INEFFICIENCIES IN PARIMUTUEL BETTING MARKETS ACROSS WAGERING POOLS IN THE SIMULCAST ERA

MARSHALL GRAMM[1]
Rhodes College

DOUGLAS H. OWENS
Texas A&M University

ABSTRACT

Simulcast wagering, where bets from across the country are taken at tracks, off-track betting facilities, casinos, by phone or online and incorporated into the same mutuel pool, has contributed to a large increase in betting volume on American horse races since the mid-1990s. This paper investigates betting market efficiency in the simulcast era focusing on whether the interrelated betting markets comprised of win, place (finishing in the top two), and show (finishing in the top three) wagering are efficiently priced. We find that the increased accessibility and betting volume associated with simulcasting has reduced, but not eliminated, the inefficiencies seen in prior studies. Despite the inefficiencies in these markets, arbitrage is not profitable since market closing prices are unknown when bets are placed.

JEL Classification Code: G14
Keywords: Market Efficiency, Parimutuel Betting, Favorite-Longshot Bias

Acknowledgements
The authors would like to thank Ramon P. DeGennaro, Robert B. Ekelund, Teresa Beckham Gramm, and Wayne Strayer for comments on the paper as well as participants at the 3rd International Equine Industry Program Academic Conference. Amanda Godbold helped with data entry.

[1] Corresponding Author. Address: Department of Economics and Business, Rhodes College, 2000 North Parkway, Memphis, TN 38103. Phone: (901) 843-3122 (office), (901) 843-3736 (fax). Email: gramm@rhodes.edu.

INEFFICIENCIES IN PARIMUTUEL BETTING MARKETS ACROSS WAGERING POOLS IN THE SIMULCAST ERA

INTRODUCTION

Parimutuel wagering has been much studied in economics and finance because it functions as a controlled repeated experiment of an asset market (see Sauer (1998) for an overview). Through parimutuel betting the public collectively establishes a price on each betting interest, and these prices have been found to be fairly accurate in representing the true value of the bet. The track acts as a market maker, extracting a fixed percentage (14-20%) from betting pools and redistributing the rest to the holders of the winning tickets. Because the market is repeated numerous times daily at tracks across the world, an abundance of data exists on betting markets. Furthermore, with the proliferation of simulcasting races, participation in the parimutuel market is no longer restricted to just those attending the races.

This paper is an empirical analysis of straight wagers, which are bets on a horse to win, place (finish in the top two), or show (finish in the top three). Most previous research on betting market efficiency has focused on win wagering and has demonstrated that racetrack betting markets are not efficient (see Thaler and Ziemba (1988), Vaughan Williams (1999) for an overview). Numerous empirical studies have found the existence of a bias such that favorites were underbet and longshots were overbet resulting in a higher expected return for low odds horses (Ali (1977), Asch, Markiel, and Quandt (1982)). However, other studies have found a reverse favorite-longshot bias (Busche and Hall (1988), Swindler and Shaw (1995)). Explanations of the inefficiency have included risk preference (Ali(1977), Golec and Tamarkin(1998)), information disparities (Hurley and McDonough (1995, 1996), Terrell and Farmer(1996), Gander, Zuber, and Johnson (2001)), transaction costs (Vaughan Williams and Paton (1998a, b), and market size (Busche and Walls (2000)). Previous studies on place and show betting have found even more pronounced biases and these findings have led to the formulation of profitable betting strategies, the most prominent being Ziemba and Hausch's "Beat the Racetrack"

(Ziemba and Hausch (1984), Asch, Malkiel, and Quandt (1984, 1986), Asch and Quandt (1986), Hausch, Ziemba, and Rubinstein (1981), Hausch and Ziemba (1985)).

The proliferation of simulcast wagering has created an environment where few betting patrons attend the races anymore, and those that do are more likely to be found in front of a television carrel watching races from around the country, rather than in the grandstand. Previously tracks would simulcast only major races a few times a year and have their own separate betting pools for these races. A betting pool at a given track for a given race would be comprised of money from people at the track and in some instances from off-track betting sites or phone accounts, both within the track's home state. Today, simulcast wagering allows bettors to play a multitude of races at many tracks across the country from their home track, casino, off-track betting hub, by phone, or online and their bets are co-mingled into the same pool as those made at the host track. This development has resulted in an explosion in the dollar volume wagered on horse racing in the last decade. From 1985 to 2002 the total wagered on thoroughbred races in North America increased from \$8.25 billion to \$15.62 billion despite the fact that the number of races fell to 59,896 from 75,687. Per race wagering more than doubled over the seventeen year period, increasing from \$109,000 to \$260,000. Adjusting for inflation total wagering increased by 21%, while per race wagering increased by 53%. Much of this can be attributed to off-track betting which accounted for 86% of all bets made in 2002.

This paper is the first comprehensive study of straight wagers since wagering pools began to be co-mingled in the mid-1990s. The fact that betting markets are accessible to horseplayers across the country should result in more efficient pricing both within and across wagering pools. We use a large dataset, consisting of all tracks available to subscribers of TVG network's online racing service, to test whether the interrelated markets of win, place, and show wagering are efficiently priced. All major racetracks are included. Despite increased participation, we find that a favorite-longshot bias still exists in each pool, with the bias being more severe in place and show wagering. Place and show bets on extreme favorites earned a positive return. With evidence of inefficiency, an experiment to arbitrage betting markets was attempted but found to be unprofitable.

The method used to try to arbitrage interrelated betting markets was the Dr. Z system, popularized in the classic book "Beat the Racetrack." Ziemba and Hausch's strategy to profit from betting market inefficiencies involved using Harville (1973) formulas. Despite the fact that there typically exists a small favorite-longshot bias in win betting, probabilities determined by the bettors, or subjective probabilities, can be a good approximation of a horse's actual probability of winning. Harville introduced methods to calculate the probability of finishing in the top two or three based upon the win probabilities of the horses in a race. A good place (show) bet may exist when a horse's probability of finishing in the top two (three) is much greater than the amount bet on the horse in the place (show) pool would justify. Unfortunately for the arbitrageur, the amount bet on each horse is typically not fully known until after the race has begun, making applications of the Dr. Z system a challenge at the racetrack. While arbitrageurs may wait until the last minute to have the best projection of the final odds, they run the risk of getting shut out at the betting window. These last frantic minutes often involve making quick but somewhat complex calculations to determine how much to bet on which horse or horses.

The analysis also shows that, based on the final odds, arbitraging these betting markets using the Dr. Z system could be profitable. However, since our dataset contains final pool totals, the question still remained as to whether profits could be made in an actual betting scenario. To test this in practice, we placed bets on 203 races from February through April of 2003. Our bets fit the criteria of a Dr. Z system and were made online at the last possible moment that betting was allowed. Even so, only about 60 percent of the final pool totals are recorded when the betting windows close. The results show that the market becomes more efficient in the minutes leading up to the race, meaning that profitable bets at post time become poor plays once the final pool totals are revealed. Overall, the experiment resulted in a small net loss.

EMPIRICAL RESULTS

DATA OVERVIEW

The authors have collected a comprehensive data set including all races available to the TVG network on-line subscribers from October 9th to December 31st of 2002. This includes 96,275 betting interests[2] in 11,361 races over 84 days at 36 racetracks. All major tracks are included. Of the 36 tracks studied, 23 hosted thoroughbred racing, 10 harness, and 3 were mixed (including thoroughbreds, quarterhorses, arabians, and even mules). Both the overall size of the dataset and the number of racetracks included make it one of the largest to be used in a betting market efficiency study. Table 1 summarizes the dataset by race meet.

<Table 1: Racetracks 10/02-12/02>

The number of horses and races are included in the table, with an overall average of 8.47 betting interests per race. The track take varies from a low of 14% at the New York tracks to 20.5% at Pompano Park in Florida. Pool Size is the average total bet on straight wagers per race. With more than $400,000 per race, Arlington Park has the highest average bet, mainly due to hosting the Breeders' Cup World Thoroughbred Championships. Prairie Meadows, one of the small tracks in the study, had just over $1,000 bet per race during their Harness meet. Average purse size gives an indication of how important the track is, and once again Arlington Park ranks at the top due to the $13 million in purses at the Breeders' Cup.

FAVORITE LONGSHOT BIAS

The favorite-longshot bias can be detected by grouping horses by favorite position and comparing the subjective probability to the objective probability. The subjective probabilities are what the bettors in aggregate feel the horses' chances are, as revealed by the odds. Objective probabilities, on the other hand, are defined as the actual percentage of winners in the group. A significant difference between subjective and objective probability for a group indicates mispricing and market inefficiency. The total amount bet to win on all horses in a race can be expressed as W, with w denoting the amount bet to

[2] Generally each horse in a race is a separate betting interest. However in some cases when horses have the same owners or trainer they are grouped together as one betting interest and are effectively treated as one horse in wagering. References to horses in this paper are actually to betting interests and coupled entries are treated as one horse.

win on an individual horse, so that $\sum_{i=1}^{n} w_i = W$ where i indexes the n individual horses in a race. The odds on a horse to win are equal to $\frac{(1-t)W}{w_i} - 1$, where t is the track take. The odds are updated every minute and payouts are based on the odds when the pools close (when the horses start running and thus the tellers stop taking bets). A horse's subjective probability of winning is $\psi = \frac{w_i}{W} = \frac{1-t}{Odds_i + 1}$. The return on a \$1 win bet is $\frac{(1-t)W - w_i}{w_i} = Odds_i$ if horse i wins and -1 otherwise. The objective probability, ζ, is the percentage of winners in each observed group. To determine whether there is a significant difference between the objective and subjective probabilities for a given group the number of wins can be viewed as a binomial statistic. For a sample of n horses a z-statistic can be computed as $z = (\psi - \zeta)\sqrt{n/\zeta(1-\zeta)}$ (see Busche and Walls (2001)). Z-statistics that are significantly different from zero provide evidence of inefficiency. A positive (negative) z-score indicates that a group is overbet (underbet) relative to its true probability.

For this analysis, subjective probabilities for place and show wagers are calculated using the Harville formulas:

$$\text{Probability that } i \text{ is first and } j \text{ is second} = \frac{q_i q_j}{(1-q_i)} \tag{1}$$

$$\text{Probability that } i \text{ is first, } j \text{ is second, and } k \text{ is third} = \frac{q_i q_j q_k}{(1-q_i)(1-q_i-q_j)} \tag{2}$$

where q represents the probability that the horse wins the race. Summing all the probabilities involving a horse either finishing first or second will yield its probability of placing, and summing the probabilities for finishing first, second, or third will yield the probability of showing. However, using subjective win probabilities for q fails to take into account what Hausch, Ziemba, and Rubinstein dubbed the "Silky Sullivan" problem after the great western closer. Silky Sullivan and horses of his ilk were all or nothing, they either won or finished out of the money. Therefore Harville formulas overestimate these horses' probabilities of placing and showing. There are other horses (western handicap horse Grey Memo comes to mind) that finish second and third on many occasions but rarely visit the winners circle. In those instances the probability that the horse placed or showed would be underestimated. Therefore, we use an adjusted version of the Harville formulas in this study. For place wagers, q_i is estimated by p_i/P, where p_i is the amount bet on horse *i* to place and *P* is the total amount wagered in the place pool. For show wagers, q_i is estimated by s_i/S where s_i is the amount bet on horse *i* to show and *S* is the total amount wagered in the place show. These adjustments allow the place and show subjective probabilities to reflect the bettors' intentions by including the amount bet in the place and show pools, as opposed to constructing them from subjective win probabilities estimated from the win pool. Since we would like to look at inefficiencies across betting pools, it is preferable to isolate all calculations involving a horse's probability of finishing in the top two to the place pool (and likewise all calculations involving a horse's probability of finishing in the top three to the show pool).

The return on a place bet depends on whether horse *i* finishes in the top two and which other horse finishes in the top two with it. The return on a \$1 place bet if horse *i* finishes in the top two with horse *j* is $\frac{(1-t)P - p_i - p_j}{2p_i}$. Similarly for show wagering, the return on a \$1 show bet if horse *i* finishes in the top three with horses *j* and *k* is $\frac{(1-t)S - s_i - s_j - s_k}{3s_i}$. Thus, while the odds that a horse will win the race are publicly available, the public does not know the probable payoff of place and show wagers. The public is able to view how much is bet on each horse in place and show pools but not probable payoffs since the probable payoffs are determined in part by who the other top two or three finishers. The more money bet on horse *j* to place reduces the place payoff on horse *i* if horses *i* and *j* are the top two finishers. Likewise, the more money bet on horses j and k to show reduces the show payoff on horse *i* if horses *i*, *j*, and *k* are the top three finishers.

Establishing the existence and the direction of a favorite-longshot bias involve comparisons of the subjective and objective probabilities between groups of horses. One method of grouping involves ranking the horses in each race from most favored (lowest odds) to least favored (highest odds). The horses are divided into nine groups by their favorite position in the race from 1 (most favored, lowest odds) to 9-14 (least favored, odds rankings of 9th and above). The 9th through 14th favorites were combined because of the (relatively) small number of observations. The results are summarized in Table 2.

<Table 2: Data Grouped by Favorite Position>

Note that fewer horses could be bet on in the place and show pools because some races with small fields do not allow show betting, and in rare instances do not allow place betting. Differences in the size of the groups is due to variation in the number of horses in each race and because horses with the same odds were given the same odds ranking.

The column labeled "Raw" in Table 2 is the raw return from betting all horses in the odds grouping not accounting for any takeout, i. e. if the track returned 100 percent of all pools. The take and breakage return column is the actual payout to the bettor accounting for the track take (typically 14-20 percent) and any breakage (rounding payouts down to the nearest nickel or dime). In win, place, and show bets the standard favorite-longshot bias was evident. The difference in returns between the lowest and highest odds horse was much greater in the place (-8% to -38%) and show pool (-7½% to -42½%) than in the win pool (-16½% to -24%). The differences between objective probability and subjective probability were significant in three positions for win wagers, six positions for place wagers, and seven positions for show wagers. To jointly test the difference in actual and expected returns across all odds groupings we use a chi-square test equal to the sum of the squared z-scores from each odds grouping. The statistic is 31.70 for win bets, 87.61 for place bets, and 185.52 for show bets, each greater than the 1% critical value of 21.67. Thus, it can be concluded that the place and show pools are less efficient than the win pool. Even so, strictly betting favorites to place or show will result in a negative return.

If the Harville formulas are correct, and win, place, and show wagers are equally efficient, then the percentage bet on a particular horse should be the same across win, place, and show wagers. As shown in Table 3, this is clearly not the case.

<Table 3: Breakdown of Wagering Pools by Favorite Position>

Of all the money bet on race favorites, 68.1% is to win, 21.5% to place, and 10.4% to show. Moving lower in the odds ranking, there is less bet to win as a percentage (down to 56.2%) and more bet to place (up to 25.8%) and show (up to 18.0%). When people bet longshots they tend to back them in the place and show pools while favorites are backed more heavily in the win pool. This is further demonstrated by the percentage of the win, place, and show pools bet on each horse. 34.7% of all win bets are on the race favorites, while only 30.7% of the place bets and 29.6% of the show bets are on race favorites. The least favorite horses receive on 1.9% of all money bet in the win pool, but 2.3% of the place pool and 3.1% of the show pool. These results are strong evidence of inefficiencies across the three wagering pools.

There have been a number of betting market studies conducted over the years and it is insightful to compare efficiency under different conditions. To compare betting market efficiency between different datasets, a simple regression of the subjective probability for each favorite grouping can be regressed on the objective probability as follows:

$$Subjective_i = \beta_0 + \beta_1 Objective_i + u_i \tag{3}$$

where *i* indexes favorite position groups. If $\beta_0 = 0$ and $\beta_1 = 1$ then the market is efficient. A standard favorite-longshot bias exists when $\beta_1 < 1$ and a reverse favorite-longshot bias when $\beta_1 > 1$. Regression results for win wagers for this data as well as five previous studies are shown in Table 4.

<Table 4: Comparisons of Studies of Betting Markets>

Chronologically, Ali (1977) looked at harness races in New York, Asch, Malkiel, and Quandt (1982) at horse races in New Jersey, Busche and Hall (1988) at horse races from Hong Kong, Gandar, Zuber, and Johnson (2001) at horse races in New Zealand, and Sobel and Raines (2003) at dog races in West Virginia. Each dataset involved parimutuel wagering and grouped participants by favorite position. The t-statistics given are for tests against a null where $\beta_0 = 0$ and $\beta_1 = 1$ respectively. In every study the intercept was not significantly different from zero. However only the two studies focusing on foreign racing (Busche and Hall, Gandar et al.) had slope coefficients which were not significantly different from one. Both Hong Kong and New Zealand racing have characteristics which differentiate them from American racing. In Hong Kong, the amount wagered per race is much bigger than in the U.S. Busche and Hall point out that per race handle in Hong Kong when their study was undertaken was $5.6 million as compared to only $87,000 for the U.S. (Busche and Hall (1998 pp. 339). New Zealand racing is overseen by the Totalisator Agency Board which, in the duration of Gandar et al.'s study, had a large national off-track presence including telephone accounts and wagering hubs in retail stores and pubs. 90% of the betting volume on New Zealand racing was done off-track (Gandar et al (2001) pp. 1622). The American studies either predate simulcasting (Ali, Asch et al.) or have a small percentage of betting volume generated off-track (Sobel and Raines). Each of these studies reveal inefficient betting markets, with Sobel and Raines' greyhound races having a reverse favorite-longshot bias. Our data exhibits the standard favorite-longshot bias despite the increased bettor participation through simulcasting. Looking at the magnitude of the slope coefficient, domestic betting markets have become only slightly more efficient over time despite increased bettor participation through simulcasting.

ARBITRAGE

Given that inefficiencies exist in the place and show pools, can a profitable wagering rule be established? This was a question that was addressed by Ziemba and Hausch in their 1984 book "Beat the Racetrack" (also see Hausch, Ziemba, and Rubinstein (1981), Ziemba and Hausch (1984), and Hausch and Ziemba (1985)). Dr. Z's system, as it came to be known, involved calculating the expected return to place and show based upon the amounts wagered on a horse in the three betting pools.

$$E(RET_{PLACE}) \approx 0.319 + 0.559\frac{w_i/W}{p_i/P} + \left(2.22 - 1.29\frac{w_i}{W}\right)(1 - t - 0.829) \quad (4)$$

$$E(RET_{SHOW}) \approx 0.543 + 0.369\frac{w_i/W}{s_i/S} + \left(3.60 - 2.13\frac{w_i}{W}\right)(1 - t - 0.829) \qquad (5)$$

Formula (4) and (5), Ziemba and Hausch's empirical estimates of the expected return for place and show wagers, are used to initially screen for horses who might be underbet in the place and show pools. If the expected return to place (show) on a horse is 1.15, then a place (show) bet will earn a predicted 15% return. Dr. Z's betting strategy involves betting horses to place or show if their expected return is above a minimum criterion and if these horses are not longshots. Ignoring any horse going to the post at greater than 8-1 odds and using 1.15 as the minimum expected return, the methodology yields a 14.87% straight profit using our data. Wagering opportunities were sparse, with only 3.03% of all betting interests exhibiting an expected return above the set threshold but below the odds cut off. Even so, this would imply a few bets a day at any given racetrack, just as indicated in "Beat the Racetrack." While the profitability of the system may seem enticing, it is important to remember that we are looking at returns on bets using final pool totals. The pool totals and odds are updated every minute and continue to be updated even after the horses leave the starting gates and the betting windows have closed. No one has the benefit of applying a system to the final pool totals so we tested the profitability of the Dr. Z system in real time.

Despite the positive return of these mythical bets, the greatest difficulty with the Dr. Z system is in its implementation. Bettors have to watch the tote board and make calculations while trying not to get shut out at the betting window. With the evolution of online wagering, monitoring pool totals and making calculations using Dr. Z's formulas are much easier.

For the purposes of evaluation, Dr. Z system bets were made on 203 races at 34 tracks in February through April of 2003. A $2 wager was made on horses with an expected return on a place (show) wager was at or above 1.15 and with win odds less than or equal to 8-1.[3] Furthermore any races likely to create a minus pool, where so much money is bet on one horse that tracks pay the minimum 5% and lose money on the race,

[3] Dr. Z's system advocated using the Kelly Criterion (maximizing expected log wealth) to determine bet size. This replication did not do so.

were not considered (see Chapter 15 of Ziemba and Hausch (1984)). Betting was postponed until the last possible moment but could be done quickly through the author's online wagering account. Expected Returns from equations (4) and (5) could be found quickly with a computer.

Bets were made at the last possible moment (save a few instances when the author was shut out) to get the closest approximation of the final odds and expected returns. Unfortunately, on average only 57% of the final pool totals are viewable on the tote board when betting on a race ends. Much of the betting occurs in the last few minutes and posted pool totals change when betting is closed. Overall, 319 $2 bets were made on 203 races. There were 113 place wagers and 206 show wagers and the net result was a $91.80 loss (-12.8%). While these bets looked attractive when made, the late money often lowered their expected return below 1.00. A $45.70 profit (7.2%) would have been made had we received the payouts based upon pool totals when the wager was made. Only 90 of the original 319 wagers meet the criterion both at the time of wager and in the final total. These bets returned $9.30 or 5.2%. Using final pool totals, 134 wagers (43 place and 91 show) meet the criteria and returning $31.40 or 11.8%. The data from the arbitrage experiment exhibited the same favorite-longshot bias found in previous studies when final pool totals are examined. Thus, it is likely that the negative returns were not an aberration.

<Table 5: Betting Application Data>

We find that late money eliminates profit from arbitraging bias between the betting pools. Table 5 shows the results of analyses that lead to this conclusion. Two subjective probabilities were calculated, one with final pool totals and another using the pool totals when wagers were made or the Post Time (PT) pool totals. In each case late money flows to the favorites in all pools. The subjective probability increases for the top two favorite positions in all pools, including a five percentage point increase for place and show wagers. The subjective probability falls for positions 4 and higher in the win and place pools and 5 and higher in the show pool. These changes in probabilities demonstrate that late money shifts the odds towards the true win probabilities. The increased efficiency reduces the number of optimal bets from 319 to 134. Figure 1 contains graphs of the estimated return line for each betting pool by favorite position. The three lines designate estimated returns using the TVG data (solid line) and data from the betting experiment, both post time (dotted line) and final (dashed and dotted line) pool totals. All estimated return lines are downward sloping with the TVG data being the most flat. In each pool the estimated return flattens as we go from post time to final pool totals indicating that the market becomes more efficient.

<Figure 1: Estimated Returns>

CONCLUSION

This paper finds that despite increased accessibility and participation due to the proliferation of simulcast wagering, betting markets continue to inefficiently price outcomes, a result that holds across wagering pools. The win pool exhibits a favorite longshot bias where favorites (longshots) tend to be underbet (overbet) relative to their true probability of winning. The size of the bias is smaller than previous studies of American racing but much greater than foreign countries where simulcast wagering is prevalent. Expanding the analysis to the place and show pools, we find an even more pronounced bias continues to exist.

Data used in the study was comprised of a large number of races over numerous racetracks and included nearly all races simulcast in the Fall of 2002. The variation in return was much larger for place and show wagers. Based on final betting pool totals a small positive profit could be earned betting extreme favorites (odds on) to place and show. However betting extreme longshots (40-1 or greater) would result in a 34 percent loss on win wagers, 46 percent loss on place wagers and 50 percent loss on show wagers.

Despite strong evidence of greater inefficiencies in the place and show pools, methods used to arbitrage betting markets resulted in a net loss. Using a modification of the Dr. Z system described in “Beat the Racetrack” resulted in a net loss of 13 percent from bets on 319 horses in 203 races in the winter of 2003. This was despite the fact that it was generally possible through online wagering to make bets in the last seconds before the races began. A positive expected return at post time disappeared as late money reduced or eliminated the inefficiencies that had appeared exploitable.

REFERENCES

Ali, M. M. (1977). Probability and Utility Estimates for Racetrack Bettors. Journal of Political Economy, 85, 803-815 .

Asch, P., Malkiel, B. G., & Quandt, R. E. (1982). Racetrack Betting and Informed Behavior. Journal of Financial Economics, 10, 187-194.

Asch, P., Malkiel, B. G., & Quandt, R. E. (1984). Market Efficiency in Racetrack Betting. Journal of Business, 57, 165-175.

Asch, P., Malkiel, B. G., & Quandt, R. E. (1986). Market Efficiency in Racetrack Betting: Further Evidence and a Correction. Journal of Business, 59, 157-160.

Asch, P. & Quandt, R. E. (1986). Racetrack Betting: The Professors' Guide to Strategies. Dover: Auburn House.

Busche, K. & Hall, C. D. (1988). An Exception to the Risk Preference Anomaly. Journal of Business, 61, 337-346.

Busche, K. & Walls, W. D. (2000). Decision Cost and Betting Market Efficiency. Rationality and Society, 12, 477-492.

Busche, K. & Walls, W. D. (2001). Breakage and Betting Market Efficiency: Evidence from the Horse Track. Applied Economics Letters, 8, 601-604.

Davidowitz, S. (2003). The American Racing Manual 2003. New York: DRF Press.

Gander, J. M., Zuber R. A., & Johnson, R. S. (2001). Searching for the Favourite-Longshot Bias Down Under: An Examination of the New Zealand Pari-mutuel Betting Market. Applied Economics, 33, 1621-1629.

Golec, J. & Tamarkin, M. (1998). Bettors Love Skewness, not Risk, at the Horse Track. Journal of Political Economy, 106, 205-225.

Harville, D. A. (1974). Assigning Probabilities to the Outcomes of Multi-entry Competitions. Journal of the American Statistical Association, 69, 446-452.

Hausch, D. B., Ziemba, W. T., & Rubinstein, M. (1981). Efficiency of the Market for Racetrack Betting. Management Science, 27, 1435-1452.

Hausch, D. B. & Ziemba, W. T. (1985). Transactions Costs, Extent of Inefficiencies, Entries and Multiple Wagers in a Racetrack Betting Model. Management Science, 31, 381-394.

Hurley, W. & McDonough, L. (1995). A Note on the Hayek Hypothesis and the Favorite-Longshot Bias in Parimutuel Betting. American Economic Review, 85, 949-955.

Hurley, W. & McDonough, L. (1996). The Favourite-Longshot Bias in Parimutuel Betting: A Clarification of the Explanation That Bettors Like to Bet Longshots. Economics Letters, 52, 275-78.

Sauer, R. D. (1998). The Economics of Wagering Markets. Journal of Economic Literature, 36, 2021-2064.

Sobel, R. S. & Raines, S. T. (2003). An Examination of the Empirical Derivatives of the Favourite-Longshot Bias in Racetrack Betting. Applied Economics, 35, 371-385.

Swindler, S. & Shaw, R. (1995) Racetrack Wagering and the Uninformed Bettor: A Study of Market Efficiency. Quarterly Review of Economics and Finance, 35, 305-14.

Terrell, D. & Farmer, A. (1996). Optimal Betting and Efficiency in Parimutuel Betting Markets with Information Costs. Economic Journal, 106, 846-868.

Thaler, R. H. & Ziemba, W. T. (1998). Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries. The Journal of Economic Perspectives, 2, 161-174.

Vaughan Williams, L. & Paton, D. (1998). Why are some Favorite-Longshot Biases Positive and some Negative? Applied Economics, 30, 1505-1510.

Vaughan Williams, L. & Paton, D. (1998). Do Betting Cost Explain Betting Biases? Applied Economics Letters, 5, 333-335.

Vaughan Williams, L. (1999). Information Efficiency in Betting Markets: A Survey. Bulletin of Economic Research, 51, 307-337.

Ziemba, W. T. & Hausch, D. B. (1984). Beat the Racetrack. San Diego: Harcourt, Brace, and Jovanovich.

Table 1: Racetracks (10/02-12/02)

Track	State	Type	Take	Horses	Races	Pool Size	Purse
Aqueduct	NY	Thoroughbred	14.0%	3,447	403	$321,428	$45,022
Arlington	IL	Thoroughbred	17.0%	1,103	137	$404,221	$122,000
Balmoral	IL	Harness	17.0%	4,587	493	$37,255	$7,762
Belmont	NY	Thoroughbred	14.0%	660	84	$356,853	$51,542
Beulah	OH	Thoroughbred	18.0%	6,322	665	$25,033	$6,800
Calder	FL	Thoroughbred	18.0%	5,233	641	$120,388	$23,845
Churchill Downs	KY	Thoroughbred	16.0%	2,280	244	$232,216	$39,598
Colonial Downs	VA	Harness	18.0%	1,452	190	$3,451	$6,752
Delta Downs	LA	Thoroughbred	17.0%	2,312	258	$26,927	$17,428
Dover Downs	DE	Harness	18.0%	4,391	545	$13,177	$14,676
Fair Grounds	LA	Thoroughbred	17.0%	1,401	160	$151,360	$27,322
Fairmount Park	IL	Thoroughbred	17.0%	458	60	$10,604	$6,667
Fresno Fair	CA	Mixed	16.8%	373	49	$43,207	$8,088
Great Lakes Downs	MI	Thoroughbred	17.0%	832	106	$13,977	$9,703
Harrington Raceway	DE	Harness	18.0%	1,797	225	$10,207	$11,811
Hollywood Park	CA	Thoroughbred	15.43%	2,175	296	$289,380	$42,233
Hoosier	IN	Thoroughbred	18.0%	3,786	409	$38,787	$14,902
Keeneland	KY	Thoroughbred	16.0%	622	70	$250,352	$44,800
Laurel Park	MD	Thoroughbred	18.0%	4,069	511	$72,429	$21,296
Lone Star Park	TX	Mixed	18.0%	2,751	304	$21,657	$15,771
Los Alamitos	CA	Mixed	15.6%	3,193	433	$26,589	$16,057
Louisiana Downs	LA	Thoroughbred	17.0%	1,722	206	$62,207	$11,388
Maywood	IL	Harness	17.0%	3,067	389	$32,565	$8,851
Monticello Raceway	NY	Harness	18.0%	4,011	542	$9,627	$2,092
Moutaineer	WV	Thoroughbred	17.3%	3,788	405	$41,689	$18,076
Northfield	OH	Harness	18.0%	5,388	633	$22,891	$4,368
Oak Tree	CA	Thoroughbred	15.43%	1,419	168	$308,508	$42,970
Pompano	FL	Harness	20.5%	3,688	457	$9,860	$5,348
Prairie Meadows	IA	Harness	18.0%	1,284	166	$1,313	$2,692
Retama	TX	Thoroughbred	18.0%	1,085	123	$36,327	$11,521
Sam Houston	TX	Thoroughbred	18.0%	2,896	319	$60,005	$18,002
Saratoga Harness	NY	Harness	18.0%	2,079	270	$5,532	$2,244
Suffolk Downs	MA	Thoroughbred	19.0%	3,361	373	$40,356	$14,314
Sunland Park	NM	Thoroughbred	19.0%	2,519	264	$12,747	$21,052
Turf Paradise	AZ	Thoroughbred	20.0%	4,588	547	$37,604	$7,521
Turfway Park	KY	Thoroughbred	17.5%	2,136	216	$76,850	$14,240

Balmoral, Fairmount Park, and Maywood all charge a 1% surtax on winning tickets.

Table 2: Data Grouped by Favorite Position

Win Pool							
Favorite Position	Runners	Winners	Objective Probability	Subjective Probability	z-stat	Raw	Take & Breakage
1	11,365	4,126	36.30%	34.75%	-3.45	3.68%	-16.43%
2	11,371	2,425	21.33%	20.78%	-1.42	1.30%	-17.70%
3	11,367	1,622	14.27%	14.63%	1.10	-3.28%	-21.11%
4	11,362	1,137	10.01%	10.39%	1.36	-5.20%	-22.43%
5	11,340	787	6.94%	7.39%	1.88	-8.76%	-25.21%
6	11,063	533	4.82%	5.22%	1.99	-10.85%	-26.91%
7	10,086	331	3.28%	3.73%	2.50	-19.26%	-33.82%
8	8,226	209	2.54%	2.72%	1.01	-8.29%	-24.57%
9-14	10,095	191	1.89%	1.89%	-0.02	-7.80%	-24.19%
Place Pool							
Favorite Position	Runners	Placers	Objective Probability	Subjective Probability	z-stat	Raw	Take & Breakage
1	11,349	6,532	57.56%	54.56%	-6.46	11.81%	-8.16%
2	11,357	4,746	41.79%	40.27%	-3.27	5.23%	-15.94%
3	11,353	3,500	30.83%	31.45%	1.44	0.38%	-20.24%
4	11,348	2,656	23.41%	24.02%	1.54	-3.37%	-23.97%
5	11,330	1,998	17.63%	18.05%	1.16	-3.69%	-24.43%
6	11,051	1,361	12.32%	13.31%	3.18	-11.40%	-30.75%
7	10,085	922	9.14%	9.86%	2.51	-13.61%	-32.34%
8	8,226	547	6.65%	7.34%	2.52	-17.05%	-34.92%
9-14	10,095	461	4.57%	5.10%	2.57	-21.63%	-38.08%
Show Pool							
Favorite Position	Runners	Showers	Objective Probability	Subjective Probability	z-stat	Raw	Take & Breakage
1	11,300	7,954	70.39%	68.93%	-3.40	10.74%	-7.52%
2	11,308	6,462	57.15%	55.06%	-4.49	9.56%	-12.63%
3	11,304	5,418	47.93%	45.88%	-4.37	8.24%	-15.73%
4	11,299	4,305	38.10%	37.62%	-1.06	2.19%	-21.79%
5	11,309	3,541	31.31%	30.52%	-1.81	2.72%	-22.49%
6	11,039	2,539	23.00%	24.20%	2.99	-7.27%	-30.22%
7	10,080	1,786	17.72%	19.19%	3.87	-11.22%	-33.47%
8	8,223	1,046	12.72%	15.39%	7.25	-24.06%	-43.14%
9-14	10,092	907	8.99%	11.07%	7.33	-24.54%	-42.73%

Table 3: Breakdown of Wagering Pools by Favorite Position

Favorite Postion	Conditioned on Favorite Position % Win	% Place	% Show	Conditioned on Wager Type % Win	% Place	% Show
1	68.1%	21.5%	10.4%	34.7%	30.7%	29.6%
2	66.6%	22.9%	10.4%	20.8%	20.1%	18.7%
3	65.0%	23.8%	11.1%	14.6%	15.1%	14.4%
4	63.6%	24.5%	11.9%	10.4%	11.3%	11.2%
5	62.1%	25.0%	12.9%	7.4%	8.3%	8.8%
6	60.6%	25.3%	14.1%	5.2%	6.1%	6.8%
7	58.8%	25.7%	15.4%	3.7%	4.5%	5.4%
8	57.2%	25.8%	17.0%	2.7%	3.3%	4.3%
9u	56.2%	25.8%	18.0%	1.9%	2.3%	3.1%

Table 4: Comparisons of Studies of Betting Markets

			Intercept			Slope		
Author	Races	Years	Coefficient	St Error	t-stat	Coefficient	St Error	t-stat
Ali	20,247	1970-1974	0.014	0.004	3.72	0.887	0.023	-4.88
Asch et al	729	1978	0.008	0.003	2.75	0.873	0.018	-7.16
Busche & Hall	2,653	1981-1987	-0.002	0.003	-0.66	1.018	0.024	0.75
Gandar et al	10,332	1994-1997	0.002	0.001	1.34	0.991	0.012	-0.76
Sobel & Raines	2,799	1996-1997	-0.018	0.006	-3.20	1.143	0.041	3.50
Gramm & Owens	11,365	2002	0.006	0.002	3.67	0.948	0.011	-4.94

Table 5: Betting Application Data

Win Pool

Favorite Position	Runners	Winners	Objective Probability	Subjective Probability	z-stat	PT Subjective Probability	z-stat
1	203	72	35.47%	34.27%	-0.36	32.37%	-0.92
2	203	48	23.65%	21.42%	-0.75	20.83%	-0.94
3	204	28	13.73%	15.23%	0.62	15.25%	0.63
4	202	18	8.91%	10.50%	0.79	10.84%	0.96
5	203	17	8.37%	7.26%	-0.57	7.69%	-0.35
6	200	9	4.50%	4.98%	0.33	5.51%	0.69
7	184	6	3.26%	3.51%	0.19	4.03%	0.59
8	145	4	2.76%	2.45%	-0.23	2.88%	0.09
9-12	195	1	0.51%	1.52%	1.96	1.93%	2.76

Place Pool

Favorite Position	Runners	Placers	Objective Probability	Subjective Probability	z-stat	PT Subjective Probability	z-stat
1	203	121	59.61%	54.16%	-1.58	49.53%	-2.93
2	203	86	42.36%	39.82%	-0.73	38.45%	-1.13
3	204	70	34.31%	33.02%	-0.39	32.51%	-0.54
4	202	32	15.84%	23.68%	3.05	24.51%	3.37
5	203	42	20.69%	18.10%	-0.91	19.48%	-0.43
6	200	23	11.50%	13.27%	0.78	14.51%	1.34
7	184	18	9.78%	9.87%	0.04	11.25%	0.67
8	145	10	6.90%	6.77%	-0.06	7.98%	0.51
9-12	195	5	2.56%	4.50%	1.71	5.50%	2.60

Show Pool

Favorite Position	Runners	Showers	Objective Probability	Subjective Probability	z-stat	PT Subjective Probability	z-stat
1	203	148	72.91%	67.55%	-1.72	62.72%	-3.26
2	203	113	55.67%	52.61%	-0.88	49.83%	-1.67
3	204	101	49.51%	47.68%	-0.52	46.91%	-0.74
4	202	67	33.17%	38.50%	1.61	37.51%	1.31
5	203	73	35.96%	31.24%	-1.40	32.69%	-0.97
6	200	48	24.00%	24.35%	0.11	26.00%	0.66
7	184	33	17.93%	19.30%	0.48	21.81%	1.37
8	145	18	12.41%	15.11%	0.98	17.73%	1.94
9-12	195	9	4.62%	10.51%	3.92	12.75%	5.42

Figure 1: Estimated Returns