bettor - and one is not. To ensure that limited cognition is indeed behind the use of round numbers, we then exploit variation in the cognitive load facing traders. This variation is created by the running of each race, and the resultant need to process rapidly arriving information - and price bets - in a short space of time.

Table 2 presents the results of the main analysis. In this table, we regress an indicator variable equalling 1 if a round number (previously defined) is quoted, on an indicator variable equalling 1 if the race was in-running at that stage. Only data on round numbers, and the odds immediately below (nearest neighbours) are included. A logit specification is used, and random effects are included for each horse (bet) in the sample.[6] In the top panel we include win market data, with place market data in the bottom panel. These regressions are carried out for Royal Ascot in 2011 and 2012, and the Cheltenham Festival in 2011 and 2012. The coefficient associated with the intercept term suggests that there is a small baseline prevalence for quoting round numbers. The treatment effect, however, captured by the in-running indicator, is of greater importance. We find that as races go in-running there is a large and significant increase in the propensity to quote a round number, for place markets, and a smaller increase for win markets (all significant at the 0.1% level). This is the case across all four horse race meetings. The coefficients imply an odds ratio of between 1.29 (Royal Ascot 2012, win market) and 5.25 (Cheltenham 2012, place market) of observing a round number quote in-running relative to prior to races. This treatment effect suggests that the clustering of quotes at round numbers is indeed a manifestation of limited cognition.[7]

[6] In addition to this specification, we also checked the results with two-way clustering of standard errors, albeit without random effects, when the two clusters were the bet and the race, and also when the two clusters were the bet and the time. Our results were qualitatively similar to those that we present.

[7] At this stage we should clarify why there are differences in the number of random effects (bets) in win

One concern at this stage is that, due to the prevalence of undercutting, we have exclusively chosen nearest neighbours which are priced just below the round number. Therefore, we may be confounding the limited cognition effect with subtle drifts in the odds which occur when a race begins. As the majority of horses do not win or place, most will see their odds increase during the in-running period. Therefore a comparison of the frequency of, for example, odds of 2.98 and 3 could find slightly more 3s when races are under way, as horses which were hitherto priced at 2.98 slowly drift out of contention.

To deal with this problem, we run a series of placebo tests on the 4 race meetings. For these tests, we define the original nearest neighbours - 1.99, 2.98, 3.95, 4.9, 5.9, 6.8, 7.8, 8.8, 9.8, 14.5, 19.5, 29, 38, and 48 - as the ‘placebo’ round numbers. We then create new nearest neighbours just one price down. These new nearest neighbours are 1.98, 2.96, 3.9, 4.8, 5.8, 6.6, 7.6, 8.6, 9.6, 14, 19, 28, 36, and 46. If our results thus far can be apportioned to a general drift in the odds during races, we should observe a similar effect when we repeat the analysis with these odds.

In Table 3, we regress a indicator variable equalling 1 if a placebo round number is quoted, on an indicator variable equalling 1 if the race in question was in-running at that stage. Only placebo round numbers, and the odds immediately below them, are included in the regressions. As before, a logit specification is used, and random effects for each horse (bet) are included. Win (place) market data is in top (bottom) panel. Judging by the eight regressions in Table 2 - one for each market in each race meeting - there is no evidence that the in-running effect is due to a subtle drift in the odds in-running. In fact, the coefficients associated with the in-running indicator are negative and significant, albeit quite small. This suggests that the round numbers absorb such a high proportion of their nearest neighbours that these neighbours are slightly less likely to arise than other non-round numbers. These results re-affirm the impression that limited cognition was driving the increase in round number quotes during races observed in Table 2.

---

and place markets for the same meeting (e.g. Ascot 2011). As we are only using the quotes at round numbers or nearest neighbours, a bet is only incorporated in our analysis if, at some point pre-race or in-running, it is priced at one of these figures. We are likely to see more bets in our place market analyses than our win market analyses, as place odds are more likely to be in the middle of the implied win probability distribution, where most of our round numbers and nearest neighbours are found.

Another approach to deal with any drift in odds during races is to consider quotes on the lay side of the bet. For lay bets, undercutting will work in the opposite direction. A quote of 2, for example, will be undercut to 1.99, rather than 2.02. Therefore, we can classify the numbers immediately above as nearest neighbours - for lay bets - and examine whether the limited cognition effect is observed despite (rather than because of) any drift in the odds during races. In Table 4, we regress an indicator variable equalling 1 if a lay bet was quoted at a round number, on an indicator variable equalling 1 if the race was in-running at that point. As before, win (place) market data is in the top (bottom) panel, a logit specification is used, and random effects for each horse (bet) are included. The intercept term suggests that lay quotes are more likely to be nearest neighbours than round numbers prior to races. If you have a back bet at a round number, the lay bet cannot be at the same round number (otherwise the two bets will execute with each other, and immediately disappear from the book). Therefore, if a bettor quotes the minimum spread (as is common prior to races), we are likely to observe many lay bets at nearest neighbours, purely as a result of the observed tendency to quote back bets at round numbers. Once races go in-running however, and spreads widen, this mechanical effect is not present. For all four events, and both markets, the in-running propensity to quote round numbers is observed (with significance at the 0.1% level) despite any natural drift in average odds during races.

We next turn to concerns regards missing data. As races begin, liquidity on the exchange can dry up in certain instances. The result is that there is a non-trivial proportion of time periods when there are no quotes on a particular horse to win or place. These non-trivial proportions can be observed in Table 1. For back bets, there are up to 5% of observations without any quote at all (for lay bets the proportion is much higher). We have implicitly assumed, in the analysis thus far, that had there been a quote in these time periods, then a round number is equally likely as a nearest neighbour. It is worth examining how robust the results for back bets are to such an assumption.

With this in mind, in Table 5 we calculate upper and lower bounds to the in-running treatment effect described in Table 2. The upper bound is calculated by assuming that all missing observations during races are in fact round number quotes. (Incidentally, we believe that this is the more plausible of the bounds, as a cognitively constrained trader is either likely to quote a round number (as Table 2 suggests), or not provide any quotes at all). Win market results are in the top panel, with place market results in the bottom panel. The upper bound results are displayed in the top sub-panel of each panel in Table 5. The lower bound is calculated by assuming that all missing observations during races are in fact nearest neighbour non-round number quotes. These results are displayed in the bottom sub-panel of each panel in Table 5. Judging by the lower bound results, the problem of missing data does not qualitatively affect the results. As we mentioned, we would lean towards believing that the missing data would disproportionately have been round numbers, rather than non-round numbers, but the in-running effect is sizeable enough that such an assumption does not need to be made, at least not with back quotes. (Repeating this analysis for lay bets would be futile as data are missing in up to 46% of time-periods in-running).[8]

The next part of analysis in the paper is an attempt to establish the drivers of increased clustering during races. The interpretation taken throughout the paper is that a cognitively constrained trader uses round numbers as a short-cut pricing method. If more time and mental capacity had been available, a more precise quote may have been provided. An alternative interpretation is that the trader submitting the limit order (i.e. providing the quote) submits a round number to take advantage of the limited processing capabilities of other traders during races. In other words, limit order traders believe that market order traders will spy relative value in a quote of 2, compared to 1.99, when races go in-running, because their limited cognition does not allow them to spot the similarities between the two prices. The empirical prediction of such a hypothesis is that the frequency of orders should jump at round numbers, relative to nearest neighbour non-round numbers, and jump specifically during races when traders are cognitively constrained.

In Table 6 we test this alternative hypothesis. We regress an indicator variable equalling 1 if a (back) order was observed in the last second, on an indicator variable equalling 1 if a round number was quoted in the previous second, an indicator variable equalling 1 if the race was in-running, and an interaction between the two aforementioned indicators. The interaction

[8]As with our earlier analyses, there are differences in the number of random effects (bets) for win and place markets. This is because a bet is only included if it is at some point quoted at a round number, a nearest neighbour - or, in the case of the Table 5 analysis - a missing value (no quote at all).

term is crucial as this captures the extent to which market order traders fail to identify the similarities between round number and nearest neighbour quotes, specifically during races. Only data on round numbers, and the odds immediately below (nearest neighbours) are included in this regression. As before, a logit specification is used and random effects are included for each horse (bet). This regression is carried out for win and place market data for the 4 horse race meetings.

There are three results to be discussed. Firstly, in all 8 cases, quotes at round numbers are slightly more likely to attract an order. This is consistent with a left-digit bias, but if this was the result of interest, a placebo test would need to be carried out to ensure that this is not simply picking up the fact that higher odds attract more frequent orders. The second result is that orders decline as races begin. Finally, and most importantly, there is no evidence that the propensity for orders to arrive for round number quotes increases during races. The coefficient is negative and significant for five of the regressions, and insignificant for the other three. In other words, rather than a tactic to induce orders from constrained counterparties, the round number appears to be used as a heuristic method for pricing bets when finer information processing is not feasible.[9]

Throughout the paper it appears that the effect of the running of each race, on the propensity to quote round numbers, is greater in the place market. There are two prior reasons to predict such a difference. Firstly, the place market is to some extent secondary in importance. Media and television will tend to focus on the odds in the win market, and volume is typically higher in that market. Therefore, if a trader is cognitively constrained during a race, they may choose to concentrate their limited attention on the win market, thereby neglecting the place market. This neglect may manifest itself in the form of a greater proportion of round, and imprecise, quotes. The second, and perhaps more important reason, is that a large proportion of horses remain in contention to place during races. While only a few horses can win a race entering its final stages - and therefore only a few bets still need

[9]While looking at orders, we also examined whether bettors are more likely to submit round trade sized orders, e.g. multiples of 10 GBP, during races. Alexander and Peterson (2007) and Dennis (2012), in studies of stock market transactions, find that trade sizes cluster at round numbers. Perhaps surprisingly, we found that trade sizes were more likely to be round numbers in the pre-race period than in the in-running period.

to be priced - the place positions are typically still within the reach of many horses as a race nears its conclusion. This should heighten the cognitive load for bettors in the place market.

To test this hypothesis, we pooled the win and place market data for a new (double) difference-in-difference analysis in Table 7. In this table, we regress an indicator variable equalling 1 if a round number was quoted, on an indicator variable equalling 1 if the race was in-running, an indicator variable equalling 1 if the odds came from the place market, and an interaction between the two aforementioned indicators. The interaction term captures any significant differences in the in-running effect across the win and place markets. As before, only data involving round numbers and their nearest neighbours are included. A logit specification is again used, but this time random effects are clustered at the bet level for each horse-market combination.

The intercept term confirms that there is a baseline prevalence for the quoting of round numbers (as in Table 2), and the in-running coefficient confirms that there is an increase in this prevalence during races, as the cognitive load increases (also in Table 2). Notably, as the interaction coefficient suggests, the largest effect can be found in the place market. As discussed earlier, there is some reason to believe that this is because a greater proportion of horses remain in contention for places throughout races - needing to be priced - and therefore the cognitive load is likely to be greater, during races, for participants in this market.

Our next piece of analysis is intended to check the breadth of the in-running effect documented throughout the paper. While the effect is robust across markets, across events, and across time (at least from 2011 to 2012), it would be instructive to examine whether the effect is robust across all of the round numbers we have specified. To this end, we retained the pooled win and place market data, and furthermore pooled the data for Royal Ascot and the Cheltenham Festival in both years. This gives the following tests substantial statistical power. In Table 8, we repeat the analysis of Table 2, but this time for each individual round number. For example, when looking at the in-running effect at the round number of 3, only quotes of 3 and 2.98 are included. The results show that the propensity to quote a round number, relative to a nearest neighbour non-round number, increases as races go in-running for all 14 round numbers used (significance at least at the 1% level). Although the size of the effect is not monotonic with the size of the odds, there is some evidence that the greatest propensity to quote round numbers - both pre-race and in-running - is found for lower round numbers, particular those from 2 to 5.

One final argument that may be put forward is that the increased incidence of round numbers during races is not due to limited cognition, but is instead due to another factor that arises due to time-pressure. Bettors may quote round numbers during races as, in most cases, round numbers require the inputting of fewer digits. In other words, rather than enter 1.99 as the quote for a back bet, a punter enters 2 - which is almost the same price - to save entering the two digits that follow the decimal place. Such a shortcut allows the bettor to update their quotes, and secure trade, much more quickly. While speed is not so important prior to races, it is during races.

To deal with such a concern, we take advantage of the fact that a subset of the round numbers in our analysis have the same number of digits as their nearest neighbours. For 15, 20, 30, 40 and 50, the nearest neighbours are 14, 19, 28, 36, and 46 respectively. In other words, it takes just as much time to input the round number price as the nearest neighbour. In contrast, round numbers of 10 and below involve fewer inputs than their nearest neighbours. (In the case of 10, there are two characters inputted for the round number, and 3 characters for the nearest neighbour, 9.6). Therefore, we argue that the treatment effect for higher round numbers - from 15 to 50 - can be more reliably attributed to limited cognition.

With this in mind, in Table 9 we present the results of two final regressions. We retain the pooled data from the Table 8 analysis - to ensure the greatest statistical power - and break the full sample down into low round numbers (2 to 10) and high round numbers (15 to 50). We then repeat the regressions of Table 2. We find that both low and high round numbers are more prevalent than their nearest neighbours, particularly during races. The running of each race does increase the incidence of lower round numbers to a greater extent than high round numbers, suggesting that limited time to input digits may well be a factor. However, the effect for high round numbers is still highly significant, leading us to believe that limited cognition is a contributor to the clustering of asset prices at round numbers.

5 Conclusion

Asset prices tend to cluster at round numbers. This clustering is found not just in stock prices, but also in currency prices (Sopranzetti and Datar, 2002), bond prices (Gwilym *et al.*, 1998a), and option prices (Gwilym *et al.*, 1998b). Furthermore, clustering appears to be resistant to changes in market design. Ikenberry and Weston (2007) find little decrease in clustering after the decimalisation of prices - (prices had previously been quoted in coarser dollar fractions) - on the New York Stock Exchange and the Nasdaq in 2001. This would suggest that the use of round numbers is a 'hard-wired' reflex action for the cognitively constrained trader.

Missing from this story, however, is any evidence that the level of clustering varies with the cognitive load facing traders. If limited cognition is truly behind price clustering, we should observe particularly pronounced levels of clustering when traders are under an intense cognitive load. Rather than test such an idea in the laboratory - where the experience of subjects may be in some doubt - we use data arising from horse race betting on a U.K. betting exchange to test this hypothesis. Information in this market is produced in such a way that there is stark variation in the cognitive load facing traders. Prior to races there is little new information to process, and traders can therefore make fine price distinctions. Once the race begins, however, and horses switch positions frequently, information must be processed - and bets newly priced - in a very short space of time. We hypothesise that round numbers will become more prevalent in this latter period.

To tease out such an effect, we use an empirical approach logically based partly on regression discontinuity (RD) methods, and partly on difference-in-difference (DID) methods. We compare the propensity for traders to quote round numbers with their propensity to quote the nearest non-round number. These two prices reflect almost identical values and therefore, if traders had unlimited cognition, each price should be equally likely to be quoted (or 'assigned', to use the RD parlance). Unlike RD analysis, we are not interested in the outcome for the assigned and unassigned variables, but instead we are interested in how the assignment decision varies across pre-race (pre-treatment) periods, and in-running (treatment) periods. If traders display a greater propensity to quote round numbers rather than their nearest neighbours during races - when the cognitive load is greater - this puts limited cognition front-and-centre as an explanation for the clustering of asset prices.

We find a significant and substantial increase in the proportion of round number quotes during races. The in-running effect is most pronounced in the place market, where a large proportion of horses remain in contention throughout races (therefore needing to be priced), and where, as a result, the cognitive load is arguably most demanding. The results are also robust to a series of placebo tests, and, due to the sheer scale of the effect, are robust to the use of bounds to deal with substantial amounts of missing data (at least with the more popular back quotes). Finally, the results also apply when inputting a round number requires as much effort as entering a nearest neighbour (i.e. when the round number has the same number of digits as its neighbour). In short, the analysis presented in this paper suggests that limited cognition does indeed lead to asset price clustering.

References

- Ahn, H., J., Cai, J., Cheung, Y., L., (2005). Price Clustering on the Limit-Order Book: Evidence from the Stock Exchange of Hong Kong. *Journal of Financial Markets*, 8, 421-451.

- Alexander, G., J., Peterson, M., A., (2007). An Analysis of Trade-Size Clustering and its Relation to Stealth Trading. *Journal of Financial Economics*, 84, 435-471.

- Bhattacharya, U., Holden, C., W., Jacobsen, S., (2012). Penny Wise, Dollar Foolish: Buy-Sell Imbalances On and Around Round Numbers. *Management Science*, 58, 413-431.

- Brown, A., (2014). Information Processing Constraints and Asset Mispricing. *Economic Journal*, 124, 245-268.

- Busse, M., R., Lacetera, N., Pope, D., G., Silva-Risso, J., Sydnor, J., R., (2013). Estimating the Effect of Salience in Wholesale and Retail Car Markets. *American Economic Review: Papers and Proceedings*, 103, 575-579.

- DellaVigna, S., and Pollet, J. M., (2009). Investor Inattention and Friday Earnings Announcements, *Journal of Finance*, 64, 709-749.

- Dennis, P., (2012). Trade Size Clustering. Working paper.

- Ginzberg, E., (1936). *Customary Prices. American Economic Review*, 26, 296.

- Gwilym, O., P., Clare, A., Thomas, T., (1998a). Price Clustering and Bid-Ask Spreads in International Bond Futures. *Journal of International Financial Markets, Institutions and Money*, 8, 377-391.

- Gwilym, O., P., Clare, A., Thomas, T., (1998b). Extreme Price Clustering in the London Equity Index Futures and Options Markets. *Journal of Banking and Finance*, 22, 1193-1206.

- Harris, L., (1991). Stock Price Clustering and Discreteness. *Review of Financial Studies*, 4, 389-415.

- Hinrichs, J., V., Yurko, D., S., Hu, J., M., (1981). Two-Digit Number Comparison: Use of Place Information. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 890-901.

- Hirshleifer, D., Lim, S., S., Teoh, S., H., (2009). Driven to Distraction: Extraneous Events and Underreaction to Earnings News. *Journal of Finance*, 64, 2289-2325.

- Ikenberry, D., L., Weston, J., P., (2007). Clustering in US Stock Prices after Decimalisation. *European Financial Management*, 14, 30-54.

- Kuo, W., Y., Lin, T., C., Zhao, J., (2015). Cognitive Limitation and Investment Performance: Evidence from Limit Order Clustering. *Review of Financial Studies*, 28, 838-875.

- Lacetera, N., Pope D., G., Syndor, J., R., (2012). Heuristic Thinking and Limited Attention in the Car Market. *American Economic Review*, 102, 2206-2236.

- Niederhoffer, V., (1965). Clustering of Stock Prices. *Operations Research*, 13, 258-265.

- Niederhoffer, V., (1966). A New Look at the Clustering of Stock Prices. *Journal of Business*, 39, 309-313.
- Poltrock, S., E., Schwartz, D., R., (1984). Comparative Judgements of Multidigit Numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 32-45.
- Pope, D., Simonsohn, U., (2011). Round Numbers as Goals: Evidence From Baseball, SAT Takers, and the Lab. *Psychological Science*, 22, 71-79.
- Rosch, E., (1975). Cognitive Reference Points. *Cognitive Psychology*, 7, 532-547.
- Simon, H., A., (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*, 69, 99-118.
- Sonnemans, J., (2006). Price Clustering and Natural Resistance Points in the Dutch Stock Market: A Natural Experiment. *European Economic Review*, 50, 1937-1950.
- Sopranzetti, B., J., Datar, V., (2002). Price Clustering in Foreign Exchange Spot Markets. *Journal of Financial Markets*, 5, 411-417.

Tables & Figures

Table 1: Summary Statistics				
Royal Ascot 2011				
Win Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	97.26%	12.28%	9.82%
	Lay Odds	97.18%	9.96%	12.14%
In-Running	Back Odds	98.15%	8.01%	4.07%
	Lay Odds	54.96%	6.66%	5.71%
Place Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	97.26%	12.20%	9.18%
	Lay Odds	97.26%	10.80%	11.04%
In-Running	Back Odds	96.45%	20.20%	4.07%
	Lay Odds	76.45%	13.93%	10.60%
Royal Ascot 2012				
Win Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	96.08%	12.90%	9.99%
	Lay Odds	95.20%	10.21%	12.90%
In-Running	Back Odds	95.43%	10.98%	6.87%
	Lay Odds	76.22%	9.01%	6.82%
Place Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	96.08%	10.95%	9.42%
	Lay Odds	96.08%	10.80%	10.17%
In-Running	Back Odds	94.65%	16.54%	4.76%
	Lay Odds	66.33%	9.99%	4.81%
Cheltenham Festival 2011				
Win Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	98.62%	12.34%	9.66%
	Lay Odds	98.36%	9.87%	12.33%
In-Running	Back Odds	98.38%	10.13%	6.69%
	Lay Odds	77.89%	8.76%	7.20%
Place Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	98.58%	11.76%	10.23%
	Lay Odds	98.58%	12.39%	10.53%
In-Running	Back Odds	98.04%	18.79%	5.23%
	Lay Odds	70.24%	12.58%	6.67%
Cheltenham Festival 2012				
Win Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	98.34%	11.59%	9.37%
	Lay Odds	97.63%	9.59%	11.50%
In-Running	Back Odds	98.11%	10.02%	6.37%
	Lay Odds	77.31%	8.63%	7.46%
Place Market		Quotes	Round Number	Nearest Neighbour
Pre-Race	Back Odds	98.34%	10.35%	9.91%
	Lay Odds	98.34%	11.66%	9.53%
In-Running	Back Odds	97.72%	19.04%	4.23%
	Lay Odds	67.04%	14.18%	5.60%

The percentage of seconds with at least one quote, and the percentage of seconds with a quote at a round number or a nearest neighbour. The round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 and 50. The nearest neighbours are the odds immediately below (for a back quote), or immediately above (for a lay quote). These statistics are compared for pre-race and in-running trading periods. Data for the win and place market for Royal Ascot and the Cheltenham Festival in 2011 and 2012 are displayed.

Table 2: Main Test Results (Logit)				
Win Market (Back Quotes)				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.475***	0.374***	0.391***	0.367***
	(.066)	(.072)	(.065)	(.057)
In-Running	0.442***	0.253***	0.281***	0.255***
	(.026)	(.027)	(.018)	(.017)
No. of Observations	115,861	120,606	137,461	146,264
ρ	0.294	0.337	0.308	0.275
No. of Random Effects (Horses)	342	340	363	421
Place Market (Back Quotes)				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.435***	0.278***	0.25***	0.288***
	(.061)	(.056)	(.055)	(.056)
In-Running	1.437***	1.237***	1.334***	1.666***
	(.028)	(.027)	(.017)	(.017)
No. of Observations	114,967	110,030	143,222	153,083
ρ	0.283	0.248	0.259	0.293
No. of Random Effects (Horses)	375	374	409	475

The main test results in the paper. An indicator variable equalling 1 if a round number is quoted, is regressed on an indicator variable equalling 1 if the race is 'in-running' at that time-point. The round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 and 50. Only quotes at these odds, or those immediately below, are included in the regressions. Win market data is in the top panel and place market data is in the bottom panel. A logit specification is used, random effects for each horse (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and * indicates significance at the 0.1% level.**

Table 3: Placebo Test Results (Logit)				
Win Market (Back Quotes)				
Dep. Var.: Odds=Placebo Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	-0.076	0.084	0.068	0.011
	(.06)	(.058)	(.052)	(.045)
In-Running	-0.197***	-0.244***	-0.187***	-0.063**
	(.028)	(.029)	(.019)	(.018)
No. of Observations	101,833	100,659	116,905	128,050
ρ	0.253	0.239	0.21	0.183
No. of Random Effects (Horses)	334	330	350	401
Place Market (Back Quotes)				
Dep. Var.: Odds=Placebo Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	-0.048	0.085	0.047	-0.099
	(.056)	(.054)	(.051)	(.056)
In-Running	-0.19***	-0.205***	-0.091***	-0.081***
	(.034)	(.034)	(.02)	(.022)
No. of Observations	95,992	96,976	119,425	127,408
ρ	0.243	0.226	0.228	0.284
No. of Random Effects (Horses)	357	353	394	450

Placebo test results connected to the analysis in Table 2. An indicator variable equalling 1 if a placebo round number is quoted, is regressed on an indicator variable equalling 1 if the race is 'in-running' at that time-point. The placebo round numbers are defined as 1.99, 2.98, 3.95, 4.9, 5.9, 6.8, 7.8, 8.8, 9.8, 14.5, 19.5, 29, 38 and 48. Only quotes at these odds, or those immediately below, are included in the regressions. Win market data is in the top panel and place market data is in the bottom panel. A logit specification is used, random effects for each horse (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and *, and ** indicates significance at the 0.1% and 1% level respectively.**

Table 4: Lay Bet Results (Logit)				
Win Market (Lay Quotes)				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	-0.316***	-0.275***	-0.269***	-0.249***
	(.062)	(.073)	(.068)	(.059)
In-Running	0.422***	0.543***	0.371***	0.367***
	(.025)	(.028)	(.018)	(.017)
No. of Observations	116,033	120,723	137,335	146,558
ρ	0.27	0.334	0.32	0.285
No. of Random Effects (Horses)	331	325	351	404
Place Market (Lay Quotes)				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.033	0.136*	0.281***	0.418***
	(.06)	(.061)	(.057)	(.06)
In-Running	0.221***	0.799***	0.444***	0.793***
	(.023)	(.03)	(.017)	(.017)
No. of Observations	117,309	109,829	141,700	153,074
ρ	0.278	0.279	0.275	0.327
No. of Random Effects (Horses)	372	369	408	472

A replication of the results in Table 2, but this time conducted with lay quotes. An indicator variable equalling 1 if a round number is quoted, is regressed on an indicator variable equalling 1 if the race is 'in-running' at that time-point. The round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 and 50. Only quotes at these odds, or those immediately *above*, are included in the regressions. Win market data is in the top panel and place market data is in the bottom panel. A logit specification is used, random effects for each horse (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and * and * indicates significance at the 0.1% and 5% level respectively.**

Table 5: Upper and Lower Bounds (Logit)				
Win Market (Back Quotes)				
Upper Bound				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.555***	0.558***	0.459***	0.475***
	(.07)	(.085)	(.071)	(.065)
In-Running	0.559***	0.328***	0.308***	0.268***
	(.026)	(.027)	(.018)	(.017)
No. of Observations	117,288	122,824	139,634	149,172
ρ	0.329	0.42	0.349	0.337
No. of Random Effects (Horses)	349	356	368	430
Lower Bound				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.381***	0.165.	0.327***	0.272***
	(.077)	(.092)	(.074)	(.068)
In-Running	0.309***	0.162***	0.225***	0.227***
	(.025)	(.027)	(.018)	(.017)
No. of Observations	117,288	122,824	139,634	149,172
ρ	0.375	0.463	0.37	0.366
No. of Random Effects (Horses)	349	356	368	430
Place Market (Back Quotes)				
Upper Bound				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.521***	0.457***	0.322***	0.393***
	(.061)	(.064)	(.058)	(.06)
In-Running	1.542***	1.313***	1.367***	1.699***
	(.028)	(.027)	(.017)	(.017)
No. of Observations	116,699	112,634	145,813	156,599
ρ	0.288	0.306	0.288	0.329
No. of Random Effects (Horses)	392	397	397	493
Lower Bound				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.248**	-0.074	0.083	0.082
	(.078)	(.088)	(.064)	(.075)
In-Running	1.033***	0.999***	1.231***	1.515***
	(.025)	(.026)	(.016)	(.017)
No. of Observations	116,699	112,634	145,813	156,599
ρ	0.408	0.467	0.337	0.443
No. of Random Effects (Horses)	392	397	417	493

An analysis of the upper bound and lower bound of the 'in-running' effect identified in Table 2. The upper bound is calculated by assuming that all in-running missing values (where there is no quote) take the value of a round number. The lower bound is calculated by assuming that all in-running missing values (where there is no quote) take a value just below a round number. As before, the round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 and 50. Only quotes at these odds, or those immediately below - plus the missing values - are included in the regressions. Win market data is in the top two panels and place market data is in the bottom two panels. A logit specification is used, random effects for each horse (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and *, **, and . indicates significance at the 0.1%, 1%, and 10% level respectively.**

Table 6: Left-Digit Bias Test Results (Logit)				
Win Market (Back Quotes)				
Dep. Var.: Order Indicator	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	-0.544***	-0.486***	-0.377***	-0.2***
	(.04)	(.041)	(.038)	(.039)
Odds=Round Number	0.14***	0.139***	0.171***	0.166***
	(.014)	(.0144)	(.014)	(.013)
In-Running	-0.836***	-0.699***	-0.667***	-0.848***
	(.045)	(.042)	(.026)	(.026)
Odds=Round Number*In-Running	0.06	-0.137*	-0.102**	-0.064*
	(.055)	(.054)	(.034)	(.032)
No. of Observations	115,861	120,606	137,461	146,264
ρ	0.124	0.129	0.123	0.141
No. of Random Effects (Horses)	342	340	363	421
Place Market (Back Quotes)				
Dep. Var.: Order Indicator	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	-1.957***	-1.867***	-1.903***	-1.671***
	(.039)	(.04)	(.038)	(.039)
Odds=Round Number	0.143***	0.087***	0.125***	0.162***
	(.02)	(.02)	(.019)	(.017)
In-Running	-0.601***	-1.044***	-0.81***	-0.907***
	(.09)	(.099)	(.052)	(.054)
Odds=Round Number*In-Running	-0.33**	0.032	0.011	-0.226***
	(.1)	(.111)	(.059)	(.06)
No. of Observations	114,967	110,030	143,222	153,083
ρ	0.116	0.121	0.125	0.152
No. of Random Effects (Horses)	375	374	409	475

The results of tests for a left-digit bias amongst bettors. An indicator variable equalling 1 if a order was placed in the last second, was regressed on an indicator variable equalling 1 if a round number had been quoted, an indicator equalling 1 if the race was 'in-running' at that time-point, and an interaction between the two aforementioned indicators. The interaction term would capture any left-digit bias that arises specifically during races. As before, the round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 and 50. Only quotes at these odds, or those immediately below, are included in the regressions. Win market data is in the top panel and place market data is in the bottom panel. A logit specification is used, random effects for each horse (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and *, ** and * indicates significance at the 0.1%, 1% and 5% level respectively.**

Table 7: Diff-in-Diff Across Markets (Logit)				
Win & Place Markets (Back Quotes)				
Dep. Var.: Odds=Round Number	Ascot 2011	Ascot 2012	Cheltenham 2011	Cheltenham 2012
Intercept	0.272***	0.712***	0.213***	0.43***
	(.052)	(.056)	(.054)	(.046)
In-Running	0.451***	0.237***	0.283***	0.25***
	(.026)	(.027)	(.018)	(.017)
Place Market	0.349***	-0.741***	0.207***	-0.205***
	(.048)	(.056)	(.062)	(.042)
In-Running*Place Market	0.979***	1.009***	1.046***	1.41***
	(.039)	(.039)	(.025)	(.024)
No. of Observations	230,828	230,636	280,683	299,347
ρ	0.297	0.312	0.289	0.286
No. of Random Effects (bets)	707	706	767	880

A difference-in-difference analysis of the 'in-running effect' across win and place markets. An indicator variable equalling 1 if a round number is quoted, is regressed on an indicator variable equalling 1 if the race is 'in-running' at that time-point, an indicator variable equalling 1 if the market is the place market, and an interaction between the two aforementioned indicators. The round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 and 50. Only quotes at these odds, or those immediately below, are included in the regressions. A logit specification is used, random effects for each horse-market combination (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and * indicates significance at the 0.1% level.**

Table 8: Individual Round Numbers (Logit)							
Win & Place Markets (Back Quotes)							
Dep. Var.: Odds=Round Number	2	3	4	5	6	7	8
Intercept	3.368***	2.442***	2.436***	1.044***	1.435***	0.664***	1.246***
	(.335)	(.167)	(.148)	(.103)	(.091)	(.095)	(.099)
In-Running	1.271***	1.276***	0.56***	1.233***	0.647***	0.46***	0.458***
	(.128)	(.067)	(.058)	(.044)	(.04)	(.04)	(.038)
No. of Observations	10,480	31,280	44,200	66,406	87,624	74,999	78,914
ρ	0.678	0.614	0.726	0.622	0.578	0.631	0.645
No. of Random Effects (Bets)	449	702	803	944	944	891	958
Dep. Var.: Odds=Round Number	9	10	15	20	30	40	50
Intercept	0.699***	1.451***	0.133.	1.182***	0.677***	0.149***	0.966***
	(.082)	(.084)	(.068)	(.079)	(.074)	(.078)	(.102)
In-Running	0.125**	0.315***	0.491***	0.256***	0.132***	0.151***	0.448***
	(.039)	(.038)	(.036)	(.036)	(.033)	(.034)	(.036)
No. of Observations	70,910	98,268	90,771	95,757	107,624	82,719	101,542
ρ	0.553	0.565	0.487	0.547	0.535	0.534	0.663
No. of Random Effects (Bets)	884	1,008	961	1,011	939	872	892

A repeat of the regressions in Table 2, but this time run for each individual round number, and including pooled win and place market data for all four meetings. An indicator variable equalling 1 if a round number is quoted, is regressed on an indicator variable equalling 1 if the race is 'in-running' at that time-point. The round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40 and 50. Only quotes at these odds, or those immediately below, are included in the regressions. A logit specification is used, random effects for each horse (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and *, ** and . indicates significance at the 0.1%, 1%, and 10% level respectively.**

Table 9: Low and High Round Numbers (Logit)		
Win & Place Markets (Back Quotes)		
Dep. Var.: Odds=Round Number	Low Round Numbers	High Round Numbers
Intercept	0.581***	0.358***
	(.03)	(.035)
In-Running	0.985***	0.523***
	(.009)	(.011)
No. of Observations	563,081	478,413
ρ	0.315	0.369
No. of Random Effects (bets)	1,994	1,775

A repeat of the regressions in Table 2, but this time run separately for low and high round numbers, and including pooled win and place market data for all four meetings. An indicator variable equalling 1 if a round number is quoted, is regressed on an indicator variable equalling 1 if the race is 'in-running' at that time-point. The low round numbers are defined as 2, 3, 4, 5, 6, 7, 8, 9, and 10, and the high round numbers are 15, 20, 30, 40 and 50. Only quotes at these odds, or those immediately below, are included in the regressions. A logit specification is used, random effects for each horse (bet) are included (with ρ representing the variation in the dependent variable captured by these random effects), and * indicates significance at the 0.1% level.**

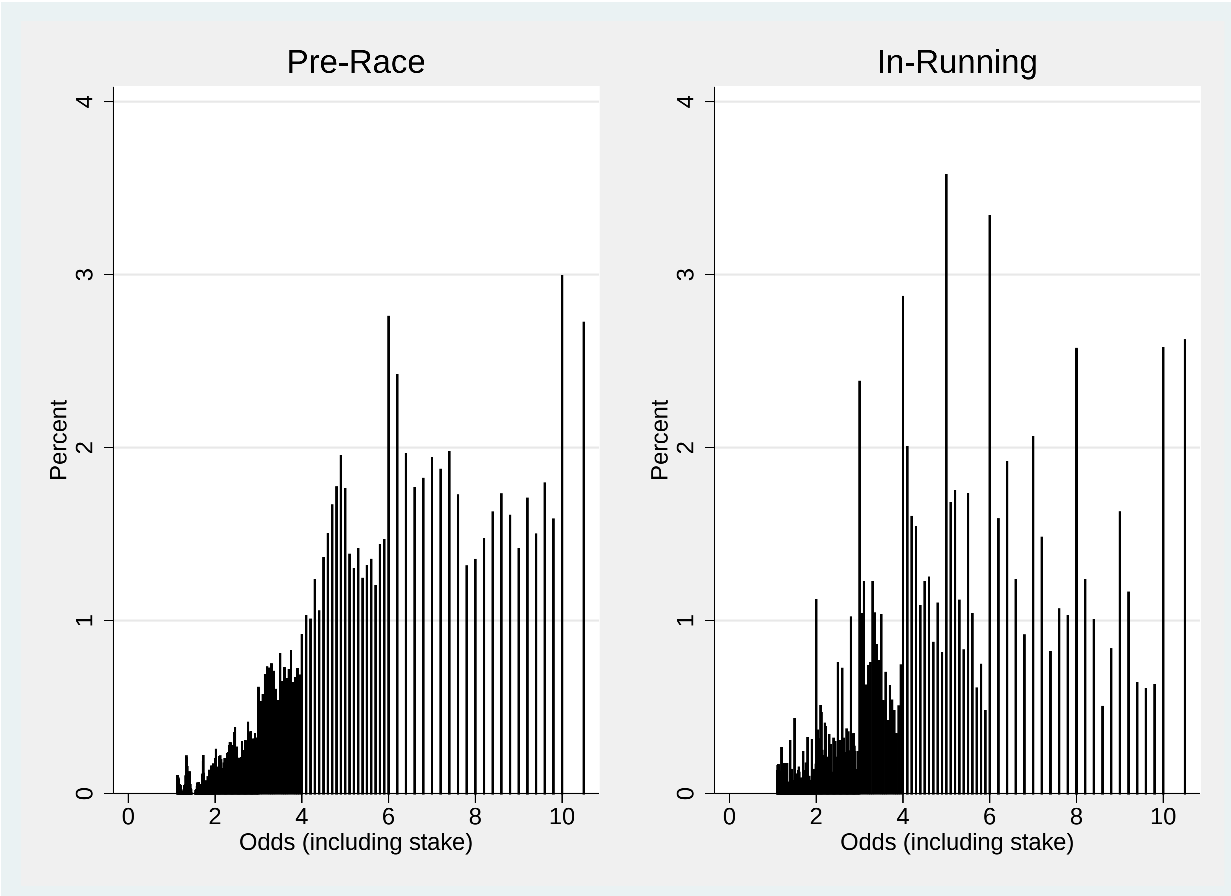
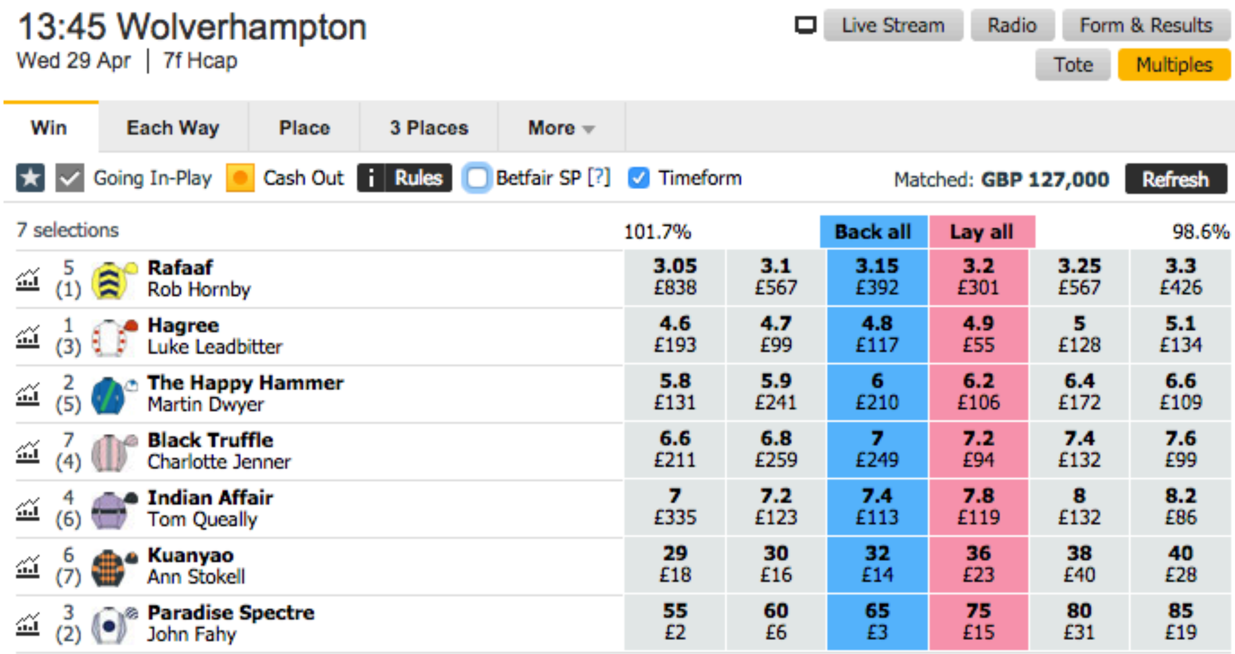


Figure 1: Bar charts of quoted win and place odds (back quotes) on horses in races at **Royal Ascot in 2011**. The left panel contains pre-race odds and the right panel contains in-running odds. The plot is restricted to odds (including the stake) of below 11.


13:45 Wolverhampton
Wed 29 Apr | 7f Hcap

Live Stream | Radio | Form & Results
Tote | Multiples

Win | Each Way | Place | 3 Places | More

Going In-Play | Cash Out | Rules | Betfair SP [?] | Timeform | Matched: GBP 127,000 | Refresh

7 selections	101.7%		Back all	Lay all		98.6%
5 (1) Rafaaf / Rob Hornby	3.05 £838	3.1 £567	3.15 £392	3.2 £301	3.25 £567	3.3 £426
1 (3) Hagree / Luke Leadbitter	4.6 £193	4.7 £99	4.8 £117	4.9 £55	5 £128	5.1 £134
2 (5) The Happy Hammer / Martin Dwyer	5.8 £131	5.9 £241	6 £210	6.2 £106	6.4 £172	6.6 £109
7 (4) Black Truffle / Charlotte Jenner	6.6 £211	6.8 £259	7 £249	7.2 £94	7.4 £132	7.6 £99
4 (6) Indian Affair / Tom Queally	7 £335	7.2 £123	7.4 £113	7.8 £119	8 £132	8.2 £86
6 (7) Kuanyao / Ann Stokell	29 £18	30 £16	32 £14	36 £23	38 £40	40 £28
3 (2) Paradise Spectre / John Fahy	55 £2	60 £6	65 £3	75 £15	80 £31	85 £19

Figure 2: A pre-race screenshot from the Betfair limit order book for the 13.45 race at Wolverhampton on Wednesday 29th April 2015 (an example outside of our sample). To bet on a horse to win, the punter would place a 'back' bet on the left, and to bet against a horse the punter would place a 'lay' bet on the right. The best back and lay quotes are displayed, for each horse, in blue and pink respectively. The volumes available at each price are indicated underneath the odds.

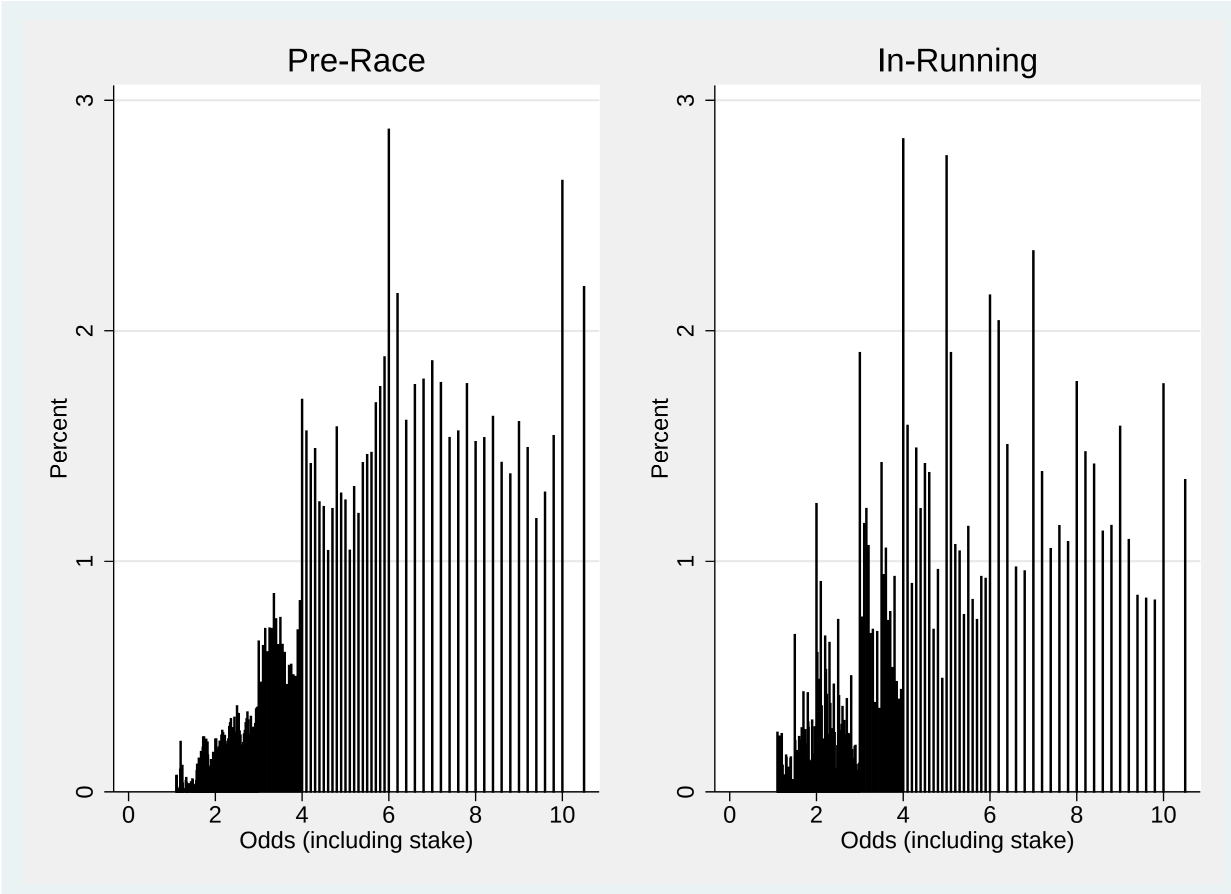


Figure 3: Bar charts of quoted win and place odds (back quotes) on horses in races at **Royal Ascot in 2012**. The left panel contains pre-race odds and the right panel contains in-running odds. The plot is restricted to odds (including the stake) of below 11.

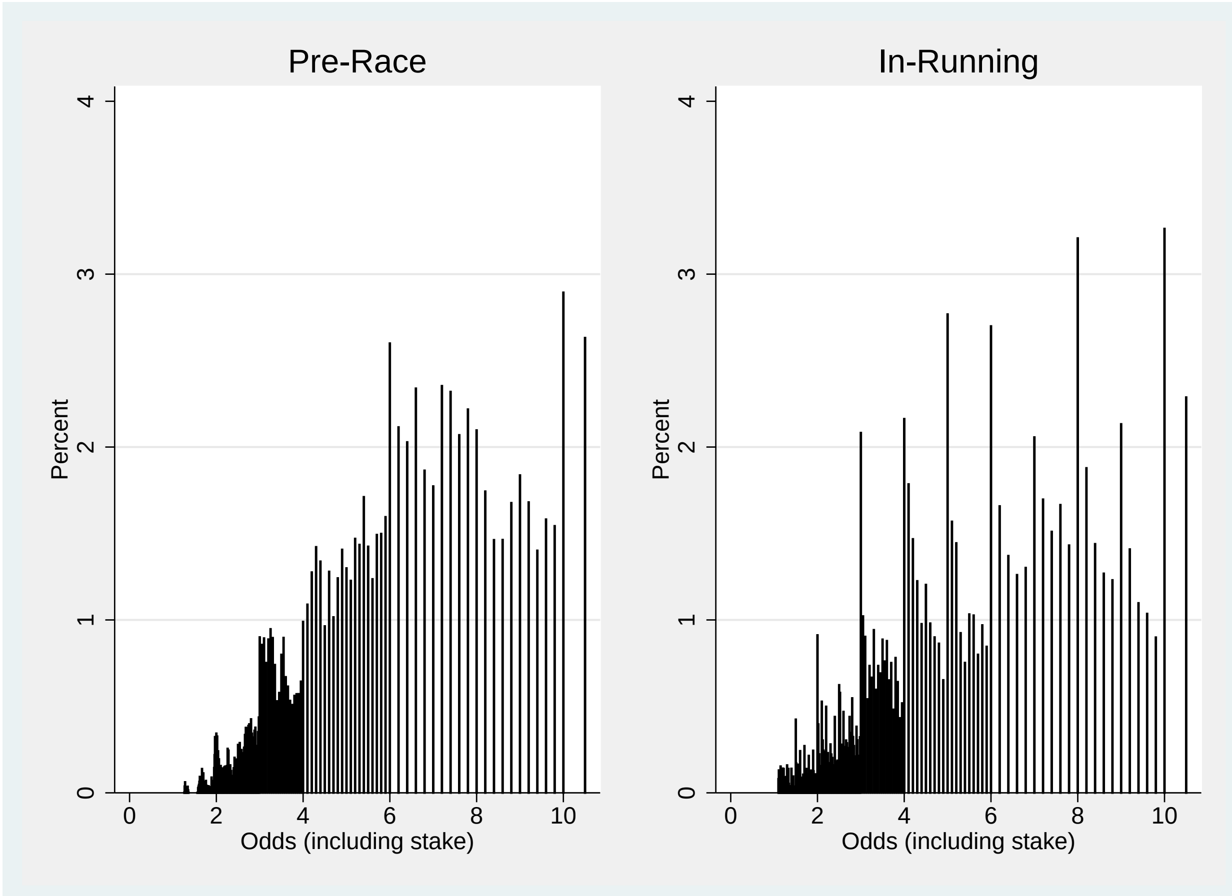


Figure 4: Bar charts of quoted win and place odds (back quotes) on horses in races at the **Cheltenham Festival in 2011**. The left panel contains pre-race odds and the right panel contains in-running odds. The plot is restricted to odds (including the stake) of below 11.

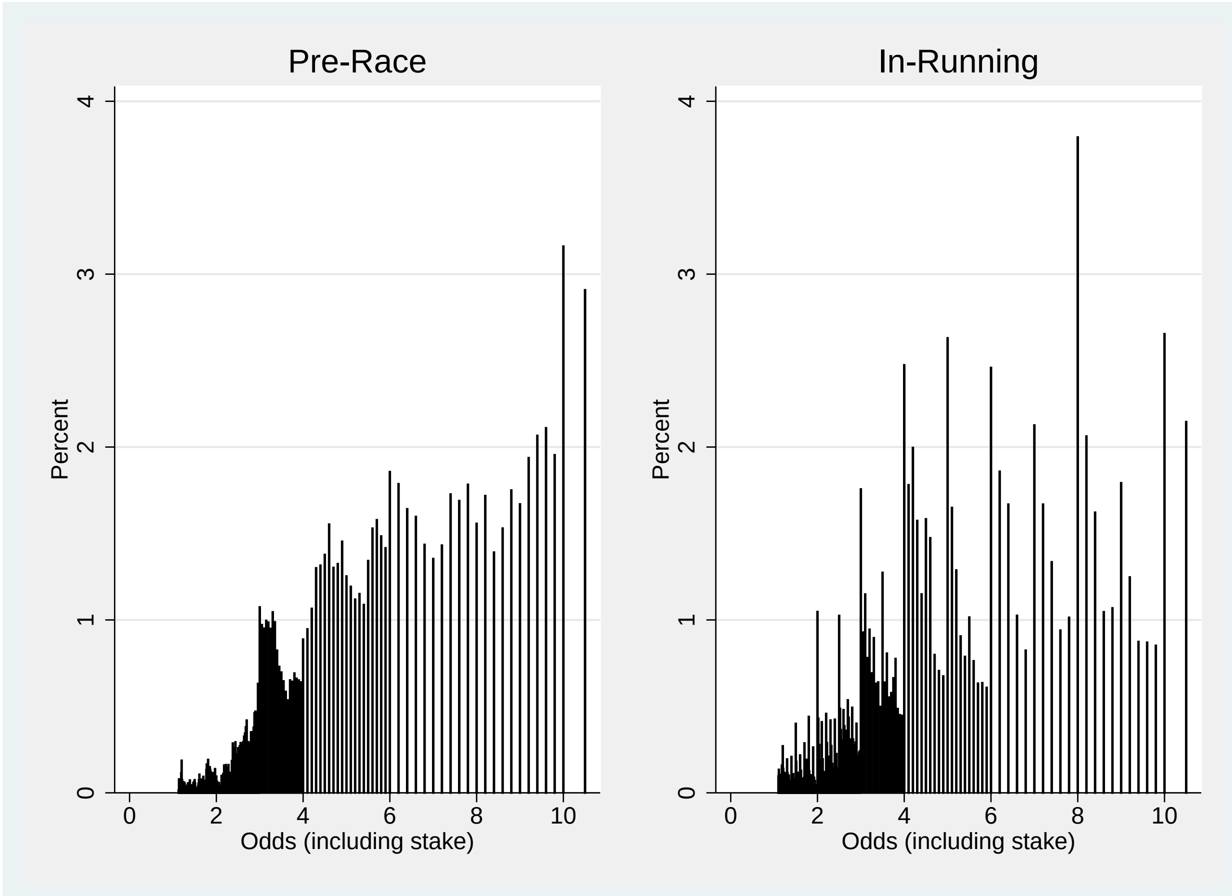


Figure 5: Bar charts of quoted win and place odds (back quotes) on horses in races at the **Cheltenham Festival in 2012**. The left panel contains pre-race odds and the right panel contains in-running odds. The plot is restricted to odds (including the stake) of below 11.